Post-pandemic landscape for central bank statistics

June 2023
Post-pandemic landscape for central bank statistics

IFC Bulletin No 58
June 2023

Proceedings of the 11th Biennial IFC Conference on “Post-pandemic landscape for central bank statistics”
BIS Basel, 25-26 August 2022

Overview
The post-pandemic new normal for central bank statistics

_Saira Jahangir-Abdoelrahman, Director Statistics Department, Central Bank of Suriname_

_Bruno Tissot, Head of Statistics and Research Support, BIS, and Head of the Secretariat of the Irving Fisher Committee on Central Bank Statistics (IFC)_

Opening speech
Measuring the past to better understand the present and chart the future: the central bank network on historical monetary and financial statistics

_Claudio Borio, Head of the BIS Monetary and Economic Department_

High-level panel speech
Time for structural change in central bank statistics? How to support the transition to a climate-friendly economy

_Claudia Buch, Vice President, Deutsche Bundesbank_

1. Addressing the impact of Covid-19
Wealth inequality and Covid-19 in the U.S.: evidence from the distributional financial accounts

_Mike Batty, Ella Deeken, Elizabeth Holmquist and Alice Henriques Volz, Board of Governors of the Federal Reserve System_

Sectoral financial positions and debt levels at the post-pandemic period

_Ahmet Tayyar Firat, Central Bank of the Republic of Türkiye_

Covid-19 crises – New rapid data for the Israel economy and making it accessible to all

_Daniel Rosenman, Bank of Israel_

The impact of central bank interventions on non-performing loans under Covid-19 pandemic – The United Kingdom and Brazilian Study Case

_Frederico Barros Diniz, University of Warwick, Warwick Business School, and João Paulo Vieira Costa, University of Brasilia, Brazil_

The effect of pandemic on Central Bank of Türkiye real sector database and post-pandemic adjustments
2. Coping with globalisation

Nationality vs residency approach: measuring the impact of MNEs production structure on corporate financial statements statistics
Klaus Gerstner and Susanne Walter, Deutsche Bundesbank, Javier Gonzalez, Santiago Martinez and Maria Garcia Riego, Bank of Spain, Florian Resch and Stefan Kischner, Austrian Central Bank, Alexandre Neves and Diogo Silva, Banco de Portugal

Challenges in the external statistics framework: how to register MNE financial restructuring operations
Nadia Accoto, Giuseppina Marocchi and Silvia Sabatini, Bank of Italy

Does firm size predict the residency status of the final investor? Evidence from Romanian FDI enterprises
Catalina-Florentina Pricope, National Bank of Romania

Who stands behind European FDI investors? A novel characterisation of pass-through within the EU
Carles Gómez-Llabrés, Fausto Pastoris and Martin Schmitz, ECB

A typology of captive financial institutions in Luxembourg: lessons from a new database
Gabriele Di Filippo and Frédéric Pierret, Central Bank of Luxembourg

3. Environmental statistics

Microdata base for sustainability indicators (ESG) developed at the Banco de España
Borja Fernández Rosillo San Isidro, Bank of Spain

Constructing forward-looking climate-related physical risk indicators – The use of private owned data for forward-looking aggregates at sector-country level for official statistics
Maurice Fehr and Elena Triebskorn, Deutsche Bundesbank, Jens Mehrhoff, IMF

Materiality of ESG factors in financial markets and financial statistics
Patrick Slovik and Farah Azman, Central Bank of the United Arab Emirates

An estimation of the carbon footprint in Spanish credit institutions’ business lending portfolio – Experimental statistics for credit institutions in Spain
Luis Ángel Maza, Bank of Spain

Measuring the emission profile of self-proclaimed sustainable exchange-traded funds
Hendrik Christian Doll, Maurice Fehr, Gabriela Alves Werb and Ece Yalcin-Roder, Deutsche Bundesbank
4. Fintech and digital assets

Sentiment analysis of user’s reviews on non-bank payment service apps
Muhammad Hafiruddin, Mohammad Khoyrul Hidayat, Arinda Dwi Okfantia and Nursidik Heru Praptono, Bank Indonesia

Fintech in statistical classifications: suggestions and tentative figures in a central bank context
Ulf von Kalckreuth and Norman Wilson, Deutsche Bundesbank, Celestino Giron and Urszula Kochanska, ECB, Enzo Buthiot and Yann Wicky, Bank of France, Luis Ángel Maza and Román Santos, Bank of Spain

Issues in reflecting digital assets in the Flow of Funds Accounts
Yoshiko Sato, Bank of Japan

Digitally-enhanced macroeconomic statistics manuals: the quest for methodological serviceability and compilation synergies
Celestino Girón, ECB

Measuring payment system policy credibility using machine learning
Muhammad Abdul Jabbar, Okiriza Wibisono and Alvin Andhika Zulen, Bank Indonesia

5. New (big data) analytics

A timely estimation of local investment trends using administrative data
Michele Loberto, Bank of Italy

Using the press as a real-time economic confidence indicator
Juan Pablo Cova M, Hugo Peralta V and Maria del Pilar Cruz N, Central Bank of Chile

Measuring macroprudential policy credibility using machine learning
Muhammad Abdul Jabbar, Nursidik Heru Praptono, Okiriza Wibisono and Alvin Andhika Zulen, Bank Indonesia

A probabilistic method for reconstructing the FDI network in search of ultimate hosting economies
Nadia Accoto, Valerio Astuti and Costanza Catalano, Bank of Italy

6. Financial microdata insights

Banks’ real estate exposures: Risk-based approach to measurement of exposures and concentrations
Patrick Slovik and Farah Azman, Central Bank of the United Arab Emirates

Examining concentration and similarity in institutional investors’ holdings
Ariel Mantzura, Bank of Israel

Statistical matching for anomaly detection in insurance assets granular reporting
Vittoria La Serra and Emiliano Svezia, Bank of Italy
Making omelettes without new eggs: a story about a self-sufficient way of producing new statistic  
*Miguel Fonseca and Sonia Mota, Banco de Portugal*

**7. New perspective for central banking statistics**

The euro short-term rate (€STR) – the new role of central bank statistics in financial markets – a financial benchmark fully based on statistical microdata  
*Ludovica Amorese, Javier Huerga and Ronald Rühmkorf, ECB*

Something old, something new: a reflexion on the new normal for statistical producers  
*Lígia Maria Nunes, Banco de Portugal*

Product level greenhouse gas contents – how to get there?  
*Ulf von Kalckreuth, Deutsche Bundesbank*

A characterisation of financial assets based on their cash-flow structure  
*Celestino Girón, ECB*

**8. Improving the quality of official statistics**

A decision-making rule to detect insufficient data quality – an application of statistical learning techniques to the non-performing loans banking data  
*Paolo Cimbali, Marco De Leonardis, Alessio Fiume, Barbara La Ganga, Luciana Meoli and Marco Orlandi, Bank of Italy*

Data quality management of entity group data: relevance and tools to address current challenges  
*Bruno Carreiras, ECB*

Big data analytics on payment system data for measuring household consumption in Indonesia  
*Renardi Ardiya Bimantoro, Mohammad Khoyrul Hidayat, Muhammad Abdul Jabbar and Alvin Andhika Zulen, Bank Indonesia*

Quantitative analysis of haircuts: evidence from the Japanese repo and securities lending markets  
*Kana Sasamoto and Kazuya Suzuki, Bank of Japan*

How can big data improve the quality of tourism statistics? The Bank of Italy’s experience in compiling the “travel” item of the Balance of Payments  
*Andrea Carboni, Costanza Catalano and Claudio Doria, Bank of Italy*

**9. Microdata disclosure control: a practical perspective**

Update on the work of the International Network for Exchanging Experiences on Statistical Handling of Granular Data (INEXDA)  
*Stefan Bender, Jannick Blaschke and Christian Hirsch, Deutsche Bundesbank*
Sharing researcher-generated code and value-added documentation in a trusted research environment

Louise Corti and Hannah Hodge-Waller, Office for National Statistics, UK

Joint secondary anonymisation of categorial and numerical variables in sensitive time series microdata – novel approach for Statistical Disclosure Control of a sensitive microdata set published in BELab data laboratory

Eugenia Koblents and Alberto Lorenzo Megia, Bank of Spain

Introduction to and application of SDC rules using self-developed tools

Jannick Blaschke, Matthias Gomolka, Christian Hirsch, Sebastian Seltmann and Harald Stahl, Deutsche Bundesbank
The post-pandemic new normal for central bank statistics

Saira Jahangir-Abdoelrahman and Bruno Tissot

Executive summary

Statisticians, including those in central banks, have quickly adapted to the consequences of the Covid-19 pandemic. But, three years after the start of the global crisis, can some normalisation be expected in the official statistical landscape, returning us to the situation prevailing before the crisis? Or, alternatively, should central banks be fundamentally rethinking the way they produce and consume data in a “new normal” state of the world?

The pandemic underlined that data producers have to provide more and more varied types of information to their users, not least by leveraging the wealth of sources available. This calls for the production of more high-frequency and timely indicators as well as for providing new types of indicator (eg sentiment indicators, policy credibility indices) on topics that are still insufficiently covered by existing official statistical frameworks. As key elements of national statistical systems (NSS), central banks have been at the forefront of global initiatives taken to address these issues.

The recent crisis was also a reminder that the statistical landscape has to permanently evolve, for instance to make the most of the opportunities provided by new big data analytics and to reach out to user groups more effectively. Central banks have been well positioned from this perspective, drawing on their unique role as producers and consumers of data to support their policies.

Moreover, recent developments have highlighted the urgency of dealing with a number of long-standing challenges. This calls in particular for the adaptation of statistical frameworks to better reflect structural developments in the global economy, especially as regards globalisation, digitalisation in finance and sustainability aspects such as climate change. Therefore, “central bank statistics have to change as well in order to support and reflect structural changes”, as argued by Claudia Buch, Vice-President of the Deutsche Bundesbank.

1 Respectively, Director Statistics Department, Central Bank of Suriname (sjahangir@cbvs.sr); and Head of Statistics and Research Support, BIS & Head of the Secretariat of the Irving Fischer Committee on Central Bank Statistics (IFC) (Bruno.Tissot@bis.org).

The views expressed here are those of the authors and do not necessarily reflect those of the Central Bank of Suriname, the BIS, the IFC or any of those institutions represented at the conference.

We thank Robert Kirchner, Ko Nakayama, Gloria Peña, Silke Stapel-Weber, Luis Teles Dias and Ulf von Kalckreuth for helpful comments and suggestions.
This suggests that central bank statisticians may be facing a quite heavy agenda to address both the new lessons of the pandemic and the longer-term challenges; in other words, they have to continue “measuring the past to better understand the present and chart the future”, as emphasised by Claudio Borio, Head of BIS Monetary and Economic Department, in his opening remarks. A number of supporting factors could be helpful in this direction.

A first one is innovation, as the new normal for central banks statisticians is likely to rely heavily on the use of data science to perform their traditional tasks. There is no such thing as a rigid, unchanging statistical framework, and statistical sources and tools have to be continuously refined to fit with evolving challenges – for instance to capture the long-term consequences of the pandemic, assess the sustainability of economic development, and track the global footprint of international firms.

Second, a clear message from the central banking community is that many of the challenges faced during recent crises could be addressed by making a better use of the large amount of micro-level information available in today’s modern societies. Here again, technological innovation can open up new perspectives, especially to facilitate the use of data that are available “organically” (i.e., independently of “designed” traditional statistical compilation exercises such as surveys and censuses) and to spur research using very granular but confidential data sets.

Third, the demand for timely, high-quality and varied statistical data is likely to remain strong, calling for statistical frameworks that can be adapted to meet evolving policy objectives and user needs. Several initiatives have been launched to make the global statistical infrastructure more flexible and efficient, for instance to develop global registers and identifiers and promote data-sharing and the access to new sources of information. Achieving rapid progress on these various fronts is a key priority.

Lastly, while central banks’ statisticians have proved responsive and innovative in the face of recent crises, further consolidating their contribution to NSS, their continued and close cooperation with other relevant stakeholders will remain essential to address the current complex challenges and further strengthen the role of official statistics in modern societies.

1. Introduction: a “new normal” for central bank statistics?

During the Covid-19 pandemic, official statistics had to quickly adapt to track new developments (Statistical Journal of the IAOS (2022)). As regards central banks, this global shock had two important immediate consequences (Tissot and De Beer (2021)). First, and reflecting their role as producers of official statistics, they had to proactively respond to the various disruptions observed in the provision of economic and financial data, in particular by making use of new information sources and analytical tools. Second, as users of data to conduct their policies, central banks faced new information needs; this reflected in particular the acute socioeconomic challenges brought about by the pandemic and the unprecedented policy measures taken in response and this needed to be carefully calibrated and monitored.

Following the pandemic, the global health situation appears to have stabilised. One view is that this would lead to some normalisation in the official statistical landscape, reverting to the pre-crisis situation. Another view, however, is that the pandemic has shed light on a number of important information challenges. From this
perspective, it has triggered a fundamental reassessment of the way central banks produce and consume data. The post-pandemic landscape for official statistics could thus be seen as a “new normal”, with significant distinctive features vis-à-vis the situation prevailing before 2020.

To shed light on these issues, the 11th conference of the Irving Fisher Committee on Central Bank Statistics (IFC) of the Bank for International Settlements (BIS) held in 2022 was devoted to the “Post-pandemic landscape for central bank statistics”. This event was an opportunity to take stock of the new perspectives facing central bank statisticians, especially as regards data production, users’ evolving information needs and ways to address them, as well as the challenges that remain.

More than 40 contributions were presented at the event, which was attended by almost 170 participants from 60 countries. These contributions, as referred to in this overview and included in this IFC Bulletin, highlighted the need for central banks to derive adequate long-term lessons from the pandemic, drawing from their dual perspective as producers of official statistics (Section 2) and as key users of economic and financial information (Section 3). They also highlighted the need to deal with a number of long-standing challenges, in particular the impact of globalisation, financial innovation and sustainability issues (Section 4). Fortunately, a number of ongoing projects in the central banking community are addressing these various points, leveraging technical innovation (Section 5) and the potential provided by the wealth of micro-level data available (Section 6). This calls in particular for further progress on the global statistical infrastructure (Section 7) and on liaising effectively with the various stakeholders involved (Section 8).

2. Longer-term impact of Covid-19: the producers’ perspective

Covid-19 had immediate, disruptive effects on the statistical production chain, leaving policymakers with acute information gaps in the short run. But the pandemic also underlined the longer-term importance for producers of official statistics to provide more and more varied types of data to their users, not least by leveraging the wealth of information sources available and big data-related tools. These elements can also help to address emerging issues that are still uncovered by existing official statistical frameworks, especially through the production of new types of statistics (Rosolia et al (2021)).

Making sense of the wealth of “organic” data available

Large and sudden disruptions at the height of the pandemic created several difficulties for the producers of official statistics, who were confronted with unavailable sources of information, methodological issues and data quality shortcomings. These challenges were first addressed in a pragmatic way to try to address the most pressing information gaps. But they were also a “wake-up call” for a general review of official statistical frameworks, not least to make a better use of existing data and sources and also of technological innovation in the longer run (Biancotti et al (2021)).

For instance, a number of new big data sources were explored by the Bank of Italy in the course of 2020 to replace missing statistics; this also proved an opportunity
to review the benefits of incorporating new data on web connectivity (eg mobile data), electronic payments (eg credit card transactions) and internet search data (eg Google Trends) to enhance the compilation of the “travel” item of the Balance of Payments (BoP). In particular, mobile phone data can be used to improve estimations on international travellers, as compared with more traditional data sources. Likewise, electronic payment records could be rather useful for data validation purposes, especially to deal with issues related to missing transactions, misclassification and insufficient granularity. Turning to Google Trends indicators, and despite the need to deal with potential breaks in time series when circumstances change, they can help improve the quality of statistics available on a complementary basis.

More fundamentally, there has been a widespread recognition among central bank statisticians that a wealth of data exist that had been too long neglected. Cases in point relate first to administrative data registers, ie those records created in an organic way as a by-product of government operations. Moreover, financial big data sources have multiplied in recent decades, reflecting the impact of digitalisation in today’s economies and the vast amount of commercial and financial operations undertaken every day. Lastly, large data collection efforts have been launched by national authorities and international organisations as part of the lessons drawn from the 2007–09 Great Financial Crisis (GFC). In particular, the Data Gaps Initiative (DGI) endorsed by the G20, and which has covered three phases so far, has been instrumental to closing the most pressing data gaps in financial markets (FSB and IMF (2009, 2015); IMF Staff et al (2023)).

As a result, central banks are working actively on making a better use of the wealth of information available, which is for instance a clear priority in Portugal. This means, first, that they are actively exploring non-conventional, still untapped sources of information. Second, they are leveraging technological innovation to improve the quality of official statistics being produced. In particular, the new statistical production processes can benefit from artificial intelligence (AI) techniques to enhance the accuracy, completeness and consistency of the reporting agents’ responses. One example relates to the treatment of regulatory requirements for banks’ non-performing loans, for which the Bank of Italy has developed a data quality management (DQM) rule based on past observations (eg number of outliers) to identify quality issues at the observation level automatically and promptly.

Providing more timely and high-frequency indicators

The Covid-19 pandemic also underlined the need for having more high-frequency and timely statistics, not least to assess today’s complex economic relationships, monitor evolving fragilities and keep policymakers informed at times of rapid changes and/or very large shocks. In general, traditional official statistics face important delays: to ensure accuracy and comparability, they are compiled on the basis of internationally accepted methodologies and standards and best practices, hence requiring relatively lengthy processes and verifications. In contrast, new types of “alternative” data can be compiled very quickly, not least by relying on innovative techniques, but often at a price in terms of the imperfect quality of the indicators produced.

While this accuracy/timeliness trade-off is not new for public statisticians, the pandemic has clearly changed views on how best to address it. Confronted with “statistical darkness”, central banks around the world have recognised the primacy of
providing timely indicators, for instance by mobilising alternative high-frequency data sources, constructing weekly or even daily indicators, and enhancing their nowcasting exercises (IFC (2021a)). From a longer-term perspective, they have also started a more fundamental review of their statistical function with a focus on data production chains. For example, the Central Bank of Türkiye has worked to increase the frequency of private institutions’ reporting, from quarterly to weekly, by combining balance sheet data with administrative records. Timeliness has also been improved by simplifying reporting formats. These enhancements have supported the creation of an in-house real time data platform as a monitoring tool of financial market conditions based on well-established processes for data streaming, integration, analysis and visualisation.

An important post-pandemic focus point has been to produce daily data. The value of such indicators is critical for both supporting public authorities’ monitoring tasks and providing faster information to society. For instance, the European Central Bank (ECB) has started to compile and disseminate daily data on the overnight interest rate for the euro (the €STR), drawing on the granular information already available in financial markets. This initiative illustrates the potential value of pre-existing statistical micro data in supporting market surveillance and monetary policy as well as in providing reference benchmarks for market participants. Another initiative by the Bank of Israel has been the setup of “new rapid” data sets by complementing various publicly available high-frequency indicators (eg developments in financial and forex markets, electricity consumption) with those available only for internal use (eg banking loans and mortgages in moratorium). The aim was also to make more data accessible to the wider public through interactive dashboards.

Producing new types of indicator

The Covid-19 pandemic also boosted the production of different types of statistics, which can be useful to address new economic and financial issues in the longer run. Central banks have been at the forefront of these initiatives, leveraging on alternative data sources and technical innovation. An obvious example relates to economic agents’ opinions, which are increasingly used for policy analysis. Significant efforts have been put on producing indicators on economic sentiment/confidence, based on text-mining tools to summarise qualitative information on agents’ perceptions or expectations in real time. While such sentiment indicators were traditionally computed from dedicated surveys of households or firms, the recent approaches have proved more flexible, easier to implement, and better suited for unexpected events – such as during the pandemic when surveys could not be conducted because of lockdowns.

For instance, the Central Bank of Chile has developed a high-frequency sentiment indicator through the application of textual analytics to capture the emotional tone of economic and opinion news, based on the computerised reading of printed media. The approach required the setup of an adequate dictionary to identify different types of emotion, the classification of the various textual inputs, and the benchmarking of the results vis-à-vis other, more traditional types of survey-based confidence indicator. The results have proved particularly useful in assessing the current state of the business cycle, the future evolution of economic conditions, and the degree of confidence among economic agents.

A key requirement for official data is to meet the test of practical utility, in the sense that they have to address the information needs of the government, the economy and the public at large. Hence an essential focus for central banks’ statisticians in the longer term is to remain relevant to users, including internal decision-makers in their own institutions.

Keeping central bank statistics relevant in the official statistical offering

Beside its short-term impact, the Covid-19 pandemic highlighted that the statistical landscape has to continuously evolve to reflect the changing economic and financial environment. This means that, to be fit for purpose, the official statistics offering needs to remain relevant in view of changing user needs. Among the many related challenges, statisticians need to be able to modify their output rapidly and flexibly; the statistical infrastructure needs to be adaptable and sufficiently granular to meet evolving requests; the trade-off between the costs of producing additional data and their benefits should be carefully assessed; and the purpose of compiling new statistics as a public good should be clearly understood by the public, which calls for strong data governance frameworks.

As regards the contribution of central banks more specifically, the pandemic showed that their data collected on the real economy and the financial system continue to be crucial to fostering monetary and financial stability and serving a wide range of users – eg policymakers, citizens, financial market participants, students and professors, as well as journalists. In that sense, the statistical responsibility of central banks has not changed fundamentally and continues to rely on a number of key elements. First, their data collection exercises must be clearly linked to their mandates. Second, the principles governing official statistics must be rigorously adhered to, not least to ensure that the correct information is used for policymaking. Third, and relatedly, is the recognition that high-quality data are key to supporting the credibility of policy decisions and in turn their effectiveness. Fourth, there must be a focus on securing public trust in this process, although this can be a challenge in the new era characterised by increasingly complex and diverse data sources.

Central bank experience shows that a number of initiatives are being pursued to further consolidate the strategic contribution that their statisticians can have to address users’ needs. One important focus area has been the broadening of the toolkits available to manage information, in particular by making the most of the opportunities provided by new big data analytics such as machine learning (ML) to support policy use cases (IFC (2022a)). Another is to strengthen their communication to various user groups, ideally based on both a multichannel and a segmented approach to reach out to the main stakeholders effectively (IFC (2023a)). Third, statistical methodological standards and products need to be constantly updated, for instance to capture more adequately the impact of digital innovation in finance, the issues posed by globalisation, and sustainable development aspects.

---

2 See UN (2013), especially its first principle: “(…) official statistics that meet the test of practical utility are to be compiled and made available on an impartial basis by official statistical agencies to honour citizens’ entitlement to public information”.
Addressing internal policy users’ needs

Central banks are themselves key consumers of data in conducting their policies. This was particularly obvious during the pandemic, when a wide range of policy measures were taken to mitigate the adverse impact of the crisis on the economy and society. Internal policymakers needed the right information at the right time to properly design, calibrate and monitor their actions (Tissot and De Beer (2021)). One example was Türkiye, where the initial response to the pandemic comprised interest rate cuts, liquidity provisioning, credit facilities and foreign currency swaps to alleviate the financial situation of households and firms. Another example was related to the various monetary and macroprudential measures taken to limit the impact of non-performing loans on the banking system so as to safeguard financial stability at the height of the pandemic, as was the case in Brazil and the United Kingdom.

One key consideration from this perspective was to be able to assess the impact of policy measures in a granular way, ie depending on the groups of economic agents of interest. For instance, the Board of Governors of the Federal Reserve System in the United States has developed “distributional financial accounts” that helped in the monitoring of the situation of different household groups through the pandemic. This analysis showed that, while reduced consumption and large stimulus packages had led to a significant increase in savings in general, important wealth effects were generated by asset price developments, resulting in significant dispersion across households.

Another important issue for central banks as policy institutions was to measure the credibility of the unprecedented measures taken during Covid-19. To address this point, one project at Bank Indonesia led to the construction of a macroprudential policy credibility index, drawing on ML and text-mining techniques based on media news. The aim was to assess the contribution of various factors (eg formulation of the measures, communication) in supporting the credibility of the actions taken to support financial stability. A similar approach was followed to assess authorities’ actions in Indonesian payment systems, representing another essential policy area.

Looking forward, central banks’ experience suggests that a number of critical aspects need to be considered to further enhance the contribution of public statisticians to evidenced-based policymaking. One is to secure the involvement of subject-matter experts in policy discussions. Second, careful prioritisation is essential, not least in view of resources constraints; this means that the statistical agenda should remain focused on economic and financial issues, the bread and butter of central banking. Third, international collaboration is key, primarily among central banks especially in the context of the IFC but also with other stakeholders eg international organisations, National Statistical Offices (NSOs), statistical associations such as the International Statistical Institute (ISI), and academia. Such cooperation can be instrumental to support the sharing of knowledge and experience as well as the global application of internationally recognised statistical standards. And, lastly, these efforts should be accompanied by a better use of the opportunities offered by micro data (Israël and Tissot (2021)).

4. Facing (not so new) challenges

In addition to addressing the longer-term lessons of the pandemic as regards the supply and demand of data, recent years have also highlighted the urgent need for
official statisticians to deal with a number of long-standing challenges. This calls in particular for central bank statistical frameworks to be adapted to better reflect structural developments in the global economy, especially as regards globalisation, digitalisation in finance, and sustainability aspects such as climate change.

Globalisation

An important challenge has been how to address the impact of globalisation on the official statistical framework. Obviously, it has become increasingly difficult to measure the footprint of multinational enterprises (MNEs) in domestic economies, due to the size of cross-border operations related to international supply chains, the complexity of corporate global ownership structures and the opaqueness of their restructuring strategies not least for tax optimisation purposes (Francois and Vicard (2023)). These elements have clearly complicated the work of official statisticians, especially in central banks, which are typically the main compilers of external sector statistics and are responsible for maintaining direct investment registers in many countries (IFC (2020a)).

Experience shows that the main difficulties faced in the compilation of external sector statistics often relate to MNEs’ four main types of data deficiency. First, the limited availability and timeliness of group-level sources data, since disclosed financial statements appear insufficient to support statistical compilation and need to be refined and validated with complementary data sets. Second, information on cross-border operations often varies between national compilers, reflecting multiple factors such as limited data-sharing and the lack of comprehensive and internationally agreed guidelines (eg for the treatment of mergers). This can lead to important inconsistencies between national BoP figures, leading to sizeable quality issues in external statistics (eg size of the errors and omissions item). Third, the treatment of special corporate restructuring cases has become increasingly complex, due for instance to the difficulties of identifying the location of corporate group activities within national borders and the role played by ad hoc units in global financial centres (eg Luxembourg), sometimes only on a temporary basis. Lastly, national data collection systems are relatively rigid, hampering statisticians’ ability to track MNEs’ evolving operations flexibly and quickly.

These limitations are clearly obvious for the compilers of foreign direct investment (FDI) statistics, especially because of the difficulty of correctly identifying MNEs’ chains of control. Complex corporate structures, with layers of intermediate holding companies, imply that the identity of the final investors is often masked. Reflecting these limitations, FDI data sets are usually compiled on the basis of the Immediate Investor Country (IIC) principle rather than the Ultimate Investing Country (UIC) one. This represents an important data gap, since the correct identification of the residence of the entities ultimately controlling FDI can be essential to understand intragroup economic relationships and improve the traceability of cross-border funding. It also complicates the interpretation of published FDI data, not least because of the inflated role played by intermediate financial centres in the global economy (Pogliani et al (2022)). Lastly, the above difficulties posed by globalisation can also have a huge impact on aggregated national accounts statistics (eg GDP, investment) especially in small open economies, as seen with the 2015 “Irish case” (OECD (2016)).
Digitalisation and financial intermediation

Another long-lasting challenge highlighted by the pandemic, in particular during the market turmoil observed in early 2020 (FSB (2020)), is the impact of innovation on financial intermediation and financial auxiliary functions, especially through the operation of non-bank entities. The so-called “fintech” topic, initially defined by the Financial Stability Board (FSB) as technologically enabled financial innovation that could result in new business models, applications, processes or products with an associated material effect on the provision of financial services (Carney (2017)), has indeed been on the agenda of the central banking community for many years. In particular, the fast-growing scope of fintech activities requires their adequate classification in statistical systems, not least to allow for their correct monitoring by public authorities. But a key obstacle is that there is currently no unique, internationally harmonised definition of fintech for statistical or other classification purposes. It is thus essential to adapt statistical frameworks to better address the challenges posed by digital innovation in finance through the correct identification of the related fintech providers and products (IFC (2020b)). International organisations have been actively working on these issues, not least in the context of the DGI as well as of the currently ongoing updating of international statistical standards – eg the System of National Accounts (SNA), the Balance of Payments Manual (BPM) and the International Standard Industrial Classification of All Economic Activities (ISIC); see Baer et al (2022).

Yet, three main issues stand out in this context. One is that a number of fintech firms (eg IT services providers) can be located outside the sector of financial firms, calling for more detailed classification in statistical systems and/or enhanced methodological guidance to solve the related identification problems. Another issue relates to the appropriate, feasible and consistent treatment in macroeconomic financial accounts of the digital assets used as a means of payment – eg central bank digital currencies (CBDCs), stablecoins and other types of cryptoassets. Lastly, an important concern for central banks, that are tasked in many jurisdictions with the mandate of promoting financial inclusion and supervising payment systems, is to monitor the use of new financial services by the population and the correct understanding of the risks involved.

Sustainable development

The pandemic has clearly reinforced the need to pay greater attention to sustainable development, with a particular focus on its environmental, social and governance (ESG) aspects. This calls for more internationally comparable statistics, based on clearly established standards and definitions to reliably inform investors, for instance to support the transition to a less carbon-intensive economy. But the current lack of international standardisation of ESG statistical frameworks and taxonomies can reduce the expected benefits of upcoming disclosure requirements, hampering the gathering of comprehensive information and the proper alignment of market incentives.

As regards central banks and supervisory authorities more specifically, the most pressing data needs appear to relate to their activities in the areas of financial stability, asset and reserve management and monetary policy, reflecting the importance of ESG issues for the solvency of financial institutions (IFC (2022b)). They have thus been at the forefront of efforts to develop sustainable finance statistics, focusing mainly on...
establishing analytical frameworks, designing sustainability indicators and actual monitoring (IFC (2021b)). In particular, the work of the Network on Greening the Financial System (NGFS) of central banks and financial supervisors has been instrumental to identify relevant information sources and address acute data needs (NGFS (2021)).

Despite these various efforts, important shortcomings remain to be addressed, especially as regards the availability and quality of granular-level ESG information. One view is that more than 100 indicators may be needed to fulfil existing information gaps. But, as emphasised by a recent study conducted by the Bank of Spain, available ESG micro-level data suffer from several limitations such as limited coverage of individual firms, heterogeneity in the standards and definition of indicators, lack of digitalisation, and verification and regulatory limitations. As a result, there is an obvious lack of granular firm/asset-level data to meaningfully measure the carbon footprint of economic and financial activities. Another data gap refers to forward-looking climate risk indicators (eg emission pathways), noting that most of the currently available ESG indicators are backward-looking. This is an important challenge, since forward-looking indicators would be key to facilitate tracking commitments towards a greener economy.3

5. Looking forward: leveraging innovation...

The above developments suggest that central bank statisticians may be facing a quite heavy agenda should they want to both address the new lessons of the pandemic as well as longer-term challenges. Fortunately, a number of factors will support their efforts.

An obvious one is innovation. The new normal for central banks statisticians is likely to rely heavily on the use of data science to perform their traditional tasks (Nymand-Andersen (2021)). The pandemic has clearly boosted interest in using “big data”, ie the multiple types of information sets often characterised by a high degree of volume, variety and velocity – the so-called “Vs” (Laney (2001)), although the big data sources of interest to central banks can be quite varied in practice, ranging from large traditional micro data sets to internet-based indicators and unstructured information such as text (IFC (2017)). Moreover, innovative analytical techniques, such as ML and other AI-based tools, can improve the efficiency and effectiveness of statistical work, by facilitating the gathering of more and better information and the dealing with complex and new complementary sources in a more automated way (IFC (2022a)). This can bring multiple benefits, in terms of granularity, flexibility, timeliness, accuracy and efficiency. For instance, it can support central bank statisticians in their quest to capture data in real time from existing sources, and help them construct innovative indicators to cover new areas of interest – see the recent project developed in Indonesia to gauge users’ satisfaction with applications for mobile payments services, which can effectively support financial inclusion.

---

3 See NGFS (2021): “Given the importance of forward-looking assessments of both physical and transition risks, the current reliance on mostly backward-looking data is unsatisfactory”. There is thus a “need to understand the point-in-time performance of an exposure against a transition pathway – hence the need for firms to disclose their transition plans – as well as the impact of adaptation and mitigation measures on the evolution of the risks.”
A first and immediate benefit is to strengthen the core statistical function of central banks. This can facilitate the dealing with the larger and more complex data sets that are in increasing demand to address the needs for more granular, timely and high-frequency information – not least to deal with events such as Covid-19. Moreover, several projects have highlighted the usefulness of ML tools for supporting DQM purposes in comparison with more traditional modelling approaches (Chakraborty et al (2017); Maddaloni et al (2022)). Better data quality can be instrumental to securing trust in the entire statistical production chain, from data collection to dissemination, and to strengthening central banks’ reputations not only as data producers but also as policymakers taking action based on the right information.

Indeed, the contribution of more and better data and advanced tools can effectively support central banks’ actions in multiple areas, such as macroeconomic analysis and forecasting, financial market monitoring and financial risk assessment (IFC (2019a)). As regards, for instance, the use of big data analytics to forecast macroeconomic indicators in near real time, a recent project pursued at the Bank of Italy aimed to use granular administrative data on motor vehicle registrations to “nowcast” business investment in transport equipment and hence capital goods. The approach also allowed for a distinctive approach across regions, which proved useful in tracking the impact of the pandemic. Another example is the construction by Bank Indonesia of a proxy for household consumption from retail payment system transaction data, based on text-mining techniques. This advanced indicator could be compiled within a few days after the end of the reference period and showed a high correlation with published (with significant lags) official household consumption numbers.

Yet the use of big data sources and tools is not without challenges. The new data sets may provide a false sense of accuracy, as they can present substantial hard-to-assess selection biases that limit their representativity of the whole population of interest (Mehrhoff (2019)). Similarly, AI tools trained on preclassified and/or inaccurate data sets may be of little help and raise the risk of “algorithm unfairness” (IFC (2021c)). Another risk is complexity, compounding the general literacy challenges faced by users of official statistics. This calls for the avoidance of “black box” messages and the development of innovative tools to inform policymakers and also researchers and the broader public. A final challenge is privacy protection and the respect of confidentiality settings, another key principle in official statistics.4 Respecting this principle can be challenging in practice when dealing with large data volumes with multiple variables, not least because of the risk of re-identification of anonymised data posed by the increased IT computing power available and the multiplication of cross-sectional and longitudinal databases that could be interconnected (IFC (2023b)).

Addressing the above issues puts a premium on adopting strong and comprehensive data governance frameworks (IFC (2021d)). It also calls for a clear recognition of the various trade-offs involved, not least given the resource constraints (e.g., budget, skilled staff, IT infrastructure) faced by public institutions. This means striking the right balance between costs and benefits when compiling data from new sources (for instance in term of the timeliness/accuracy trade-off), as compared with

---

4 See UN (2013), especially its sixth principle: “Individual data collected by statistical agencies for statistical compilation, whether they refer to natural or legal persons, are to be strictly confidential and used exclusively for statistical purposes.”
making better use of existing data. Lastly, users need to have a clear and comprehensive picture of the analytics provided, including the degree of uncertainty associated with the data and techniques used. Any shortcomings of the new big data and tools used should thus be communicated transparently to them, not least to highlight the degree of uncertainty surrounding the environment and hence policy decisions.

6. ... making a better use of micro data...

A clear message from the central banking community is that many of the challenges highlighted by the pandemic could be addressed by making better use of the large amount of micro-level information already available in today’s societies (IFC (2021e)). The growing supply of granular data reflects the impact of digitalisation (the digital footprints left by a large number of activities), the importance of government activities (e.g. the development of administrative data sets), and the statistical follow-up to the last GFC as well as the Covid-19 pandemic – with the launch of a number of coordinated initiatives such as the DGI to collect more granular economic and financial data to support decision-makers. Supply was also spurred by technology innovation, which has clearly facilitated the collecting, processing, validation, integration and analysis of large data sets. This has further boosted the use of existing microdata and the compilation of new statistical indicators as well as the application of innovative analytics. And, indeed, various central bank projects have highlighted the benefits of making a better use of the wealth of micro-level information, in particular to deal with globalisation, ESG issues, and financial stability.

As regards first globalisation, one initiative developed in the context of the European Committee of Central Balance Sheet Data Offices (ECCBSO) has highlighted the new possibilities offered by the integration of micro-level indicators on firms’ performance and group structures to generate nationality-based statistics (i.e. based on the country of control principle) that can clearly facilitate the assessment of the footprint of global groups outside their jurisdictions of residency (Tissot (2016)). Another focus has been on analysing the international chain relationship for FDI, which is often intermediated by financial centres. For instance, the Bank of Romania has used granular data from FDI registers to clarify the global network of direct investment linkages across different types of firm by industry and the impact of investment decisions made by their ultimate controlling parents. Similarly, an ECB study was conducted to assess who stands behind European FDI investors using company-level data on corporate ownership structure to determine the importance of non-EU corporations in ultimately controlling FDI activity, including its intra-EU component. Turning to the Bank of Italy, it has developed a methodology based on innovative techniques to identify ultimate hosting economies in bilateral FDI statistics, which is often a challenge due to lack of detailed data. The ECB has also developed a specific metric (the “significance multiplier”) to enhance the quality of the information on group structures contained in its granular Register of Institutions and Affiliates Data (RIAD). The approach has also helped to evaluate changes in corporate structures motivated by specific business decisions (e.g. mergers) on a high-frequency basis. Lastly, recent geopolitical uncertainties have raised questions about the pace and/or features of globalisation and highlighted the usefulness of micro data to better understand these developments. One question relates for instance to the possibility of “near-shoring”/“friend-shoring”, i.e. the reconfiguration of global value chains in favour of countries that are geographically closer (Panetta (2023)).
of analysis may be difficult to pursue based on macro data solely, and one solution can be to use non-traditional data sources instead, such as expatriate travel, land prices in key areas or transportation volumes.

Turning to the ESG area, a number of granular data sets exist that can shed useful light, as highlighted by the Bank of Spain’s project to estimate the carbon footprint of banks’ credit portfolios and track the evolving impact of their loans. Moreover, a key objective has been to properly analyse the consequences of sustainability issues on the financial system, with two types of need: first, to assess the impact of ESG factors on the risk profile and performance of specific institutions; second, to measure the contribution of sustainable finance – understood as the process of taking ESG considerations into account when making financing decisions with the aim of fostering long-term investment in sustainable economic activities (IFC (2022b)). Indeed, as highlighted by a recent review by the Central Bank of the United Arab Emirates (UAE), a growing number of institutions already integrate ESG criteria into their investment strategies. The use of green finance instruments is in fact likely to increase further, not least in view of the development of the Principles for Responsible Investment by an UN-supported international group of institutional investors. Similarly, the Deutsche Bundesbank has analysed the investment strategy of self-proclaimed ESG exchange-traded funds (ETFs) using company-level data to analyse the impact of their financing; in this context, accessing granular information proved particularly important to cope with the diversity of sustainability objectives pursued by the various types of fund. A parallel approach allowed for the extraction of relevant forward-looking risk indicators at a sectoral and country level, using existing micro data available from a private provider that helped to overcome the challenges posed by, in particular, limited comparability across data sources and over time.

A third key contribution of micro data to support central bank policies is to shed light on the functioning of the financial markets especially as regards idiosyncratic fragilities and contagion effects. For instance, the UAE Central Bank has developed a holistic risk-based approach to assess banks’ real estate sector exposures and concentration risks. Similarly, the Bank of Israel has worked to identify potential vulnerabilities based on a granular data set of assets’ holdings of institutional investors, which allowed the analysis of diversification effects, systemic factors, and the influence of initial conditions. More generally, granular reports from supervised institutions can facilitate the monitoring of specific markets that play an important role in financial stability. For instance, the Bank of Japan has utilised granular transaction data to analyse the market structure and haircuts for repo and securities lending markets. This type of information is highly valuable for authorities willing to take adequate measures preventively and/or in reaction to specific events. It also puts a premium on the automated detection of reporting errors, as done for instance at the Bank of Italy for insurance companies’ reports through ML techniques – underscoring again the value of advanced analytical tools to support micro-DQM processes efficiently, in particular when human resources are limited.

5 The related paper from the Bank of Italy on “Statistical matching for anomaly detection in insurance assets granular reporting” (La Serra and Svezia (2022)) received the IFC award for the best paper presented at the conference by a young statistician.
7... making progress in statistical infrastructure...

The central bank community is supporting initiatives to improve general aspects related to the global statistical infrastructure (e.g., registers, identifiers and statistical standards). One example relates to the Legal Entity Identifier (LEI), a 20-digit reference code to uniquely identify legally distinct entities that engage in financial transactions, which has already become mandatory in several jurisdictions. It has been argued that this standardised and unique entity identification played an important role during the Covid-19 crisis, for instance in the United Kingdom by addressing (i) the difficulty of verification of merchants on online platforms and (ii) the drain on public sources due to fraudulent applications for public funds that were designed to support firms (GLEIF (2020)).

An important focus has been on the methodologies supporting official statistics compilation, since central banks’ data production has to take place in the general context of globally standardised methodologies. This requires continued international collaboration to support the harmonisation of standards and the exchange of learning experiences, not least within the central bank community but also with relevant stakeholders at national (e.g., NSOs) and international levels (e.g., OECD, IMF and the UN). A key goal of the IFC is thus to promote knowledge-sharing and international cooperation on statistics-related methodologies, initiatives and training (see IFC (2023c)). Another important dimension is to further develop statistical standards and support related ongoing initiatives—especially in the context of the new DGI and of the revision of international statistical manuals. This calls also for improvements in the usability of these manuals, whose lifecycles often run over several decades, for instance by developing digital versions that can facilitate navigation across different domains and improve methodological consistency (e.g., reduction of overlaps and inconsistencies). This would also allow for flexibly adding methodological extensions as new, specific statistical issues arise.

One particular topic of interest to central banks in this context relates to globalisation, with the current updating of the IMF Balance of Payments and International Investment Position manual (BPM6), which aims in particular to improve the statistical treatment of corporate restructuring operations in external statistics and to favour the reduction of bilateral asymmetries. Another key domain relates to the fintech universe, which combines novel types of financial services with more traditional ones performed in novel ways and activities that are not financial. This calls for a proper classification of the various types of activity involved (e.g., cryptoassets, banktech, insurtech, tech facilitators/infrastructure providers).

Some challenges may also require a more fundamental overhaul of conceptual approaches. A case in point relates to the residency principle (country of location) underlying national statistics (IAG (2015)). Since MNEs’ operations cannot be captured comprehensively by isolated, country-based statistical systems, nationality-based statistics could be developed as a useful complement. This approach has been already followed for the compilation of a number of BIS statistical products, such as the international debt security statistics (IDS) and the international banking statistics (IBS). But it has to be reinforced by a common statistical infrastructure supporting the cross-border sharing of firm-level information, as already organised in the context of financial institutions under the BIS umbrella (Tissot (2017)). Fortunately, similar initiatives are also under way as regards information on non-financial corporates, especially by the OECD for MNEs activities (AMNE Database) and Trade in Value Added (TiVA Database) as well as by Eurostat for statistics on the types of enterprise...
engaged in international services – the Services Trade by Enterprise Characteristics (STEC) experimental statistics, produced by linking international trade in services micro-data with the business register at enterprise level.

8... and fostering cooperation

The complex and various challenges facing official statisticians suggest that central banks cannot act in isolation and should continue to pursue their cooperation with relevant stakeholders.

Reaching out to all relevant stakeholders

A key group of interest to central banks statisticians are their counterparts in national statistical systems (NSS), especially NSOs, as well as the statistical departments of international organisations. Their unique expertise is instrumental for developing data standards and new methodologies. Moreover, good cooperation helps reduce overlaps and maximise synergies, an important issue for public authorities facing resource constraints. This collaboration can take various forms, from the exchange of experience and technical assistance to the setup of digital platforms for information-sharing, including statistical repositories of models (where to find data and metadata). For instance, the NGFS has emphasised the importance of setting up a dedicated repository of climate data to support financial sector stakeholders to share best practices in identifying related data need, sources and gaps. Similarly, the IFC has actively promoted the exchange of experience in the official statistical community, building on its broad, global membership of central banks and its position as an affiliated member of the ISI.

Another important group of stakeholders for central banks relates to the private data providers that are playing an increasingly role in the current search for alternative data sets, especially for the purpose of closing existing gaps in official statistics. One essential area relates to ESG issues, for which a good understanding of both the information market and the interests of alternative data providers is essential. Establishing partnerships can be the best way to access these data in a continuous way, design related regulations and standards to prevent monopolies, and properly incentivise the private sector to develop, create and deliver novel (technical) solutions for the provision of new indicators. For their part, public statisticians have to improve the useability of these complementary data so that they can be incorporated into the official statistical framework in an agile and timely manner. This can be achieved, for instance, by embedding public statisticians in external data providers so that they can better understand the management and supply of their alternative indicators.

A third important target group are the users of statistical information. Central banks, like other compilers of official statistics, are well aware that communication is a core statistical task to ensure that the data are fit for purpose and that their value is maximised. This puts a premium on developing statistical literacy in the population, ensuring that users understand and accept what statisticians are doing, and securing public trust in the overall official statistical framework.
Facilitating data access for research purposes

As a key user group targeted by central banks, academic researchers study internal granular data sets for scientific and analytical purposes – a demand that has been growing rapidly in recent years. However, accessing such detailed information can be quite sensitive because of the confidentiality and privacy issues involved. The rules and practices for protecting confidentiality have thus to be revisited to make micro data more accessible. Two key workstreams are involved here: first, the clarification of the procedures defining how authorised personnel can use the data; this aspect has become a key element of central banks’ governance frameworks. Second, the development of specific processes for data dissemination, including research code-sharing, remote execution of computation exercises based on confidential data, and in-house dedicated data research centres.

To shed light on these issues, several central banks with a number of NSOs and international organisations have been involved in INEXDA, the International Network for Exchanging Experience on Statistical Handling of Granular Data (IFC (2019b)). Its work comprises a stocktaking of the granular data sets available across countries, the review of best practices for granting access to open software solutions and data and metadata, and the identification of common features across jurisdictions with a view to the potential harmonisation of data access procedures. A key focus has been on the exchange of experience on the accessibility of micro administrative data and related data protection techniques, with the development of machine-readable description of access procedures, the review of the technical, administrative and organisational features for accessing granular data for research purposes, and the promotion of a metadata schema to provide metadata for microdata on the data set level (Bender et al (2019)). This schema is based on (i) the data set aspects of access procedures (i.e. information collected to unambiguously identify a data set, its origins, and its access modes); (ii) information regarding the aggregation of granular data, especially when it involves combining different data sets; and (iii) information on users of this information and related projects.

The INEXDA initiative has highlighted a number of important features related to the governance of data access for research purposes. First, output control has typically become mandatory for non-publicly available data, and is usually performed by researchers and subsequently checked by the data providers. Second, the often preferred access mode is on-site access, when researchers have to work in a secure environment based in the premises of the institution providing the data. Third, access modes can differ depending on the available data sets or the degree of anonymisation required: in particular, while formal anonymisation appears to be the most commonly used method, there can be various approaches governing data access for scientific research, with the involvement of different decision bodies in these processes (supported by specific confidentiality agreements or memorandums of understanding) and based on various techniques – e.g. remote computation,

6 One typically distinguishes between the risks of “direct disclosure” (i.e. disclosure of information with identifiers) and “statistical disclosure” (i.e. the risk of re-identification of anonymised data, a risk that is growing with the increased IT computing power available and the multiplication of cross-sectional and longitudinal databases that could be interlinked); see National Research Council (2005).
identity masking tools, use of so-called public use files (PUFs) that consist of micro-
level records prepared in such a way that individual entities cannot be identified.7

Such knowledge-sharing exercises have highlighted the key role of data
protection procedures to protect confidentiality (UNECE (2019)). One example is the
Statistical Disclosure Control (SDC) method developed at the Bank of Spain’s BE\Lab
data laboratory to ensure the anonymisation of sensitive time series microdata such
as loan-level information. The approach is based on an extensive use of big data
analytics and aims to mitigate the risk of disclosing the identity of specific borrowers
resulting from the combination of multivariate information, since debtors may have
multiple loans and loans may have multiple debtors (hence raising the risk of debtor
re-identification even when the data set has been anonymised).

Another important area relates to code-sharing, which is increasingly considered
a best practice in empirical scientific research, not least for transparency and
accountability reasons. In particular, a growing number of journals employ data
editors to re-run code and validate results of published articles. The implication is
that authors have to provide all the data, code, and processing instructions necessary
to exactly replicate the analysis and obtain identical results. But, as highlighted by a
recent pilot project undertaken for the UK NSO Secure Research Service, the wider
sharing of code outside a closed researcher environment requires the safeguarding
of statistical confidentiality, the validation of its quality, and adequate disclaimers to
protect the reputation of data owners’ and publishers. Moreover, the reproducibility
of research results obtained in a trusted environment is not without challenges, as it
requires the review of the related data access rights, the consideration of supporting
auxiliary material that has not been disclosed, and the review of the code to be shared
outside the secure environment.

A final lesson is that preparing research-ready data can require a significant level
of resources. These are required to deal with large (numbers of) data sets, complex
data structure (eg multiple identities), and heterogeneous legal frameworks (eg when
specific rules depend on the data considered). Hence, the onboarding of external
researchers in internal research data centres (RDCs) can be cumbersome and
resource-intensive. This puts a premium on developing tools to support the pre-
selection of the information that can be shared in a more automated way, an
approach currently under consideration at the Deutsche Bundesbank.

7 Noting that there are many ways in practice to protect data confidentiality (IFC (2023b)).
References


Opening speech

Measuring the past to better understand the present and chart the future: the central bank network on historical monetary and financial statistics

Claudio Borio,
Head of Monetary and Economic Department,
Bank for International Settlements

---

1 This speech was prepared for the conference. The views expressed are those of the author and do not necessarily reflect the views of the BIS, the IFC or the central banks and other institutions represented at the event.
Measuring the past to better understand the present and chart the future: the central bank network on historical monetary and financial statistics

Claudio Borio
Head of the BIS Monetary and Economic Department

Remarks at the 11th Biennial IFC Conference
Basel, 25 August 2022

It is a pleasure to welcome you all to the Biennial Conference of the Irving Fisher Committee (IFC). This is already the 11th in the series, all hosted by the BIS. I am especially glad to see that former IFC Chairs, the current Chair and the Chair-to-be are participating. This is a clear sign of the importance of the event. I am sure that, as in past, the conference will provide plenty of food for thought to better inform our discussions and help implement the IFC’s agenda. Last but not least, it is great that, finally, we can all meet in person.

The programme and the IFC’s agenda are naturally concerned with the challenges of the day and with the statistics that will help policymakers address them. Today, however, I would like to stand back and reflect on how the past can inform the present and on the role that statistics can play in that context.

This will also allow me to introduce the “new kid on the block” – the latest addition to the IFC’s “enlarged” family. This is the central bank network on historical monetary and financial statistics (HMFS), which brings together central bank statisticians and academics. I am particularly excited about this development.

I will address three questions. Why do we need historical economic statistics? What is the network about? And what has been its most concrete output to date? Consider this an open invitation to get more involved in the network’s activities.

The value of historical economic statistics

They say that history is to society what memory is to an individual. Memory is what provides continuity to us as individuals. One could even say that it is the basis of consciousness, as we can only exist in time. It helps define who we are.

1 The views expressed are my own and not necessarily those of the BIS.
Of course, history has a connotation of “distant memory”. But that, too, helps define who we are and determine what we do. Moreover, the dividing line between “history” and “recent past” – or “yesterday” – is fuzzy and context-dependent. Often, the line is arbitrarily drawn at the point when our own experience begins. Introspection comes in handy here. When I was a teenager, I thought of the 1920s as distant history; today, I think of the 1970s as yesterday. And yet, they are both separated by half a century.

What is true of history in general is also true of economic history, be it the history of economic thought or the history of events. And all study of economic history must be based on statistics, ie the “facts” or data points that inform our interpretation of what happened. Hence their critical importance in understanding the past and in drawing lessons for the present.

Although I did not have much exposure to economic history in my years at university, I came to embrace it during my professional life. Personally, I found it essential to shed light on the preoccupations of current policymakers. The list of issues, as reflected in my work, is not a short one. It includes issues such as the hidden perils of the so-called Great Moderation; the costs or, in fact, non-costs of deflation; the usefulness of the concept of the natural interest rate; the great power but also great limitations of monetary policy; the waxing and waning of central bank independence; and, more generally, how policy regimes shape, and are shaped by, the economic environment in an interaction that can spring challenges from unsuspected quarters – the Great Financial Crisis of 2008–09 being the most notable example.

It goes without saying, reading history correctly is tricky. Drawing lessons for today requires identifying what can and cannot be inferred given the difference in context. In turn, this calls for knowledge that goes beyond narrow economic understanding and a degree of imagination, to avoid projecting onto the past today’s intellectual baggage and vice versa. This is also true for the statisticians that develop the raw material that economists work with. I will get back to this in a minute.

The HMFS network

What, then, is the HMFS network? It is an informal group that brings together central bank statisticians, economists with a strong interest in statistics and academics to exchange views and share their experience in the development and use of historical monetary and financial statistics. The objective is to help develop those statistics and to stimulate their production more broadly. The group is very much a “coalition of the willing”. At present it involves 10 central banks and 2 academics.

Importantly, the focus of the group is not on collection, but on design and production. Participants have been brought together by the recognition that production of historical statistics is both hard and a public good. As a result, there is a strong disincentive to create them. Participants also recognise that a robust transnational methodology is necessary to guide the production of high-quality statistics. Hence the core concern with methodology and the aspiration to delineate standards of good practice. In the absence of such guidance, it is all too easy for compilation of national statistics into panels to involve series that are not only imperfect, but like
Tolstoy’s unhappy families, which are all unhappy in their own way, to all be imperfect in their own way.

The group has two guiding principles. One is the importance of comparability of statistics across countries and time. The other is the importance of transparency in how those statistics are produced, in “how the sausage is made”, as it were. Transparency is essential to address the major obstacles and pitfalls involved as well as to give rise to an “open-ended” process. This is how knowledge is transferred and statistics can be improved over time.

Transparency is of the essence to improve not just production, but also consumption. All too often economists, as main consumers, take statistics at face value. Sometimes I have been guilty of this sin myself! It is important that statisticians do not tire of raising awareness. This is true not only of historical statistics, but of current ones as well!

As the group notes, producing high-quality historical statistics requires “statistorians” – professionals who combine technical statistical know-how with an understanding of history. This allows them to place the original statistics in their proper institutional data-generating context and hence to understand sources and the sources’ limitations. We need more of them. Hopefully, the efforts of the network will stimulate their emergence.

The first report

The group first met at the BIS in October 2016 and, subsequently, at roughly yearly intervals, interrupted by the pandemic. I guess you will be asking yourselves: “What about any concrete output?”.

The first visible output is a report or monograph to be released soon entitled “Historical monetary and financial statistics for policymakers: towards a unified framework”. As with any first child, its birth has proved challenging but also very rewarding!

The report does three things: it provides context on the history and purpose of the group; it lays out the key methodological principles; and it applies them to the construction of interest rate, credit and real estate price statistics. The principles are then illustrated concretely by the statistical series put together by participating central banks. These include those of the United States, Japan, France, the United Kingdom, Italy, Canada, Austria, Sweden, Norway and Denmark.

Why the choice of those economic variables to start with? Three reasons.

First, they have come to prominence in policymaking owing to the historical re-emergence of major financial cycles. After being dormant for several decades, these cycles have been at the heart of business cycles since the mid-1980s. But they had also been common in the late 19th century all the way to the Great Depression. There are clearly lessons to be learnt.

Second, statistics on credit and real estate prices are comparatively scarce and, surprisingly, rather poorly understood (we were struck by this). And those that do exist have significant shortcomings.
Third, from a methodological perspective, the three series shed light on different issues. One is the deceptive simplicity of the construction of (short- and long-term) interest rate series (a financial price). For example, the construction of benchmark interest rates involves sometimes subtle questions concerning the structure and operation of markets, the specific nature of contracts as well as those of pricing practices and conventions. Another issue is the huge complications that hinder the production of consistent credit aggregates (a financial quantity) – a financial variable that had been neglected until recently in favour of its close cousin, monetary aggregates. Yet another issue is the complexity of aggregating into an index highly heterogenous assets (real estate), which can have first-order effects on the corresponding series.

What are the key takeaways of the report? Many! I have already mentioned some and would strongly encourage you to read the whole study, which is one of a kind. But in the time available, let me mention two. One is conceptual and often overlooked. The other is empirical and largely novel.

The conceptual one is that, fundamentally, building historical statistics requires dealing with synthetic countries and synthetic objects. Synthetic countries, because the borders of nation states have been in flux. This has important and often neglected implications for the interpretation of the statistics. Synthetic objects, because the same term can be applied to highly different variables. Just think of how much what constitutes the right “policy interest rate” varies across countries and has changed over time! Both issues require careful treatment.

The empirical takeaway concerns what one might call “missing credit”. This largely reflects a focus on regulated institutions and, more specifically, banks – sometimes only a subset of them – in the construction of the statistics, an instance of “look under the lamppost” syndrome. The most widely used historical credit statistics miss large chunks of credit. What today would be termed “shadow banking” was typically big, and so was “peer-to-peer” lending in several countries, in the form of mortgage credit often intermediated by notaries. Not such new phenomena after all! As a result, our understanding of the degree of “financial deepening” or of the information content of credit aggregates for financial crises has probably been distorted, across both countries and time, despite efforts to overcome the drawbacks in the data.

**Conclusion**

Let me stop here.

I hope I have encouraged you to reflect more on the value of historical statistics and of their careful construction. If something is a public good, as the production of these statistics is, it is worth investing in it. Critically, the importance of a transparent approach cannot be emphasised enough, as a means to bring to light the statistics’ strengths and limitations, to allow for their improvement over time and to facilitate their proper interpretation and use.

The message of the network is both sobering and optimistic. It is sobering, because it offers a vigorous discussion of the limitations of existing measures of our macro-financial history. It is
optimistic, because while the gaps and imperfections are extensive, they can be overcome. The network offers a way forward.

You can see how the IFC having the HMFS network operate under its aegis – “adopting” this young kid, as it were – is mutually beneficial. It offers the network a welcome and cosy new home, and it offers the IFC a new vehicle to develop policy-oriented statistics and to nurture links with academia. We are grateful for your decision.

I very much hope, too, that my remarks have whetted your appetite for the work of the network. Consider this an open invitation to become more involved in its activities!
Wealth inequality and Covid-19 in the U.S.: evidence from the distributional financial accounts

Mike Batty, Ella Deeken, Elizabeth Holmquist and Alice Henriques Volz,
Board of Governors of the Federal Reserve System

---

1 This presentation was prepared for the conference. The views expressed are those of the authors and do not necessarily reflect the views of the BIS, the IFC or the central banks and other institutions represented at the event.
Wealth Inequality and COVID-19 in the U.S.

Evidence from the Distributional Financial Accounts

Mike Batty, Ella Deeken, Elizabeth Holmquist and Alice Henriques Volz

Abstract

There have been many questions about how different U.S. household groups have fared economically through the severe disruption of the COVID-19 pandemic. Using the Distributional Financial Accounts (DFAs), we consider how household wealth has evolved in this turbulent time. However, the unprecedented events of the pandemic give reason to question the relevance of the historical relationships used for the last two years in the DFAs, particularly due to the vast amount of fiscal support and reduced consumption that resulted in significant excess savings, particularly in the early quarters of the pandemic. We examine a few alternate scenarios for how the increase in deposits may be allocated, providing plausible bounds for the household wealth distribution for recent quarters. We find that the effects of alternative distributions of deposits on the overall wealth distribution are small but modestly alter the estimates of changes in wealth experienced by less affluent households.

Keywords: Wealth Distribution, Wealth, Pandemic

JEL classification: D31

Contents

Introduction ............................................................................................................................................... 2
Aggregate Household Wealth during COVID-19 ............................................................................... 2
DFA Results during COVID-19 ............................................................................................................ 4
Alternative Distributions of Excess Savings ................................................................................... 8
Conclusion ................................................................................................................................................ 11
References ................................................................................................................................................ 12

1 This is an update to the FEDS Note “Wealth Inequality and COVID-19: Evidence from the Distributional Financial Accounts” by Batty, Deeken, and Volz published August 30, 2021 by the Board of Governors of the Federal Reserve System, https://doi.org/10.17016/2380-7172.2980. Special thanks to Harrison Karp for assistance with data and exhibit preparation.
Introduction

The COVID-19 pandemic severely disrupted a wide range of economic activity, while leaving some economic activities remarkably resilient despite extraordinary restrictions placed on our physical interactions. The unique pattern of income losses, early spending reductions and substitutions, and government relief have raised many questions about how different household groups fared economically. This note discusses how household wealth may have evolved over the COVID-19 pandemic. We first describe how aggregate household wealth, as reported in the Financial Accounts of the United States (FA), changed since the pandemic began. Next, we discuss how these changes in wealth are apportioned across the distribution in the Distributional Financial Accounts (DFA), which use historical relationships between macroeconomic aggregates and micro survey distributions to extrapolate from the distribution of wealth measured by the 2019 Survey of Consumer Finances. However, the unprecedented events of the pandemic give reason to question the relevance of these historical relationships in this period. The DFA projections do not explicitly take the unique circumstances into account; for example, they do not consider the vast amount of fiscal support and reduced consumption that has resulted in a significant amount of excess savings, particularly in the earliest quarters of the pandemic. Since we do not directly observe the distribution of savings after the 2019 SCF, we estimate the amount of excess savings that occurred and present a few alternate scenarios for its distribution, thus producing plausible bounds for the distribution of household wealth in the quarters since.2

Aggregate Household Wealth during COVID-19

Despite the fall in equity markets that drove a sharp decline in household wealth in 2020:Q1, equity prices rebounded quickly after the Federal Reserve, U.S. Treasury, and Congress took steps to stabilize financial markets and the economy, and house prices began to increase rapidly. Even though 2022 has seen equity markets retrace some of their prior gains, as of 2022:Q1, household wealth has increased $31 trillion since the beginning of 2020. This 28% gain was driven by asset accumulation much more than by debt paydown.3 Further, asset-price increases ("revaluations") were the dominant source of wealth accumulation, accounting for nearly 80% (Figure 1). Equity performance was strong through 2021:Q4, following the pandemic-related crash in 2020:Q1, producing a net gain of 47% for the Dow Jones U.S. Total Stock Market Index over the seven quarters; however, global tensions, inflation concerns, and rising interest rates reversed some of these gains in 2022:Q1. Real estate valuations have also risen strongly since the beginning of 2020, which contributed 93% of the 32% increase in household real estate holdings, a trend that continued into 2022:Q1.4

---

2 Excess savings may have elevated balances in other assets, but the evidence thus far suggests that vast majority of excess savings by households resides in deposit accounts.

3 Although there was some reduction in household credit card debt through 2020, mortgage debt did not decrease, likely due to strong refinance activity and rising house prices (Barnes et al, 2022).

4 The valuation model is based upon sales data from Zillow. For more details, see, “A New Measure of Housing Wealth in the Financial Accounts of the United States” by Hall, Nielsen, and Sommer (2018).
Although not the primary contributor to the aggregate gain in wealth, savings ("net transactions") also surged during the pandemic.\(^5\) This increase in savings reflects the resilience of certain sectors of the economy, reduced consumption, and the large amount of fiscal relief and debt forbearance provided by the government.\(^6\) Notably, increased household savings largely flowed to deposits, which saw unprecedented gains (Figure 2). In contrast, net transactions for other assets and liabilities were within their recent historical ranges through 2020 and beyond. Equities and real estate had somewhat elevated transactions in 2021 but remained far below those for deposits.\(^7\)

---

\(^5\) More formally, savings here can be defined as the net acquisition of assets less the net incurrence of liabilities.

\(^6\) The CARES Act provided fiscal support in 2020, and the Omnibus Appropriations and Coronavirus Relief Package and the American Rescue Plan provided support in 2021. The Paycheck Protection Program which provided forgivable loans to eligible businesses and forbearance programs that allowed the borrowers to pause on payment on federally backed mortgages and student loans also supported increased savings for some households.

\(^7\) Transactions for debt securities are calculated as a relatively small residual (roughly ten percent) from the total issuance and holdings of other sectors. Thus, modest measurement error in these other sectors can have a significant effect on households, which likely explains the heightened volatility of that series.
DFA Results during COVID-19

The aggregate changes in household wealth provide a useful lens through which to analyze the DFA estimates during COVID-19.\(^8\) The net transactions for assets that account for almost all the wealth gained through market price increases—corporate equities, mutual funds, pensions, and real estate—have not been elevated through most of the pandemic, and they were significantly lower than their respective net revaluations. Thus, the distributions for these assets measured in the 2019 SCF are likely still informative, and there is little reason to expect the DFA apportionment of the price-driven gains since 2020:Q1 would be substantially biased.\(^9\) In contrast, the DFA models have no relevant historical precedent for the large inflows of deposits that were influenced by pandemic-specific factors. In this section, we describe the recent evolution of the wealth distribution as estimated by the DFA and turn to alternate distributions of deposits in the next section.

We begin by reviewing the balance sheet composition across the distribution. In 2019:Q4, approximately half of the assets held by households in the bottom half of the wealth distribution ("Bottom 50") were real estate, with pensions and durables

---


\(^9\) Even though the amounts of transactions for these assets were affected by the pandemic much less than those for deposits, it is possible that their distributions were atypical. However, given that these transactions were small compared to both the levels and the revaluations for these assets, the effect on their overall distributions would be muted. Note, the Financial Accounts does not show intra-sector transactions, and thus would not reveal if the pandemic caused a large increase in asset sales between households. More granular transactions data suggest this likely does not materially bias the DFA. According to the National Association of Realtors, existing home sales increased 5.6% in 2020, which is less than the increase in real estate transactions in the Financial Accounts. Similarly, trading volumes on the NYSE increased in 2020 over 2019 by a smaller amount than the increase in net transactions of household’s corporate equities and mutual funds in the Financial Accounts.
each comprising less than 20%. Liquid assets were less than 10% of Bottom 50 assets. In contrast, for those in the top one percent (“Top 1”) and families between the 90th and 99th percentile of the wealth distribution (“Next 9”), real estate comprises less than 20%, while equities and mutual fund shares are nearly one third of asset portfolios, rising to nearly one half at the top of the distribution. Pensions are approximately 30% of assets for the Next 9.

Wealth levels fell for all groups in the wealth distribution in 2020:Q1. The decline was almost all due to the stock market revaluations, which mostly operate through corporate equity and mutual funds but also through the defined contribution plans included in pension entitlements. Due to their greater exposure to equity markets, the Top 1 and Next 9 lost 9 and 6 percent of their wealth in 2020:Q1, respectively, whereas the “Next 40” (households between the 50th and 90th percentiles) and the Bottom 50 lost 2 and 5 percent.

---

10 Pension entitlements include both 401(k)-type accounts and defined benefit accumulations, but not individual retirement accounts (IRAs) with 401(k)-type accounts directly affected by revaluations.
Figure 3. Change in Net Worth by Wealth Group, by Asset and Liability Type

After the 2020:Q1 decline, all wealth groups saw gains during the “Rebound” (the net change between 2020:Q1 and 2022:Q1). Stocks recovered quickly and then surpassed previous highs (before slipping in the latest quarter). These market movements drove rapid increases in wealth for the top 10 percent of the wealth distribution. For the Top 1, nearly 70% of total wealth gain in the Rebound is from corporate equities, while equities made up about half of the gains for the Next 9. For the Next 9, increases in pension entitlements was an additional 20% of their total wealth gain. The combination of pensions and equities were also a large driver of asset increases for the Next 40, making up nearly 40%. This share falls to 20% for the Bottom 50. It is also important to note, however, that a minority of the Bottom 50 benefitted from rising stock market prices since only one-third of these households own any public equity.¹¹

¹¹ Including exposure to equities held through defined contribution pension plans.
Real estate assets were also a major contributor to wealth gains, as house prices rose more than 35 percent, accelerating beginning in 2020:Q3 and continuing into early 2022. However, compared with corporate equities, housing is a relatively small piece of the Rebound, making up about a quarter of the total net revaluations. That said, real estate assets comprise the majority of assets for the Bottom 50; thus, the strong house price growth is the largest driver of wealth gains for that group, accounting for almost 80% of their gains. However, these wealth gains are not evenly spread across that group as only 40% of Bottom 50 own their home.

Since the SCF reveals that the Bottom 50 consistently holds a larger share of deposits than they do other non-pension financial assets, the large increase in their checking and savings balances also contributed meaningfully to wealth gains for the Bottom 50 as projected in the DFA. Between 2019:Q4 and 2022:Q1, the Bottom 50 groups' holdings of deposits grew by more than 50% and increased as a proportion of non-pension financial assets from 33% to 41%. Overall, the wealth of the Bottom 50 doubled from 2019:Q4 levels, which is about 65 percentage points more than any other wealth group. However, since their wealth share is quite small, they received only 6% of the total wealth gain during the pandemic.

Figure 4. Net Worth Shares by Wealth Group

In total, changes in the concentration of wealth were not large in magnitude. Consistent with Bricker et al. (2022), the share of wealth going to the Top 1% moved procyclically, increasing 2.2 percentage points in the Rebound period after falling 1.1 percentage points in 2020:Q1. Despite their wealth growing at the fastest pace, the small wealth level for the Bottom 50 resulted in their share only increasing by 1.0 percentage points between 2019:Q4 and 2022:Q1.

In the DFA, the concentration of wealth across the income distribution was also not substantially reshaped in the pandemic. However, in contrast to the lower wealth households, lower income households saw somewhat slower wealth growth compared to higher income households. The bottom 40 percent of the income distribution (Bottom 40) and the middle quintile group (families between the 40th and 60th percentiles) experienced growth of about 28% over the rebound and recovery period after only small decline in 2020:Q1. Meanwhile, the top 20 percent
group by income and households in the 60th to 80th percentiles of the income distribution experienced growth of 36% and 41%, respectively. The primary difference between the DFA projections for the bottom wealth and income groups is how the models, based upon historical data, allocate the large increases in deposits. Specifically, they project a slight increase in the share of deposits for the Bottom 50 percent of the wealth distribution but a slight decrease in the share for the Bottom 40 income group. This highlights both how the distributions of wealth and income are distinct, but also the challenges of modelling excess savings during the pandemic, which we turn to in the next section.

While we use the DFA as the starting point for studying wealth during COVID-19, we acknowledge there are major aspects of the pandemic for which the DFA estimation approach does not account. To begin to explore this uncertainty, it is useful to evaluate what portion of the DFA estimates come from assets and liabilities that are most likely still represented accurately in these abnormal times. Seeing that transaction volumes for assets and liabilities other than deposits were not considerably abnormal after 2019, we assume that there was little active rebalancing of portfolios by households (i.e. the 2019 SCF provides a valid description for these assets). Further, 74% of total wealth gain since 2019:Q4 was from the revaluations of real estate and corporate equities. This grows to 97% when including pensions and business wealth, which were also driven by stock market and real estate market movements, respectively, since 2019:Q4. This suggests we should have greater confidence in the DFA results for groups whose portfolios were comprised of more of these assets in the 2019 SCF. The Top 1 percent of the wealth distribution held around 80% of their assets in equities, businesses, pensions, and real estate in 2019, but this falls to 69% for the Bottom 50. The pattern across the income distribution is similar.

Alternative Distributions of Excess Savings

We now turn to exploring how the true evolution of the wealth distribution during the pandemic may deviate from the DFA projections. The net effect of the pandemic-specific factors on the distribution of savings is unclear. For example, the categories of spending that declined the most during the early lockdowns, such as travel and entertainment, are skewed towards affluent households (Bureau of Labor Statistics, 2022). Further, many of the service sector jobs most disrupted by COVID-19 pay relatively low wages. However, the replacement rates provided by expanded unemployment insurance were often above 100% for lower income workers (Ganong, Noel, and Vavra, 2020), and the Economic Impact Payments phased out for couples earning more than $150,000.

Our goal in this section is to gauge how this uncertainty might alter the DFA measurement of the wealth distribution as of 2022:Q1. To do so, we estimate how much of household savings stemmed from pandemic-specific factors (“excess savings”), run the DFA models without these excess savings, and then study the range of outcomes when the excess savings are distributed under alternative scenarios.
As a starting point, we define excess savings as the net increase in asset and liability transactions over the rate for 2019, which accumulates to $1.8 trillion from 2020:Q1-2022:Q1, with the great majority of the accumulation occurring in 2020.\textsuperscript{12} Our estimates for 2020 are similar in magnitude to others, such as Blanchard (2021), Briggs and Mericle (2021), and Aladangady et al. (2022) who also study this period. Figure 2 strongly suggests that the vast majority of excess savings flowed into bank accounts rather than were used to pay down debt or were otherwise invested (for additional evidence, see Briggs and Mericle, 2021). Thus, for our counterfactual DFA, we subtract the quarterly estimates of excess savings from the levels of savings and checking deposits. To complete the exercise, we add the excess savings to deposits under three alternative distribution scenarios: 1) roughly halfway between an equal distribution and the baseline DFA, which is heavily skewed towards the wealthy, 2) equal across the population, and 3) entirely to the bottom 50% of the wealth and income distributions. We construct these scenarios for each quarter, and aggregate across the pandemic to present them in Table 1, along with the baseline DFA increase [from the previous section] in deposits for comparison. Briggs and Mericle (2021) use the limited, available data to distribute the spending cuts, income losses, and government support to quintiles of the income distribution. They find that although excess savings was far from equally distributed, it was considerably less skewed to the top than is estimated by the baseline DFA. Greig, Deadman, and Noel (2021) come to a similar conclusion studying the account balances of JP Morgan Chase banking customers. Thus, we view the “somewhat equal” scenario as the most realistic, but we include the others, particularly the unrealistic, “all to the bottom 50%” scenario to establish plausible bounds.

<table>
<thead>
<tr>
<th>Distribution of 2022-Q1 Excess Savings</th>
<th>DFA</th>
<th>Somewhat Equally Distributed</th>
<th>Equally Distributed</th>
<th>All to the Bottom 50%</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Wealth Group</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Top 1</td>
<td>32%</td>
<td>8.6%</td>
<td>1%</td>
<td>0%</td>
</tr>
<tr>
<td>Next 9</td>
<td>32.2%</td>
<td>16.3%</td>
<td>9%</td>
<td>0%</td>
</tr>
<tr>
<td>Next 40</td>
<td>28.3%</td>
<td>38.6%</td>
<td>40%</td>
<td>0%</td>
</tr>
<tr>
<td>Bottom 50</td>
<td>4.5%</td>
<td>36.4%</td>
<td>50%</td>
<td>100%</td>
</tr>
<tr>
<td><strong>Income Group</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>99-100</td>
<td>29.9%</td>
<td>8.1%</td>
<td>1%</td>
<td>0%</td>
</tr>
<tr>
<td>80-99</td>
<td>37.8%</td>
<td>24.3%</td>
<td>19%</td>
<td>0%</td>
</tr>
<tr>
<td>60-80</td>
<td>17.5%</td>
<td>21.3%</td>
<td>20%</td>
<td>0%</td>
</tr>
<tr>
<td>40-60</td>
<td>9.7%</td>
<td>17.4%</td>
<td>20%</td>
<td>0%</td>
</tr>
</tbody>
</table>

\textsuperscript{12} We use the 2019 transactions as the counterfactual for simplicity, and because it aligns well with other estimates of excess savings for this period. If we instead project the counterfactual using the positive, but insignificant, trend in net transactions between 2016 and 2019, the excess savings would fall to $1.2 trillion, and the results below would roughly scale accordingly. The $1.8 trillion of excess savings is a net effect of a $2.39 trillion increase in asset inflows, less a $553 billion increase in liability inflows. Although we run the counterfactual through the latest period, 2022-Q1, we note that the greatest deposit inflows (excess savings) occurred in the first half of 2020.
Tables 2 and 3 present the results. The first takeaway is that the significant amount of excess savings did not substantially reshape the wealth distribution. Even under the extreme assumption that all excess savings went to the bottom 50%, the gains in the wealth share of the bottom wealth and income groups are very small compared to the degree of wealth inequality. Although there was a large increase in savings during the pandemic, it was somewhat overshadowed by large price increases in asset classes that are heavily skewed towards the wealthy and remains small compared to the level of household wealth.

Table 2. Net Worth Shares and Growth Rates by Wealth Group

<table>
<thead>
<tr>
<th>Wealth Group</th>
<th>DFA 2019:Q4</th>
<th>DFA 2022:Q1</th>
<th>Somewhat Equally Distributed</th>
<th>Equally Distributed</th>
<th>All to the Bottom 50%</th>
</tr>
</thead>
<tbody>
<tr>
<td>Top 1</td>
<td>30.7%</td>
<td>31.8%</td>
<td>31.4%</td>
<td>31.3%</td>
<td>31.3%</td>
</tr>
<tr>
<td>Next 9</td>
<td>38.7%</td>
<td>37.3%</td>
<td>37.1%</td>
<td>37.0%</td>
<td>36.9%</td>
</tr>
<tr>
<td>Next 40</td>
<td>28.8%</td>
<td>28.1%</td>
<td>28.3%</td>
<td>28.3%</td>
<td>27.8%</td>
</tr>
<tr>
<td>Bottom 50</td>
<td>1.8%</td>
<td>2.8%</td>
<td>3.2%</td>
<td>3.4%</td>
<td>4.0%</td>
</tr>
</tbody>
</table>

Net Worth Growth Rate:

<table>
<thead>
<tr>
<th>Wealth Group</th>
<th>DFA 2019:Q4</th>
<th>DFA 2022:Q1</th>
<th>Somewhat Equally Distributed</th>
<th>Equally Distributed</th>
<th>All to the Bottom 50%</th>
</tr>
</thead>
<tbody>
<tr>
<td>Top 1</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Next 9</td>
<td>32.9%</td>
<td>31.3%</td>
<td>30.9%</td>
<td>30.8%</td>
<td></td>
</tr>
<tr>
<td>Next 40</td>
<td>23.8%</td>
<td>23.0%</td>
<td>22.7%</td>
<td>22.3%</td>
<td></td>
</tr>
<tr>
<td>Bottom 50</td>
<td>99.4%</td>
<td>129.5%</td>
<td>142.1%</td>
<td>188.8%</td>
<td></td>
</tr>
</tbody>
</table>

Table 3. Net Worth Shares and Growth Rates by Income Group

<table>
<thead>
<tr>
<th>Income Group</th>
<th>DFA 2019:Q4</th>
<th>DFA 2022:Q1</th>
<th>Somewhat Equally Distributed</th>
<th>Equally Distributed</th>
<th>All to the Bottom 50%</th>
</tr>
</thead>
<tbody>
<tr>
<td>99-100</td>
<td>25.1%</td>
<td>26.8%</td>
<td>26.9%</td>
<td>26.9%</td>
<td>26.8%</td>
</tr>
<tr>
<td>80-99</td>
<td>45.2%</td>
<td>43.1%</td>
<td>43.4%</td>
<td>43.4%</td>
<td>43.1%</td>
</tr>
<tr>
<td>60-80</td>
<td>15.0%</td>
<td>15.5%</td>
<td>15.8%</td>
<td>15.8%</td>
<td>15.5%</td>
</tr>
<tr>
<td>40-60</td>
<td>7.5%</td>
<td>7.4%</td>
<td>7.6%</td>
<td>7.7%</td>
<td>7.7%</td>
</tr>
<tr>
<td>0-40</td>
<td>7.3%</td>
<td>7.1%</td>
<td>7.5%</td>
<td>7.6%</td>
<td>8.2%</td>
</tr>
</tbody>
</table>

Annual Net Worth Growth Rate:

<table>
<thead>
<tr>
<th>Income Group</th>
<th>DFA 2019:Q4</th>
<th>DFA 2022:Q1</th>
<th>Somewhat Equally Distributed</th>
<th>Equally Distributed</th>
<th>All to the Bottom 50%</th>
</tr>
</thead>
<tbody>
<tr>
<td>99-100</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>80-99</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>60-80</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Income Group</th>
<th>DFA 2019:Q4</th>
<th>DFA 2022:Q1</th>
<th>Somewhat Equally Distributed</th>
<th>Equally Distributed</th>
<th>All to the Bottom 50%</th>
</tr>
</thead>
<tbody>
<tr>
<td>99-100</td>
<td>37.4%</td>
<td>38.0%</td>
<td>37.5%</td>
<td>37.4%</td>
<td></td>
</tr>
<tr>
<td>80-99</td>
<td>22.6%</td>
<td>23.5%</td>
<td>23.3%</td>
<td>22.6%</td>
<td></td>
</tr>
<tr>
<td>60-80</td>
<td>32.6%</td>
<td>34.9%</td>
<td>34.8%</td>
<td>32.6%</td>
<td></td>
</tr>
<tr>
<td></td>
<td>40-60</td>
<td>60-80</td>
<td>80-100</td>
<td>100-130</td>
<td></td>
</tr>
<tr>
<td>------</td>
<td>-------</td>
<td>-------</td>
<td>--------</td>
<td>---------</td>
<td></td>
</tr>
<tr>
<td>0-40</td>
<td>25.9%</td>
<td>32.6%</td>
<td>35.1%</td>
<td>44.4%</td>
<td></td>
</tr>
</tbody>
</table>

In the baseline DFA, wealth for the bottom of the wealth distribution grew at the fastest rate, followed by that of the Top 1. Our alternative scenarios for excess savings do little to alter that conclusion. However, due to their low levels of wealth, this exercise reveals nontrivial uncertainty in the growth rate of wealth held by the bottom half of the wealth distribution. The DFA shows wealth for the Bottom 50 growing by 99% in the pandemic, which increases to 130% in our preferred “somewhat equal” excess savings scenario, 140% if excess savings were equally distributed, and 189% if the savings all went to the bottom. On a per household basis, wealth of the Bottom 50 increased from $31,000 in 2019Q4, to $60,000, $69,000, $73,000, and $86,000 in 2022Q1 under the DFA and three alternative scenarios. Thus, although there is material uncertainty in the evolution of wealth for the Bottom 50 during the pandemic, the qualitative finding of rapid wealth growth for the Bottom 50 relative to other groups holds for the DFA and the alternative scenarios.

When segmenting the population by income, this exercise again reveals little variation in the rate of wealth gain for the top of 60 percent of the distribution, but some uncertainty for the Bottom 40 percent. However, the range of plausible growth rates for the Bottom 40 of income is lower than that for wealth because per household wealth levels for the lower income groups are substantially higher than for the bottom half of the wealth distribution.

Conclusion

The distribution of wealth has been relatively stable recently despite the massive economic upheaval caused by the COVID-19 pandemic. This is true both in the baseline DFA estimates and in our alternative scenarios for the distribution of excess savings. The stability of the wealth distribution stems from the fact that most of the wealth gained comes from appreciation of assets that were heavily concentrated before the pandemic. The gains on these assets were so large that they dwarf the historic increase in household savings that also occurred and leave limited margin for the uncertainty of the distribution of excess savings to alter the overall wealth distribution. Our simulation exercise also reveals the challenges of relying upon historical relationships to model the wealth distribution during unprecedented circumstances. Since it has no information about the distributional patterns of income loss, early spending declines, or the government response, the DFA likely assigns too much of the large increase in savings to high wealth and income households.

Although the effects on the overall wealth distribution are small, the uncertainty due to the pandemic does modestly alter the estimates of changes in wealth experienced by less affluent households. In particular, our exposition of the alternate distributions of excess savings highlights the significance of the government response in supporting economic well-being of low wealth and income households. Now, as the final government stimulus checks dried up with the end of 2021 and consumers face higher prices due to inflation in 2022, a new set of considerations surround the potential drawdown in savings that may follow.
References


Wealth Inequality and COVID-19 in the US

Note: The analysis and conclusions set forth are those of the authors and do not indicate concurrence by other members of the research staff, the Board of Governors, or the Federal Reserve System.

Elizabeth Holmquist, Federal Reserve Board of Governors | 11th Biennial IFC Conference, BIS, Basel, 25 August 2022
Outline

1. Changes in aggregate household wealth during the COVID-19 pandemic
2. The composition and distribution of wealth over COVID-19 using the Distributional Financial Accounts
3. Implications of pandemic-driven excess savings on the distribution of wealth
4. Conclusion
U.S. Households gained $31 trillion since 2019

- Revaluations account for nearly 80% of wealth creation
- Equity markets rebounded strongly after the 2020:Q1 crash
- Real estate price growth accelerated in 2020:H2

Level and Change in Aggregate U.S. Household Net Worth

Note: Other changes in volume, which account for a much smaller portion of wealth change, are omitted.
Household savings also surged during the pandemic

- Resilience in certain sectors of the economy, reduced consumption, and large amounts of fiscal relief fueled savings
- Increased savings largely flowed to deposits, which saw unprecedented gains in net transactions over the pandemic

Change in Net Worth by Wealth Group

Change in Net Worth by Asset and Liability Type

<table>
<thead>
<tr>
<th>Wealth Group</th>
<th>19:Q4</th>
<th>22:Q1</th>
<th>Change</th>
</tr>
</thead>
<tbody>
<tr>
<td>Top 1</td>
<td>30.7%</td>
<td>31.8%</td>
<td>1.1%</td>
</tr>
<tr>
<td>Next 9</td>
<td>38.7%</td>
<td>37.3%</td>
<td>-1.4%</td>
</tr>
<tr>
<td>Next 40</td>
<td>28.8%</td>
<td>28.1%</td>
<td>-0.7%</td>
</tr>
<tr>
<td>Bottom 50</td>
<td>1.8%</td>
<td>2.8%</td>
<td>1.0%</td>
</tr>
</tbody>
</table>

Alternative Distribution Scenarios for Excess Savings: Wealth Groups

**Excess Savings**: Net increase in asset and liability transactions over the 2019 rate

<table>
<thead>
<tr>
<th>Period</th>
<th>Excess Savings ($ bill)</th>
</tr>
</thead>
<tbody>
<tr>
<td>2020:H1</td>
<td>1,417</td>
</tr>
<tr>
<td>2020:H2</td>
<td>132</td>
</tr>
<tr>
<td>2021:H1</td>
<td>591</td>
</tr>
<tr>
<td>2021:H2</td>
<td>-140</td>
</tr>
<tr>
<td>2022:Q1</td>
<td>365</td>
</tr>
<tr>
<td><strong>Total</strong></td>
<td><strong>1,837</strong></td>
</tr>
</tbody>
</table>
Excess savings did not substantially reshape the wealth distribution
## Net Worth Growth Rate

<table>
<thead>
<tr>
<th>WEALTH GROUP</th>
<th>DFA 2022:Q1</th>
<th>SOMEWHAT EQUALLY</th>
<th>EQUALLY</th>
<th>ALL TO BOTTOM 50%</th>
</tr>
</thead>
<tbody>
<tr>
<td>Top 1</td>
<td>32.9%</td>
<td>31.3%</td>
<td>30.9%</td>
<td>30.8%</td>
</tr>
<tr>
<td>Next 9</td>
<td>23.8%</td>
<td>23.0%</td>
<td>22.7%</td>
<td>22.3%</td>
</tr>
<tr>
<td>Next 40</td>
<td>25.3%</td>
<td>26.2%</td>
<td>26.2%</td>
<td>23.9%</td>
</tr>
<tr>
<td>Bottom 50</td>
<td>99.4%</td>
<td>129.5%</td>
<td>142.1%</td>
<td>188.8%</td>
</tr>
</tbody>
</table>

Rapid relative growth for Bottom 50 across the scenarios but material uncertainty about growth rate
Conclusion

• The distribution of wealth has been relatively stable despite the massive upheaval caused by the pandemic.

• Gains on assets concentrated towards the top of the distribution were so large that they dwarf the historic increase in savings and leave limited margin for uncertainty on the distribution of excess savings.

• This simulation exercise reveals the challenges of relying on historical relationships to model the wealth distribution during unprecedented circumstances.

https://www.federalreserve.gov/releases/z1/dataviz/dfa/
11th Biennial IFC Conference on “Post-pandemic landscape for central bank statistics”
BIS Basel, 25-26 August 2022

Sectoral financial positions and debt levels at the post-pandemic period

Ahmet Tayyar Fırat,
Central Bank of the Republic of Türkiye

---

1 This presentation was prepared for the conference. The views expressed are those of the author and do not necessarily reflect the views of the BIS, the IFC or the central banks and other institutions represented at the event.
Sectoral Financial Positions and Debt Levels at the Post-Pandemic Period

Eleventh IFC Conference on

“Post-pandemic landscape for central bank statistics”

BIS Basel, 25 and 26 August 2022

Ahmet Tayyar Fırat

Abstract

Central Bank of the Republic of Türkiye (CBRT) has acted very swiftly after the first Covid-19 case was seen in Türkiye in March 2020 in order to mitigate the potential impact of the pandemic on the economy especially on the financial, non-financial and household sectors. At this paper, the outcomes of these decisions and the effect of the pandemic on these sectors in Türkiye will be analysed by comparing the sectoral financial account tables for the pre and post pandemic periods. Cross - country debt comparisons for this period will be also made by using Financial Accounts in order to contribute to the use of the Financial Accounts Statistics.

Keywords: Sectoral Financial Accounts, Debt Levels, Post Pandemic Period

Contents

Sectoral Financial Positions and Debt Levels at the Post-Pandemic Period ............... 1
Introduction ................................................................................................................................. 2
1. Covid–19 in the World & Türkiye ....................................................................................... 2
2. CBRT Policy Responses to Pandemic and Its Outcomes ............................................. 4
   2.1 CBRT Policy Responses ............................................................................................... 4
   2.2 Potential Targets of These Measures ......................................................................... 5
   2.3 CBRT Policy Outcomes ............................................................................................... 5
   2.4 G20 Central Banks’ Responses to Covid -19 .............................................................. 6
   2.5 G20 Countries GDP Growth Rates ........................................................................... 7
3. Sectoral Financial Accounts of Türkiye ..................................................................... 8
4. Debt Levels at Post Pandemic Period ......................................................................... 9
Introduction

The first positive Covid-19 case was seen in Türkiye on March 11th 2020, a short while after the world had faced the disease. Like other governments and central banks, Türkiye also started to take proactive measures in order to save the public from the adverse effects of the pandemic. These measures mainly have two dimensions. The first one is related to the health of the public. Lock downs, using face masks, quarantining, physical or social distancing, banning of mass gathering events, closing schools and universities, working remotely, closing borders and international travel restrictions etc. are all part of the first dimension, saving the health of the public. The second dimension of these measures is related to the health of the economy. Cutting interest rates, adjusting the liquidity in the market, unemployment restrictions, social contributions to households etc. are all the examples of the second dimension, more specifically saving the economic health of the sectors like households, financial and non-financial corporations. It was observed that these two dimensions are totally linked to each other. For example, any travel restriction decision has negative impacts on the non-financial sector companies specifically on the tourism, airline and service sectors. As a domino effect, this also affects its employees, so the household sector. These sectoral interactions can be diversified into many different cases.

This paper comprises of four parts. Firstly, it shows the recent figures of both Covid cases and deaths in the world and Türkiye. Secondly, it summarizes the responses of Central Bank of the Republic of Türkiye (CBRT) to the pandemic and expresses its outcomes on the balance sheet. It also analyses some G20 Central Banks’ responses to Covid-19 at this part. Thirdly, it reviews sectoral financial positions of Turkish Economy with the distinction of households, government, non – financial and financial sectors in the pandemic and post pandemic period. Finally, the paper discusses cross country comparisons of sectoral debt levels in the post pandemic period.

1. Covid–19 in the World & Türkiye

The Covid-19 virus was firstly identified in Wuhan, China in December 2019. After it was seen in other countries and regions from the different parts of the world, it was declared as a pandemic by the The World Health Organization (WHO) in March 2020. Most of the authorities and governments took drastic actions in order to mitigate the adverse effects of the pandemic both on the health of the public and economies. The first year of the pandemic, 2020, can be accepted as the pandemic year with the global fight to the Covid -19. After starting the widely usage of the vaccination, 2021 and afterwards can be marked as the post – pandemic period.
Graph 1: Total Top 20 Covid Cases in the World

As of the end of July 2022, total Covid cases in the world is about 591 million and the countries in the first three rows are USA, India and Brazil respectively. Türkiye is the 10th place with the 15.5 million cases (Graph 1).

Graph 2: Total Top 20 Covid Deaths in the World

Source: Worldometers - Access: July 2022
Total Covid – 19 deaths in the World is about 6.4 million people as the end of July 2022 and like the Covid cases, the first three countries are same, USA Brazil and India, just a change at the 2nd and 3rd places. Türkiye takes the 19th place at the death ranking with about 100,000 deaths (Graph 2).

Both the cases and death figures have a direct relationship with the populations of the countries. The crowded countries would most likely to face much more cases than the others. It is a fact that there is also a link with the cases and the death numbers. It is observed in this period that the health system of the countries and the vaccination rates are be the drivers of these relationship. Being at below level for the death ranking is a sign of the strong health system of Türkiye and its high vaccination rates.

2. CBRT Policy Responses to Pandemic and Its Outcomes

2.1 CBRT Policy Responses

Measures taken by Central Bank of the Republic of Türkiye in the pandemic period can be classified in the below four headings:

2.11 Cutting the Interest Rates
- 17 March 2020 - The policy rate was cut by 100 bps to 9.75%
- 22 April 2020 - The policy rate was cut by 100 bps to 8.75%
- 21 May 2020 - The policy rate was cut by 50 bps to 8.25%

From March to May 2020 the central bank policy rate was decreased totally 250 bps from 10.75% to 8.25%.

2.12 Increasing Liquidity in the Market
- Primary dealers' open market operation limits were increased.
- FX reserve requirement ratios were reduced by 500 bps for banks meeting certain loan growth criterion.
- The primary dealer banks were allowed to sell the government securities they bought from the Unemployment Insurance Fund to the CBRT.
- The collateral pool for both TL and FX lending was expanded to include asset-backed and mortgage-backed securities.
- The limits for lower cost and longer-term lending facilities were increased.
- Swap transaction quotas for banks were raised by 50%.
- Upper limit of government security purchases was raised to 10% of the CBRT asset size.

2.13 Widening Rediscount Credit Facilities
- Maturity and repayment dates for CBRT exporter rediscount credits (which is borrowed in Turkish Lira and repaid in FX) were extended.
- The limits for lower cost and longer-term lending facilities were increased.
• Launched a new low-cost facility (150 bps below the policy rate) to support TRY loans to exporters.

2.14 Swap agreements
• Bilateral currency swap agreement between the CBRT and Qatar Central Bank was increased by $10 billion.

2.2 Potential Targets of These Measures
The targets of the above policy responses can be summarized as follows:
• Enhancing predictability by providing banks with flexibility in Turkish Lira and foreign exchange liquidity management,
• Offering targeted additional liquidity facilities to banks to secure uninterrupted credit flow to the corporate sector,
• Boosting cash flow of exporting firms through arrangements on rediscount credits,
• Strengthen the monetary transmission mechanism by boosting the liquidity of the Government Domestic Debt Securities (GDDS) market,
• Enhance banks’ flexibility in Turkish Lira and foreign exchange liquidity management,
• Secure uninterrupted credit flow to the corporate sector, and broadly support the goods and services exporting firms.

2.3 CBRT Policy Outcomes

Quantitative easing is defined as a monetary policy strategy used by central banks by purchasing securities in an attempt to reduce interest rates which spurs economic growth with increasing the supply of money and drive more lending to households and non-financial companies.

Like other central banks, CBRT also preferred to have quantitative easing in the pandemic period by purchasing government bonds. The fundamental target of this decision is to provide much liquidity to the market.
Total government bond portfolio of CBRT was less than 20 billion TRY before March 2020. With the start of the pandemic, it was increased swiftly to TRY 80 million level which corresponds to more than 400% increase (Graph 3).

Quantitative easing and the other measures show their effects especially on the Turkish Lira denominated items in the CBRT balance sheet (Graph 4).

2.4 G20 Central Banks’ Responses to Covid -19

Bank of Canada had the strongest respond to Covid – 19 in 2020 within the selected G20 countries by its huge asset purchasing program. At the pandemic year, in 2020, it has 358% increase at its balance sheet.
CBRT balance sheet grown 58% in 2020 and %101 in 2021. The main driver of the increase at post pandemic period, namely in 2021, is the upside valuation of FX assets (Graph 5).

Graph 5: Asset Changes of G20 Central Banks

![Asset Changes of G20 Central Banks](source: BIS)

2.5 G20 Countries GDP Growth Rates

In 2019, before the pandemic period most of the selected G20 countries have positive growth rates based on the different dynamics of their own economies. In the pandemic year, this picture had negatively changed and most of them faced with economic recession. Türkiye is one of the two G20 countries with China which had positive growth rates in this period. Furthermore, at the post pandemic period all the selected economies have high positive growth rates especially due to the base effect (Graph 6).

Graph 6: Asset Changes of G20 Central Banks

![GDP Growth Rates](source: OECD)
3. Sectoral Financial Accounts of Türkiye

Financial Accounts Statistics provide systematic information of the main sectors (non-financial corporations, financial corporations, general government, households) in the economy. It also shows linkages between the economy and the rest of the world.

Graph 7: Sectoral Financial Accounts of Türkiye

![Sectoral Financial Net Worths (Billion TRY)](image)

Source: CBRT

Turkish economy is financial borrower over the years. Households and the Rest of the World sectors are creditors, and Non-Financial Corporations and the General Government sectors were debtors (Graph 7).

Graph 8: Sectoral Financial Accounts of Türkiye

Financial accounts also serve the details on the stocks, transactions and valuations. Transaction figures give inside information on the sectoral net lendings and borrowings (Graph 8).

![Sectoral Net Lending / Borrowings (Billion TRY)](image)

Source: CBRT
Graph 9: Sectoral Financial Accounts of Türkiye

Valuation changes shows how the exchange rates and market prices effect the sectoral positions (Graph 9).

Source: CBRT

4. Debt Levels at Post Pandemic Period

Debt is a part of the liabilities side and here it comprised of the loans and issued debt securities. With the effect of the pandemic, total debt of resident sectors to GDP ratio reached to 167% level in 2020 Q3. At the post pandemic period, it has been going down to 150% levels thanks to the fast recovery in the economy (Graph 10).

Graph 10: Sectoral Financial Accounts of Türkiye

Source: CBRT
In order to measure the impact of the pandemic on the sectoral debt levels, the changes on the “debt to GDP ratios” for the selected countries are compared for the pre and post pandemic periods. During these two years period, total debt to GDP ratios of economies were increased except Norway, Belgium and Czech Republic. Furthermore, Türkiye has totally 23.9% ratio increase, mainly driven by the financial and non-financial sectors (Graph 11).

**Graph 11: Cross-Country Comparison of Debt/GDP**

![Graph 11: Cross-Country Comparison of Debt/GDP](image)

Source: OECD

**References**


https://www.worldometers.info/coronavirus/#countries

https://www.oecd.org/newsroom/g20-gdp-growth-fourth-quarter-2021-oecd.htm


BIS DBSOnline / Macroeconomic Data Set / Financial sectors balance sheets / Central Banks
Sectoral Financial Positions and Debt Levels at the Post-Pandemic Period

Ahmet Tayyar FIRAT

Central Bank of the Republic of Turkey
Data Governance and Statistics Department

Eleventh IFC Conference on “Post-pandemic landscape for central bank statistics”

BIS Basel, 24-26 August 2022
CONTENTS

▪ Covid-19 in the World & Türkiye
▪ Policy Reponses and Outcomes
▪ Sectoral Financial Positions
▪ Debt Levels at Post Pandemic Period
Covid in the World & Türkiye

In terms of total covid cases Türkiye is in the 10th place in the world.

For the death ranking Türkiye takes the 19th place.

Source: Worldometers - As of end of July 2022
CBRT Policy Responses to Covid-19

1) Cutting Interest Rates
2) Increasing Liquidity in the Market
3) Widening Rediscount Credit facilities
4) Swap agreements

CBRT Bond Portfolio (Billion TRY)

CBRT Total Assets (Billion TRY)

Source: CBRT
G20 Central Banks’ Responses to Covid-19

Bank of Canada had the strongest respond to Covid – 19 in 2020 by its huge asset purchasing program.

Source: BIS
Türkiye is one of the two G20 countries with China which had positive growth rates in 2020.

Source: OECD
**Türkiye Sectoral Financial Accounts**

Turkish economy is financial borrower over the years.

Households and the Rest of the world sectors are creditors, and Non-Financial Corporations and the General Government were debtors.

<table>
<thead>
<tr>
<th>Year</th>
<th>Non-Financial Corporations</th>
<th>Households</th>
<th>Financial Corporations</th>
<th>General Government</th>
<th>Rest of the World</th>
</tr>
</thead>
<tbody>
<tr>
<td>2015Q1</td>
<td>-8,000</td>
<td>-1,000</td>
<td>2,000</td>
<td>3,000</td>
<td>1,000</td>
</tr>
<tr>
<td>2015Q2</td>
<td>-7,000</td>
<td>-1,500</td>
<td>2,500</td>
<td>3,500</td>
<td>1,500</td>
</tr>
<tr>
<td>2015Q3</td>
<td>-6,000</td>
<td>-2,000</td>
<td>3,000</td>
<td>4,000</td>
<td>2,000</td>
</tr>
<tr>
<td>2015Q4</td>
<td>-5,000</td>
<td>-2,500</td>
<td>3,500</td>
<td>4,500</td>
<td>2,500</td>
</tr>
<tr>
<td>2016Q1</td>
<td>-4,000</td>
<td>-3,000</td>
<td>4,000</td>
<td>5,000</td>
<td>3,000</td>
</tr>
<tr>
<td>2016Q2</td>
<td>-3,000</td>
<td>-3,500</td>
<td>4,500</td>
<td>5,500</td>
<td>3,500</td>
</tr>
<tr>
<td>2016Q3</td>
<td>-2,000</td>
<td>-4,000</td>
<td>5,000</td>
<td>6,000</td>
<td>4,000</td>
</tr>
<tr>
<td>2016Q4</td>
<td>-1,000</td>
<td>-4,500</td>
<td>5,500</td>
<td>6,500</td>
<td>4,500</td>
</tr>
<tr>
<td>2017Q1</td>
<td>0</td>
<td>-5,000</td>
<td>6,000</td>
<td>7,000</td>
<td>5,000</td>
</tr>
<tr>
<td>2017Q2</td>
<td>1,000</td>
<td>-5,500</td>
<td>6,500</td>
<td>7,500</td>
<td>5,500</td>
</tr>
<tr>
<td>2017Q3</td>
<td>2,000</td>
<td>-6,000</td>
<td>7,000</td>
<td>8,000</td>
<td>6,000</td>
</tr>
<tr>
<td>2017Q4</td>
<td>3,000</td>
<td>-6,500</td>
<td>7,500</td>
<td>8,500</td>
<td>6,500</td>
</tr>
<tr>
<td>2018Q1</td>
<td>4,000</td>
<td>-7,000</td>
<td>8,000</td>
<td>9,000</td>
<td>7,000</td>
</tr>
<tr>
<td>2018Q2</td>
<td>5,000</td>
<td>-7,500</td>
<td>8,500</td>
<td>9,500</td>
<td>7,500</td>
</tr>
<tr>
<td>2018Q3</td>
<td>6,000</td>
<td>-8,000</td>
<td>9,000</td>
<td>10,000</td>
<td>8,000</td>
</tr>
<tr>
<td>2018Q4</td>
<td>7,000</td>
<td>-8,500</td>
<td>9,500</td>
<td>10,500</td>
<td>8,500</td>
</tr>
<tr>
<td>2019Q1</td>
<td>8,000</td>
<td>-9,000</td>
<td>10,000</td>
<td>11,000</td>
<td>9,000</td>
</tr>
<tr>
<td>2019Q2</td>
<td>9,000</td>
<td>-9,500</td>
<td>10,500</td>
<td>11,500</td>
<td>9,500</td>
</tr>
<tr>
<td>2019Q3</td>
<td>10,000</td>
<td>-10,000</td>
<td>11,000</td>
<td>12,000</td>
<td>10,000</td>
</tr>
<tr>
<td>2019Q4</td>
<td>11,000</td>
<td>-10,500</td>
<td>11,500</td>
<td>12,500</td>
<td>10,500</td>
</tr>
<tr>
<td>2020Q1</td>
<td>12,000</td>
<td>-11,000</td>
<td>12,000</td>
<td>13,000</td>
<td>11,000</td>
</tr>
<tr>
<td>2020Q2</td>
<td>13,000</td>
<td>-11,500</td>
<td>12,500</td>
<td>13,500</td>
<td>11,500</td>
</tr>
<tr>
<td>2020Q3</td>
<td>14,000</td>
<td>-12,000</td>
<td>13,000</td>
<td>14,000</td>
<td>12,000</td>
</tr>
<tr>
<td>2020Q4</td>
<td>15,000</td>
<td>-12,500</td>
<td>13,500</td>
<td>14,500</td>
<td>12,500</td>
</tr>
</tbody>
</table>

Source: CBRT
Sectoral Net Lending / Borrowings (Billion TRY)

Transaction figures give inside information on the sectoral net lendings and borrowings.

Valuation changes shows how the exchange rates and market prices effect the sectoral positions.

Source: CBRT
Türkiye Sectoral Financial Accounts

Total debt of resident sectors has been going down at the post pandemic period

Total Debt of Resident Sectors/GDP* (%)
Sectoral Financial Accounts

Pandemic Effect on the Total Debt of Resident Sectors

Cross-Country Comparison of Debt/GDP by Sectors (2019Q4 - 2021Q4, % Change)

Source: OECD
Thank you...

Ahmet Tayyar FIRAT
tayyar.firat@tcmb.gov.tr
Covid-19 crises –
New rapid data for the Israel economy and making it accessible to all\(^1\)

Daniel Rosenman,
Bank of Israel

---

\(^1\) This presentation was prepared for the conference. The views expressed are those of the author and do not necessarily reflect the views of the BIS, the IFC or the central banks and other institutions represented at the event.
COVID19 crises
New Rapid data for the Israel economy and Making it accessible to all

Daniel Rosenman
Bank of Israel - Information and Statistics department

11th. biennial IFC Conference
August 2022
The way to the success

1. The outbreak of the crisis: Mid March 2020
2. The immediate task: To find and to Bridge the information gap
3. Engagement to work together To achieve the main mission
4. The output
1 A
New rapid daily data for internal purposes
- Sectorial Activity on FOREX
- Funds Outflows
- Unemployment benefit claims
- Banking credit and mortgages in moratorium
- Bounced checks (NSF)
- Etc.

1 B
New rapid data published daily at our website
- Credit Card expenditure
- The Public’s Mobility Patterns
- Electricity Consumption in Israel
- Trends in the Activity in Cash

2
Dashboards
- First – Internal Dashboard
- Second - Public basic publication
- Late – Public dashboard in our website
Internal Dashboard
Graphic presentation of policy measures taken, displayed in a timeline

13/12/2020 - Increase in weekly amounts for food money allowances for families

22/10/2020 - Adjustment of income at an annual rate of 1.5% and 0.1% for families with up to 35 million shekels

06/07/2020 - Grant of a one-time financial assistance of 15 million shekels to each family

06/04/2020 - Reduction in interest rate for housing loans to 0.15% and 0.01% for families with up to 20 million shekels

14/01/2021 - Increase in child benefits for families with children

22/10/2020 - Increase in child benefits for families with children up to 15 million shekels

06/07/2020 - Increase in child benefits for families with children up to 15 million shekels

06/07/2020 - Grant of a one-time financial assistance of 15 million shekels to each family

06/04/2020 - Grant of a one-time financial assistance of 15 million shekels to each family
First Public access to the data

<table>
<thead>
<tr>
<th>Indicator name</th>
<th>Worksheet</th>
<th>Link to press release</th>
</tr>
</thead>
<tbody>
<tr>
<td>Credit Card Expenditure</td>
<td>xlsx</td>
<td>Special analysis: The changes in credit card purchases as a result of the coronavirus crisis. Updated data on changes in credit card expenditures as a result of the coronavirus crisis.</td>
</tr>
<tr>
<td>The Public’s Mobility Patterns</td>
<td>xlsx</td>
<td>Special analysis by the Bank of Israel Research Department: Initial economic insights from indices of changes in mobility patterns in Israel</td>
</tr>
<tr>
<td>Analysis of Electricity Consumption in Israel</td>
<td>xlsx</td>
<td>Analysis of electricity consumption in Israel as a rapid indicator of economic activity</td>
</tr>
<tr>
<td>Trends in the Activity in Cash</td>
<td>xlsx</td>
<td>The trends in withdrawals and deposits in shekels and in the currency circulation</td>
</tr>
</tbody>
</table>

https://www.boi.org.il/en/Pages/Indicators.aspx
Final Public access to the data

Rapid economic indicators

Credit to the public

Financial and FOREX markets

Real activity

https://bankipedia.boi.org.il/dashboard/
Forward looking

Accessibility

Flexibility

Transparency
Thank you

Email: daniel.rosenman@boi.org.il
WhatsApp: +972544968149
The impact of central bank interventions on non-performing loans under Covid-19 pandemic – The United Kingdom and Brazilian Study Case¹

Frederico Barros Diniz, University of Warwick, Warwick Business School, and João Paulo Vieira Costa, University of Brasília, Brazil

¹ This presentation was prepared for the conference. The views expressed are those of the authors and do not necessarily reflect the views of the BIS, the IFC or the central banks and other institutions represented at the event.
The Impact of Central Bank Interventions on Non-Performing Loans Under the COVID-19 Pandemic

The United Kingdom and Brazilian Study Case

Frederico Barros Diniz, University of Warwick, Warwick Business School

João Paulo Vieira Costa, University of Brasilia, Brasilia, Brazil

Abstract: This paper aims to assess the impact of the Central Bank’s monetary policies implemented to minimise the effect caused by the COVID-19 Pandemic on the Credit Lending environment. Those policies are reflected in many economic variables, including the number of non-performing loans rate. Thus, this work can be defined as a quantitative case study using the Ordinary Least Squares (OLS) methodology to evaluate the British and Brazilian banks’ NPL behaviour from a macroeconomic perspective and assess government interventions and the economic response. The work concludes that implementing a conventional macroeconomic policy parallels quantitative easing that positively influences the Brazilian economy. The outcome from the model developed is consistent with what is observed in other countries; estimating the trajectories and levels with a certain confidence level throughout the NPL time series is possible.

Keywords: Banking, Credit Risk, IFRS 9, Non-Performing Loan, COVID-19.
## Contents

The United Kingdom and Brazilian Study Case ................................................................. 1

Introduction ....................................................................................................................... 3

Central Bank Mandates and Early Response to the Covid-19 ................................... 3

The Response to the Covid-19 Bank of England and Brazilian Central Bank .......... 4

Non-Performing Loans (NPL) ......................................................................................... 7

Literature Review ........................................................................................................... 8

Credit Risk during the COVID-19 Pandemic ................................................................. 8

Non-Performing Loan Modelling ................................................................................. 9

Ordinary Least Squares (OLS) Models on Credit Risk ............................................... 9

Methodology .................................................................................................................. 10

Data and Scope ............................................................................................................. 11

Exploratory Analysis and Econometric Modelling ..................................................... 11

Results and Analysis .................................................................................................... 11

Model Development .................................................................................................... 11

NPL Forecast and Discussion ..................................................................................... 13

Conclusions ................................................................................................................... 15

References ..................................................................................................................... 16
Introduction

Default probabilities are essentially compelled by how companies are tied to economic, domestic and foreign cycles and how they are linked to their business conditions (Pesaran et al., 2006). Thus, regulators have implemented macroeconomic and credit-specific factors in forecasting default since numerous studies have explored the benefits of dynamic factor analysis in credit risk modelling. Basel Accords for Unexpected Loss Estimation and IFRS 9 and CECL for Expected Loan Loss are good examples of how exogenous factors can be modelled to predict default.

Those two international specifications were the regulator’s response to the global Financial Crisis (GFC) of 2007 - 2009. Basel III (Basel Committee on Banking Supervision 2010) corresponds to the regulatory capital banks should hold to absorb the unexpected loss. IFRS 9 (International Accounting Standards Board 2014) and the Current Expected Credit Losses (CECL), the two international accounting standards, drive banks to adopt a forward-looking approach to calculating their expected loss using macroeconomic factors.

The countries can use tools to manage their financial systems and their monetary policies. It can vary from country to country, and as expected, there are individual contrasts specific to the Brazilian and British economies, which the Central Bank led. The most common actions to manage monetary policy objectives are modifying the interest rate, buying and selling government bonds, regulating foreign exchange rates, and changing the number of money banks must maintain as reserves (Domanski, Kohlscheen, and Moreno 2016). However, actions should be more robust during a global recession, as observed during the COVID-19 Pandemic.

The economic dynamics are a significant driver of the evolution of arrears in the pool of loans granted to non-financial companies or individuals, followed by the financial pressure induced by the monetary conditions. However, lending allows companies and householders to provide resources for investment projects or debt repayments. Thus, the lending policy’s role in commercial banks’ exercise is essential, as it may impact the cost of credits and the loan portfolio quality. However, governmental intervention can drive how financial institutions design their lending policy during deep recessions, promoting economic reaction.

Central Bank Mandates and Early Response to the Covid-19

Central Banks have as their primary objective financial stability or stability in the value of their currency. The Macroeconomic theory argues that the best way to keep inflation under control is to give the responsibility to set monetary policy to an independent central bank (Castillo-Martinez, 2019). However, monetary policy decisions are implemented quite differently in different countries, and the level of development of their economies is a significant factor for implementation.

Financial institutions provide essential services to the real economy, such as deposits and lending, for retail and commercial purposes. However, during a deep recession, it is common to see an increase in unemployment and a reduction in consumption and output, causing a natural deflation. Thus, once the COVID-19 Pandemic started, the Brazilian Central Bank (BCB) and the Bank of England (BoE) and most of their global peers intervened in the economy to avoid the deflation trap and preserve jobs.
The GFC, followed by quantitative easing by major central banks, slow growth in European economies, and a growing recognition of social inequality, has led some observers to argue for changes to central bank mandates. The more orthodox notes that controlling inflation was insufficient to prevent the GFC. That line of debate leads to discussions concerning a central bank’s role in financial stability, its potential use of macroprudential tools and whether it should be the prudential Supervisor. However, the reaction to COVID-19 utilising further quantitative easing was essential for the Central banks to deliver the expected financial stability.

Table 1 below shows the critical social-economic differences between Brazil and the United Kingdom, which is considered a developed economy, while Brazil is a developing economy.

<table>
<thead>
<tr>
<th>Metric (2020)</th>
<th>Brazil</th>
<th>United Kingdom</th>
</tr>
</thead>
<tbody>
<tr>
<td>Unemployment rate:</td>
<td>13.7%</td>
<td>3.7%</td>
</tr>
<tr>
<td>Base Rate:</td>
<td>4.5%</td>
<td>0.8%</td>
</tr>
<tr>
<td>Inflation Rate:</td>
<td>3.2%</td>
<td>1.0%</td>
</tr>
<tr>
<td>Cost of Living: (USA = 100%)</td>
<td>49.8%</td>
<td>95.5%</td>
</tr>
<tr>
<td>Commercial taxes and contributions:</td>
<td>65.1%</td>
<td>30.6%</td>
</tr>
<tr>
<td>GDP per capita</td>
<td>6,970 US$</td>
<td>40,320 US$</td>
</tr>
<tr>
<td>GDP (current US$)</td>
<td>$ 1449 bi</td>
<td>$ 2700 bi</td>
</tr>
<tr>
<td>Government debt (% of GDP):</td>
<td>92.0%</td>
<td>104.0%</td>
</tr>
<tr>
<td>Government debt (% of GDP):</td>
<td>88.0%</td>
<td>103.0%</td>
</tr>
<tr>
<td>IMF Financial Development Index (Rank)</td>
<td>20</td>
<td>5</td>
</tr>
</tbody>
</table>

Source: World Bank Data

The Response to the Covid-19 Bank of England and Brazilian Central Bank

The policymakers reduced bank exposure to liquidity, solvency, and financial stabilisation issues. The approach used by the BCB and the BoE to put in place an effective crisis management regime was based on four main aspects:

- **Supervision** - Early intervention involves the Supervisor’s ability and willingness to act early in problem banks cases, in which predictive actions can be applied.
- **Resolution** - The power to apply measures that minimise stability risks for the rest of the system at the minimum cost while protecting assets' value.
- **Deposit Insurance** - Provides certainty to small depositors, mitigating the risk of runs and contagion to similar institutions.
- **Emergency Liquidity Assistance** - The provision of emergency liquidity, on a bilateral basis, to solvent banks to preserve financial stability.

The BCB implemented changes through conventional monetary policy; however, the policymakers also called for unprecedented unconventional monetary policy measures (or Quantitative Easing) due to the severity of the crisis. The section below summarises the list of actions taken by the BCB in response to the COVID-19 Economic crisis.
**Liquidity Measures**

The Brazilian Central Bank dealt with seizing the money markets by providing liquidity to the Brazilian banking system. These measures imply a potential expansion in the system liquidity of BRL 1,217 bn, amounting to about 16.7% of GDP.

<table>
<thead>
<tr>
<th>Support Measure</th>
<th>Support Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Loans backed by L.F. guaranteed by credit operations</td>
<td>670, 122-149</td>
</tr>
<tr>
<td>New Term Deposit with Special Guarantees - DPGE</td>
<td>200, 36 - 44</td>
</tr>
<tr>
<td>Reduction in reserve requirement ratio on time deposits (from 31% to 25%) and Regulation enhancement on Liquidity Coverage Ratio (LCR)</td>
<td>135, 25-30</td>
</tr>
<tr>
<td>Loans backed by debentures</td>
<td>91, 17-20</td>
</tr>
<tr>
<td>Additional reduction in reserve requirement ratio on time deposits (from 25% to 17%)</td>
<td>70.00, 13-16</td>
</tr>
<tr>
<td>Change on reserve requirement on savings deposits</td>
<td>55,8, 44905</td>
</tr>
<tr>
<td>One-year term repurchase backed by federal securities</td>
<td>50, 44874</td>
</tr>
<tr>
<td>More flexibility on LCA regulation</td>
<td>2,2, 0.4 - 0.5</td>
</tr>
<tr>
<td>Total</td>
<td>1,274,00, 232 - 283</td>
</tr>
<tr>
<td>Total as a % Of GDP*</td>
<td>17,50%</td>
</tr>
</tbody>
</table>

* As a percentage of the 2019 Brazilian GDP.

** F.X. range: from BRL 4.5/USD to BRL 5.5/USD, given the high uncertainty.

Sources: BCB 2020

The Bank of England focused on reducing the Bank Rates and a New Term Funding Scheme and applied similar measures to the Central Bank of Brazil.

<table>
<thead>
<tr>
<th>Support Measure</th>
<th>Description</th>
<th>Support Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Bank rate</td>
<td>A reduction of the Bank Rate by 50 basis points</td>
<td>0.25%</td>
</tr>
<tr>
<td>Term Funding Scheme</td>
<td>provide funding for a term of four years of at least 5% of market participants' real economy lending stock at rates equal to, or around, the Bank Rate</td>
<td>£5 of funding for every £1 of positive net-lending to SMEs</td>
</tr>
<tr>
<td>Bond Purchases</td>
<td>Maintenance of the BOE’s stock of sterling non-financial investment-grade corporate bond purchases (valued at £10 billion) and government bond purchases (valued at £435 billion)</td>
<td>£ 445 billion</td>
</tr>
<tr>
<td>Mortgage holidays</td>
<td>Mortgage providers will grant three-month mortgage holidays</td>
<td>N/A</td>
</tr>
<tr>
<td>SME loans</td>
<td>provide loans of up to £5 million to SMEs, with no interest due for the first six months.</td>
<td>£79 billion</td>
</tr>
</tbody>
</table>

Source: Bank of England
Capital Measures

Financial institutions may face losses after the COVID-19 crisis. BCB adjusted the regulation on capital requirements to provide financial institutions with better conditions to sustain credit flow. These measures can increase credit supply by BRL 1,197 billion, amounting to about 16.4% of GDP.

<table>
<thead>
<tr>
<th>Support Measure</th>
<th>Support Value</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>BRL (bn)</td>
</tr>
<tr>
<td>Reduction of ACP factor (from 2.5% to 1.25%)</td>
<td>637</td>
</tr>
<tr>
<td>Overedge of investments (tax effects)</td>
<td>520</td>
</tr>
<tr>
<td>Reduction in the capital requirement for credit operations for SMEs</td>
<td>35</td>
</tr>
<tr>
<td>Temporary reduction of capital requirement for smaller financial institutions</td>
<td>16,5</td>
</tr>
<tr>
<td>Reduction of capital requirement on DPGE exposures</td>
<td>12,7</td>
</tr>
<tr>
<td>Working Capital Program to preserve business continuity (CGPE)</td>
<td>127.00</td>
</tr>
<tr>
<td><strong>Total</strong></td>
<td><strong>1,348,20</strong></td>
</tr>
<tr>
<td>% of GDP*</td>
<td><strong>16,70%</strong></td>
</tr>
</tbody>
</table>

* As a percentage of the 2019 Brazilian GDP.

**F.X. range: from BRL 4.5/USD to BRL 5.5/USD, given the high uncertainty.

Sources: BCB

Other Measures

BCB implemented some regulatory changes to facilitate the application and acceptance of Credit Facilities for householders and corporates.

<table>
<thead>
<tr>
<th>Support Measure</th>
<th>Support Value</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>BRL (bn)</td>
</tr>
<tr>
<td>Swap lines with U.S. Federal Reserve</td>
<td>637</td>
</tr>
<tr>
<td>Creation of a unique credit line for SMEs (PESE)</td>
<td>520</td>
</tr>
<tr>
<td>Real estate-backed loans</td>
<td>35</td>
</tr>
<tr>
<td>Purchase of private securities by BCB in the secondary market</td>
<td>16,5</td>
</tr>
</tbody>
</table>

* As a percentage of the 2019 Brazilian GDP.

**F.X. range: from BRL 4.5/USD to BRL 5.5/USD, given the high uncertainty.

Sources: BCB

The BoE has also applied some supervisory measures to reduce the pressure on banks and allow the institutions to provide customers with financial support via supervisory actions.
The Impact of Central Bank Interventions on Non-Performing Loans Under the COVID-19 Pandemic

BoE Supervision Support

<table>
<thead>
<tr>
<th>Support Measure</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>investigation of open-ended funds</td>
<td>suspended until further notice</td>
</tr>
<tr>
<td>Non-critical obligations</td>
<td>on site-visits and deadlines for firms and FMIs that are not critically postponed</td>
</tr>
<tr>
<td>Senior Management Function</td>
<td>They are easing the review of applications for Senior Management Functions to reduce regulatory burdens.</td>
</tr>
</tbody>
</table>

Source: Bank of England

The effectiveness of those monetary policies in the economy has an immediate effect, and it can be crucial to define the recovery from the crisis. Thus, the Central Bank’s primary goal is to preserve jobs and maintain financial stability. Those policies are reflected in many economic variables, including the number of non-performing loans.

Non-Performing Loans (NPL)

Most of the published papers prove the high correlation between the macroeconomic environment and the non-performing loan rate, constating that default rates rise once the economy starts a downturn, and the expected loss amount, commonly called provisions, also tends to increase. It is not different in Brazil and the United Kingdom.

Graph 1 shows that NPL levels and provisions follow the same trend. However, since the Pandemic began, NPL loans have been reduced. However, credit loss provision rates have increased, which is not natural behaviour and can be explained by the macroprudential rules adopted by the BCB.

Provision and NPL Ratios - March 2011 to January 2021 (Source: BCB) Graph 1

The ratio between the NPL and the provision of the Brazilian banking system, shown in Graph 1, makes this gap between the perception of risk and the NPL clearer. The forward-looking component and the uncertainty about the pandemic outcome drive these abnormal behaviours.
In this context, this paper aims to evaluate the NPL behaviour of Brazilian banks from a macroeconomic perspective and clarify the behaviour of NPL in the COVID-19 Pandemic.

In addition to this introduction, this work has four sections. In the literature review, a general context about NPL modelling is sought. The methodology describes the steps followed to achieve the objective of this work. The result and analysis show the developed models. In conclusion, there are the main findings, limitations, and future work.

A similar comparison for the British banking sector could not be performed due to the lack of information from the BoE or other official entities.

Literature Review

Since the global financial crisis (GFC) 2007-2009, there has been an evolution concerning the guidelines for managing credit portfolios. The two most recent international requirements, IFRS 9, CECL and Basel III, also resulted from the recession. Equally to the GFC, the COVID-19 Pandemic produced reflections on the adequacy of the current methodologies adopted and opened an international discussion on possible improvements to the existing instruments; This is the topic for the following section.

Credit Risk during the COVID-19 Pandemic

All countries that have taken measures to reduce the adverse effects caused by the COVID-19 Pandemic, particularly by monetary policies adopted by their respective central banks, such as the European Central Bank (ECB) and the Federal Reserve (FED), observed some positive outcomes from those actions (Teresienė, Keliuotytė-staniulienė, and Kanapickienė 2021). The authors used panel data regression to estimate this paper's main factors influencing long-term loans. As a result, the effects of the COVID-19 Pandemic on financial stability by increasing liquidity and reducing profitability and the banking sector solvency.

(Acharya and Steffen 2020) They conducted a study to investigate the effect of credit risk on the Bank’s liquidity during the COVID-19 Pandemic, using a sample of American companies with financial variables available in Capital I.Q. The authors highlight the significant impact of credit risk on financial institutions’ liquidity in this period.

(Gubareva 2020) assessed the effects of the COVID-19 Pandemic, specifically on the liquidity of emerging market securities. In this case, the protection of companies in the financial sector is more resistant to liquidity shocks, associating this situation with the lack of legislation in non-financial companies. However, the author highlights the slow recovery of credit spreads to pre-crisis levels due to increased credit risk from COVID-19.

As discussed in previous sections, the outcome of central banks and governments’ actions to contain the negative impacts of COVID-19 has been lectured in recent works and evaluated the Pandemic’s effect on Small and Medium Enterprises (SMEs). (Corredera-Catalán, di Pietro, and Trujillo-Ponce 2021) They observed the credit quality and the Spanish guaranteed system’s effectiveness on
portfolio evaluation. The authors stated that one way of addressing SMEs' vulnerabilities might be through Mutual Guarantee Systems (MSGs), which guarantee access to financing with credit risk mitigation providing less impact on the public budget.

Those collaterals' disposition to SMEs leading is essential to provide the necessary cash to guarantee these companies' continuance amid social distance and lock-down measures during hibernation (Didier et al., 2021). There were measures to make labour relations more flexible in Brazil and provide greater credit access to help SMEs during the COVID-19 Pandemic (Central Bank of Brazil 2020).

In the context of the COVID-19 pandemic emergency, effective crisis resolution in banks is critical to avoiding adverse effects during this period, mainly due to the significant increase in NPL caused by the borrower’s repayment capacity (Bodellini and Lintner 2020). The COVID-19 pandemic's effect on the NPL was discussed by (Hardiyanti and Aziz 2021); modelling a simple regression, the authors concluded the impact of COVID-19 on commercial banks' NPL in Indonesia.

Going beyond the simple NPL hypothesis assessment, as in the paper by (Hardiyanti and Aziz 2021), other studies identified factors connected with NPL and econometric methods. That research could quantify the factors' significance in predicting the NPL, a subject addressed in the following subsection.

Non-Performing Loan Modelling

Even before the 2007-2009 crisis and the consequent improvements incorporated into the risk management and credit risk quantification mechanisms due to the recession, NPL modelling was already a way of understanding the economy. (Chang et al. 2008) conducted a study on the relationship between the Brazilian NPL and credit concentration. Using a panel data approach, the authors identified a significant effect of banking concentration on the NPL.

As shown in Figure 1: Provision and NPL Percentage Series - March 2011 to January 2021 (Central Bank of Brazil), the theme became more popular after the global economic crisis, and financial institutions and their regulators recognised the need to understand these relationships. (Louzis, Vouldis, and Metaxas 2012) Lists macroeconomic variables such as GDP, unemployment, interest rate and public debt, and the bank-specific information among NPL. (Messai and Jouini 2013) Brought a similar approach with the use of information from banks in European countries (Italy, Greece and Spain), finding hostile relations between the NPL and GDP and a positive with the unemployment rate and the real interest rate of the economy, among similar results from (Makri, Tsagkanos, and Bellas 2014), (Chaibi and Fititi 2015), (Bholat et al., 2018) Furthermore, (Bolognesi et al. 2020) started discussions and brought reflection, including the IFRS 9 perspective.

Ordinary Least Squares (OLS) Models on Credit Risk

The Ordinary Least Squares (OLS) estimator is, according to (Hayashi 2011), the most straightforward econometric procedure but with properties that make it applicable to different sample sizes. The use of this approach for modelling NPL and aggregate levels. Table 7 shows papers published in that context during the last ten years.
This paper aims to evaluate the NPL behaviour of the British and Brazilian banks from a macroeconomic perspective. NPL factors were assessed using The Ordinary Least Squares methodology in this context. The following section sets out the steps followed to meet the proposed objective.

Methodology

This paper can be characterised as a case study with a quantitative approach. In summary, the steps taken to achieve the objective of this research can be defined as follows:

**Exploratory data analysis:** The time series’ trends, characteristics and the selection of an econometric model were analysed.

**Econometric Modelling:** To assess the factors influencing NPL in the Brazilian banking sector.

**Projections and comparisons:** to verify the differences between the series projected by the econometric model developed and those observed after the policy measures to minimise the economic impacts of COVID-19.

For this work’s graphic representation, Matplotlib (Hunter 2007) and the applications of tests and econometric models, EViews were used.

---

1 EViews12 Student Version Lite, IHS Global Inc., Irvine, CA, USA
Data and Scope

The data used in this work were obtained from the Time Series Management System (Central Bank of Brazil n.d.), where BCB stores aggregated information regarding credit-related data in Brazil and information on economic activity and monetary policy, among others.

The data used for the United Kingdom information was extracted from the Bank of England MLAR Statistics Report in conjunction with the Financial Conduct Authority (FCA) and considered all the Residential loans to individuals: Regulated and Non-regulated accounts information.

The NPL time series used in this work comprises the total balance of credits in default over 90 days on the total credit balance amount. The NPL series are available monthly from March 2011 to February 2021. The Gross Domestic Product (GDP) data are granted quarterly in most NPL modelling works. Thus, the projections adopted a quarterly-based period for econometric modelling, covering from 2011Q2 to 2019Q4. However, the projections data were based on data from 2020Q1 to 2020Q4.

The data for modelling was selected for econometric techniques given the need for interpretation and consistency of each macroeconomic factor—the results and effects in the series and the predominance of these methods in NPL modelling studies.

Exploratory Analysis and Econometric Modelling

The exploratory analysis was conducted to verify the time series’ characteristics to choose the most appropriate modelling methodology. The time series variables were evaluated to assess the collinearity and their stationarity. For the assessment of correlation, the Johansen test was used. For stationarity, the Augmented Dickey-Fuller (ADF), Phillips-Perron (P.P.) and Kwiatkowski-Phillips-Schmidt-Shin (KPSS) tests (KPSS) were applied (Hayashi 2011).

These results enabled identifying the most appropriate variables for modelling using the OLS econometric approach. It was considered only the most functional macroeconomic characteristics, and the consistency of the model’s parameters was analysed, following the macroeconomic theory. The results obtained in the tests and the modelling are described in more detail in the following section.

Results and Analysis

The hypotheses regarding the variables used to develop the models in this work were based on the literature’s best usage. The input was the seasonal adjusted GDP growth, actual interest rate (Selic) and unemployment rate provided by the BCB (Central Bank of Brazil n.d.). This last variable was necessary to apply seasonality adjustment using the R language (R Core Team 2020).

Model Development

Table 8 presents the champion model used for the Brazilian Central Bank, and it brings the results of the tests using the most significative variables and the p-value on the relationship between each independent and dependent variable.
The Impact of Central Bank Interventions on Non-Performing Loans Under the COVID-19 Pandemic

OLS Champion Model Equation (Brazil) Table 8

<table>
<thead>
<tr>
<th>Variable</th>
<th>Coefficient</th>
<th>Std. Error</th>
<th>t-Statistic</th>
<th>Prob.</th>
</tr>
</thead>
<tbody>
<tr>
<td>C</td>
<td>0</td>
<td>0.02</td>
<td>-0.25</td>
<td>0.8</td>
</tr>
<tr>
<td>GDP Growth</td>
<td>-3.94</td>
<td>0.77</td>
<td>-5.09</td>
<td>0</td>
</tr>
<tr>
<td>D_NPLm(-1)</td>
<td>0.65</td>
<td>0.08</td>
<td>7.7</td>
<td>0</td>
</tr>
<tr>
<td>D_selic(-1)</td>
<td>0.04</td>
<td>0.02</td>
<td>1.96</td>
<td>0.06</td>
</tr>
<tr>
<td>DUMMY</td>
<td>-0.27</td>
<td>0.08</td>
<td>-3.41</td>
<td>0</td>
</tr>
</tbody>
</table>

Where the NPL stands for non-performing loan, Selic stands for Real Federal securities interest rate, and U.R. stands for Unemployment rate season corrected.

This result suggests that, in this period, other factors are not captured in the variables that impacted NPL, which can be interpreted as the result of actions to contain the impacts of the COVID-19 Pandemic on NPL.

Model Measure - OLS Champion Model (Brazil) Table 9

<table>
<thead>
<tr>
<th>Model Measurement Criteria</th>
<th>Coefficient</th>
<th>Calculation Method</th>
<th>Prob.</th>
</tr>
</thead>
<tbody>
<tr>
<td>R-squared</td>
<td>0.78</td>
<td>Mean dependent var</td>
<td>-0.05</td>
</tr>
<tr>
<td>Adjusted R-squared</td>
<td>0.76</td>
<td>S.D. dependent var</td>
<td>0.21</td>
</tr>
<tr>
<td>S.E. of regression</td>
<td>0.1</td>
<td>Akaike info criterion</td>
<td>-1.55</td>
</tr>
<tr>
<td>Sum squared residual</td>
<td>0.35</td>
<td>Schwarz criterion</td>
<td>-1.33</td>
</tr>
<tr>
<td>Log-likelihood</td>
<td>33.65</td>
<td>Hannan-Quinn criteria.</td>
<td>-1.47</td>
</tr>
<tr>
<td>F-statistic</td>
<td>28.83</td>
<td>Durbin-Watson stat</td>
<td>1.97</td>
</tr>
<tr>
<td>Prob(F-statistic)</td>
<td>4.28</td>
<td>Breusch-Godfrey - LM 4 lags</td>
<td>0.78</td>
</tr>
</tbody>
</table>

The tests applied to the residuals generated by model two indicate the absence of autocorrelation. All the independent variables’ importance can be considered significant at the level of significance of 10%. Thus, Model two will be considered for the following projections and evaluations because of the residues’ best diagnostic results.

OLS Champion Model Equation (UK) Table 10

<table>
<thead>
<tr>
<th>Variable</th>
<th>Coefficient</th>
<th>Std. Error</th>
<th>t-Statistic</th>
<th>Prob.</th>
</tr>
</thead>
<tbody>
<tr>
<td>YoY GDP Growth</td>
<td>0.0005</td>
<td>0.00017</td>
<td>2.8</td>
<td>0.0085</td>
</tr>
<tr>
<td>Employment and Unemployment Hazard Ratio</td>
<td>0.638</td>
<td>0.04</td>
<td>14.7</td>
<td>0</td>
</tr>
<tr>
<td>Real Disposable Income</td>
<td>-0.00089</td>
<td>0.000265</td>
<td>-3.3</td>
<td>0.0019</td>
</tr>
</tbody>
</table>

Analog to the results presented above on the Brazilian case, the methodology was applied to the United Kingdom data. The best model tested was selected to represent the forward-looking perspective on Non-performing loans in the British economy. Table 10 shows the results of the model and table 11 the Model Measures.
NPL Forecast and Discussion

After evaluating the factors related to the NPL series before the impacts of the COVID-19 Pandemic in Brazil and the UK, the projections for 2020 aim to assess the NPL trend explained by the independent variables (GDP and interest rate). Graph 2 and 3 shows the dynamic estimates of the NPL variation made in the model compared to the observed values, in addition to the confidence intervals calculated with two standard errors.

<table>
<thead>
<tr>
<th>Model Measure - OLS Champion Model (UK)</th>
<th>Table 11</th>
</tr>
</thead>
<tbody>
<tr>
<td>Model Measurement Criteria</td>
<td>Coefficient</td>
</tr>
<tr>
<td>R-squared</td>
<td>0.88</td>
</tr>
<tr>
<td>Adjusted R-squared</td>
<td>0.87</td>
</tr>
<tr>
<td>S.E. of regression</td>
<td>0.07</td>
</tr>
<tr>
<td>Sum squared residual</td>
<td>0.17</td>
</tr>
<tr>
<td>Log-likelihood</td>
<td>45.05</td>
</tr>
<tr>
<td>Durbin-Watson stat</td>
<td>1.69</td>
</tr>
</tbody>
</table>
It is essential to highlight the 2020Q1 period when the measures to contain the coronavirus's spread and the economic actions to prevent its effects on the economy started in March 2020. This quarter has a minor impact compared to the others regarding the results of the macroeconomic variables that sustain the model. Two critical movements should be highlighted: a reduction in the real Selic rate and two significant variations in GDP, with a considerable contraction in 2020Q2. A substantial recovery followed them in 2020Q3. These results are reflected in an increase in the NPL gap in 2020Q2 and a sharp reduction in 2020Q3.

It is possible to notice that during the entire period before the COVID-19 Pandemic, the model's estimates were close to the observed values projected from the factors that affect the NPL. It is noted that in 2020Q1. The comparative calculations can be attributed to the brief period affected by the pandemic scenario. 2020Q2, a period with a Zero Lower Bound (ZLB), follows a decreasing trajectory with the most significant GDP contraction in the historical series. There is a similar trajectory comparing the estimated and the observed value. However, the NPL variation estimate is more exacerbated than the empirical. It is within the interval of two standard errors.

In 2020Q3, the GDP recovery and the Selic trend reduced NPL. Thus, based on the relationship between the variables observed in the available historical series, the estimate is outside the confidence interval constructed from the standard error for the first time in the period. Finally, it is noted that there is a reduction in magnitude outside the observed pattern.

Another relevant movement is that observed in 2020Q4. The model estimates a rare moment of inversion of the observed trajectory by considering the significant effect of the lag of the response variable. The estimated model's expected outcome would be a drop, while a positive variation of the NPL was observed.

This result suggests that, in this period, other factors are not captured in the variables that impacted NPL, which can be interpreted as the result of actions to contain the impacts of the COVID-19 Pandemic on NPL.
Conclusions

This study aimed to assess the factors related to NPL in the British and Brazilian banking systems and the behaviours expected during the COVID-19 Pandemic. Several measures to contain the negative impacts on the economy and credit have been adopted by central banks worldwide.

It is possible to conclude that there is a positive influence on the real interest rate and a negative one between GDP over NPL and the lag of the dependent variable itself. These relationships make it possible to estimate the trajectories and levels with a certain confidence level throughout the NPL time series.

Considering the relationships’ estimation between the variables, a movement was out of the pattern observed in the rest of the series. It becomes robust in the period of solid GDP recovery. This result may indicate additional factors to the model influencing the curves, resulting from direct actions to reduce the impact on the credit circumstance.

It induces that the NPL observed in 2020 impacts the measures adopted to control the Pandemic. It is important to note that an increase below the expected by the model in 2020Q2 and a sharp reduction in 2020Q3 beyond the confidence interval of the model estimate in 2020Q4 when part of the measures to contain the Pandemic ceased to be in force. There was a trajectory reversal and compensation for the sharp drop beyond what was predicted in the previous movement.

Because of the current scenario, the authorities and banks must analyse this situation with caution, given the continuity of the Pandemic’s effects in Brazil and the smaller scope for applying similar measures and include some new factors such as the current geopolitical risk, the global rise in inflation and current political environment.

There are limitations in this work that are important to mention. The projected results start from a premise of adequate model adjustment for the presented comparisons. However, they may indicate the model’s calibration in an out-of-period and outside the modelling space for some independent variables. Relationships in similar directions to other literature works may minimise these relationships. Nevertheless, the magnitude of the estimated effect is different.

Updating the model with new periods to minimise these limitations are alternative for future work. Concluding, using variables related to the Pandemic and more robust models considering Brazilian and British regional data and panel data is a suggested exercise.
References


The Impact of Central Bank Interventions on Non-Performing Loans Under COVID-19 Pandemic
The United Kingdom and Brazilian Study Case.

Fred B. Diniz
+44 7860 405138
frederico.barros-diniz@warwick.ac.uk
Objectives

Assess the impact of the Bank of England (BoE) and the Brazilian Central Bank’s (BSB) monetary policies implemented to minimise the effect caused by the Covid-19 pandemic on the country's Non-Performing Loans.

Non-Performing Loans: is a bank asset that is subject to late repayment or is unlikely to be repaid by the borrower in full.

“Non-performing loans represent a major challenge for the banking sector, as it reduces the profitability of banks and is often presented as preventing banks from lending more to businesses and consumers, slowing down economic growth.”
Introduction

Economic Cycle
- Early Recession
- Full Recession
- Early Recovery
- Late Recovery

Recession Impact
- Macro Models
- Exposure at Default
- Default Probability
- Loss Given Default
- Asset Pricing
- Earnings
- Liquidity

Recession Outcome
- Losses Increase
- Capital Stress
- Management Actions

Affected Parameters

11th Biennial IFC Conference, 25th August 2022
Actions are taken by the BoE and BCB in response to the Covid-19 Economic crisis.

**Brazilian Central Bank**
1. **Liquidity Support**
   - Offer to Commercial banks and building societies long-term funding.
2. **Capital Relief** – BCB adjusted the regulation on capital requirements to provide financial institutions with better conditions to sustain credit flow.
3. **Regulatory changes** to facilitate the application and acceptance of credit facilities for householders and corporates.

**Bank of England**
1. **Monetary Policy** – Cut our interest rate to 0.1%
2. **Liquidity Support**:
   - Offer to Commercial banks and building societies long-term funding.
   - Helped businesses pay their staff and suppliers.
3. **Capital Relief** – Helped banks to expand lending
Methodology

- The steps taken to achieve the objective of this research can be defined as follows:
  - **Exploratory data analysis**: the time series' trends, characteristics and the selection of an econometric model were analysed.
  - **Econometric Modelling**: to assess the factors influencing NPL in the British and Brazilian banking sectors.
  - **Projections and comparisons**: to verify the differences between the series projected by the econometric model developed and those observed after the policy measures to minimise the economic impacts of COVID-19.

Data and Scope

- **Econometric modelling**: 2007Q2 to 2019Q4
  - UK: Exploring 25 Variables to predict Non-Performing Ratio over time. (Source ONS and BoE)
  - Brazil: Exploring 35 Variables to predict Non-Performing Ratio over time. (SFN Bacen)
- The NPL time series used in this work comprises the total outstanding balance of credits with 90 days in arrears on the total credit balance amount.
- Projection data: 2020Q1 to 2021Q4.
Non-Performing Loans (NPL)
Provision and NPL Percentage Series - Mar 2007 to Jan 2021


## Model Results

### United Kingdom
- YoY GDP – Annualised GDP Growth
- UR Hazard – UK: Hazard Rates: Employment to Unemployment (SA, %)
- Real Disp. Income – Household real disposable income

### Brazil
- NPL – Non-Performing Loans
- GDP – Seasonally adjusted GDP growth
- SELIC – Real interest rate (Selic)
- UR – Seasonality adjustment Unemployment Rate

### Comparison (C.I.) Between Predicted and Actual Values

**Differences in**

<table>
<thead>
<tr>
<th>Variable</th>
<th>Coefficient</th>
<th>Std. Error</th>
<th>t-Statistic</th>
<th>Prob.</th>
</tr>
</thead>
<tbody>
<tr>
<td>YoY GDP</td>
<td>4.94E-06</td>
<td>1.77E-06</td>
<td>2.798609</td>
<td>0.0085</td>
</tr>
<tr>
<td>UR Hazard</td>
<td>0.638709</td>
<td>0.043372</td>
<td>14.72643</td>
<td>0.0000</td>
</tr>
<tr>
<td>Real Disp. Income</td>
<td>-8.92E-06</td>
<td>2.65E-06</td>
<td>-3.370723</td>
<td>0.0019</td>
</tr>
<tr>
<td>R-squared</td>
<td>0.876323</td>
<td>0.199620</td>
<td>4.454444</td>
<td>0.0000</td>
</tr>
<tr>
<td>Adjusted R-squared</td>
<td>0.868828</td>
<td>0.199620</td>
<td>4.454444</td>
<td>0.0000</td>
</tr>
<tr>
<td>S.E. of regression</td>
<td>0.072298</td>
<td>-2.336387</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Sum squared resid</td>
<td>0.172491</td>
<td>-2.204427</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Log likelihood</td>
<td>45.05496</td>
<td>-2.290329</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Durbin-Watson stat</td>
<td>1.685265</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

### Brazil

**Comparison (C.I.) Between Predicted and Actual Values**

<table>
<thead>
<tr>
<th>Variable</th>
<th>Coefficient</th>
<th>Std. Error</th>
<th>t-Statistic</th>
<th>Prob.</th>
</tr>
</thead>
<tbody>
<tr>
<td>C</td>
<td>-0.0080</td>
<td>0.0186</td>
<td>-5.4303</td>
<td>0.0000</td>
</tr>
<tr>
<td>GDP Growth</td>
<td>-39.783</td>
<td>21.160</td>
<td>18.801</td>
<td>0.0000</td>
</tr>
<tr>
<td>D_NPL.m (-1)</td>
<td>0.7071</td>
<td>0.0934</td>
<td>75.676</td>
<td>0.0000</td>
</tr>
<tr>
<td>D_SELIC(-1)</td>
<td>0.0520</td>
<td>0.0256</td>
<td>20.326</td>
<td>0.0000</td>
</tr>
<tr>
<td>R-squared</td>
<td>0.7427</td>
<td>Mean dependent var</td>
<td>0.0341</td>
<td></td>
</tr>
<tr>
<td>Adjusted R-squared</td>
<td>0.7161</td>
<td>S.D. dependent var</td>
<td>0.0341</td>
<td></td>
</tr>
<tr>
<td>S.E. of regression</td>
<td>0.1050</td>
<td>Akaike info criterion</td>
<td>0.0341</td>
<td></td>
</tr>
<tr>
<td>Sum squared resid</td>
<td>0.3196</td>
<td>Schwarz criterion</td>
<td>-13.754</td>
<td></td>
</tr>
<tr>
<td>Log likelihood</td>
<td>296.866</td>
<td>Hansen-Quinn alter.</td>
<td>-14.957</td>
<td></td>
</tr>
<tr>
<td>F-statistic</td>
<td>279.057</td>
<td>Durbin-Watson stat</td>
<td>21.142</td>
<td></td>
</tr>
<tr>
<td>Prob(F-statistic)</td>
<td>0.0000</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

**Breusch-Godfrey Correlation LM Test 4 lags: F(4.25) p-value 0.2949**
Conclusions

● It can be induced that the NPL observed in 2020 impacts the measures adopted to control the pandemic. Bearing in mind that:

● There was an increase below the expected by the model in 2021Q2;

● There was a sharp reduction in 2020Q3, beyond the confidence interval of the model estimate, and in 2021Q4, when part of the measures to contain the pandemic ceased to be in force.

● There was a trajectory reversal and compensation for the sharp drop beyond what was predicted in the previous movement.
The effect of pandemic on Central Bank of Türkiye real sector database and post-pandemic adjustments

Merve Artman, İsmail Onur Yılmaz, Mustafa Tanyer and Fatih Yalçın Mete,
Central Bank of the Republic of Türkiye

---

1 This presentation was prepared for the conference. The views expressed are those of the authors and do not necessarily reflect the views of the BIS, the IFC or the central banks and other institutions represented at the event.
The effect of pandemic on Central Bank of Türkiye (CBRT) real sector database and post-pandemic adjustments

Merve Artman, İsmail Onur Yılmaz, Fatih Yalçın Mete, Mustafa Tanyer

Abstract

Timely data was crucial for taking proper decisions during the effects of pandemic. This paper aims to present the process of compiling CBRT’s highest frequency real sector data under pandemic conditions and post-pandemic harmonization studies. For the last four years, the CBRT has compiled detailed financial statement data (by making an amendment to its law) in currency distinction from companies with high foreign exchange liabilities. The data has been collected quarterly, however, during the pandemic period, companies delayed reporting their balance sheet data. This situation required CBRT to fully adapt the reporting system to pandemic conditions. So, CBRT has changed the reporting standard from IFRS to national accounting standards, reduced the frequency to weekly, and compiled some part of reporting from administrative records (like derivatives). Because CBRT received the data weekly, the balance sheet data could be used together with other high frequency loan and spot foreign exchange data. Weekly balance sheets are the highest frequency real sector balance sheet data CBRT can add to database so far. To move faster, IT infrastructure was developed in-house by the Statistics Department. After the pandemic conditions relaxed, CBRT changed the reporting from weekly to monthly, meanwhile it was a win to integrate company data with other high frequency databases, now researchers can work on this integrated dataset. And also after the pandemic conditions relaxed, CBRT can have time to integrate more complex reporting system to replace previous in-house infrastructure. The addition of other high frequency data such as quarterly temporary income statements and monthly VAT returns is now carried out within the scope of post-pandemic studies.

Keywords: post-pandemic, official statistics, real time data, data collection

JEL codes: C80, C81
Contents

The effect of pandemic on Central Bank of Türkiye (CBRT) real sector database and post-pandemic adjustments ................................................................. 1

1. The Motivation for New Datasets .............................................................. 3

2. The History of Dataset ............................................................................. 3

3. The New Weekly Dataset ....................................................................... 4

4. The Post-Pandemic Adjustments ............................................................. 8
1. The Motivation for New Datasets

During the pandemic period, central banks took many measures for companies to continue their activities in a healthy way. Timely monitoring of the effects of these regulations has become crucial. However, delays in company reporting did not make this possible. For this reason, some changes had to be made in frequency and structure of reporting in order to obtain timely data. This period has also been an important time for a detailed understanding of data and analysis needs. For this reason, post-pandemic statistics were adapted according to needs. With this motivation, CBRT built a real-time data platform as a monitoring tool with a well-established data streaming process. This process contains all necessary components to enable CBRT to integrate data for advanced analysis. High frequency raw data collected in real time from multiple data providers, cleaned, processed, matched and stored safely with other financial and non-financial CBRT data thanks to the data highway established on big data platform. Decision makers in CBRT can benefit from this enhanced data by visualization tools. These visualization tools link all the necessary information to overview market conditions via customized dashboards. Our most prominent motivation is to prove more complete, timely and granular information as a complement to traditional macroeconomic indicators. It is important to strengthen analysis for decision-making and data is valuable only if we can manage to extract value from it. To make this possible, we create a data highway on big data platform and this makes it easier to combine different data sets and extract necessary information for decision-making.

2. The History of Dataset

CBRT decided to establish a company-specific dataset where risks can be monitored in detail and a new regulatory framework of FX risk management to increase the resilience of Turkish economy against exchange rate volatility. For this purpose, Systemic Risk Monitoring System for FX assets and liabilities of non-financial companies was established by CBRT to monitor FX risk of the companies at micro level. The current databases enabled to reach information on FX liabilities of the companies at micro level but not on FX assets. The new system would collect the detailed FX assets information besides liabilities for better understanding of FX debt burden on individual companies and their capability of debt management. Non-financial firms that had FX loans USD 15 million and over were in the scope of the system, ca.2000 firms.

CBRT had made some legal arrangements to establish the system. With the amendment to article 44 of CBRT Law: “The Bank shall, in order to monitor the operations of real and legal persons which affect their foreign exchange position, be authorized to request all kinds of information and documents all kinds of information and documents from real and legal persons that it shall determine.” Also, Directive on Systemic Risk Monitoring System for FX Positions of Companies published in Official Gazette on February 17, 2018. CBRT cooperated with Credit Registry Bureau of Turkey for installation of system infrastructure, software, maintenance, operation and support services. Web-based reporting was designed for the reporting of FX positions.
According to the Directive, all companies except banks and financial institutions that had FX cash loans USD 15 million and over were in the scope of the system. So, the system covered all non-financial companies including public institutions like municipalities. The system required quarterly individual financial statements prepared according to international reporting standards for timely and proper understanding of the financial position and the performance of the companies. All financial statements should be under independent external auditing.

This system was the first comprehensive dataset in Turkey that contains FX assets, liabilities, derivatives, export and import information on currency and maturity (0-3, 3-6, 6-12 months and over 1 year) distinction. FX credit information included domestic loans, direct loans abroad, guaranteed loans abroad, intermediation loans abroad and FX indexed loans from banks or financial institutions. Also, this was the first database with separate derivative form on transaction basis and the inventory information based on the market value of the inventories in FX assets.

CBRT could follow the assets and liabilities in USD dollar, Euro, Turkish lira and total of other currencies. Also, companies could declare their assets and liabilities under the projection for the following 3 months, 6 months, This data set covered 83 percent of all FX credits in Turkey and 78 percent of CBRT rediscount credits. Also, the important part of total investments came from these companies although the share of the number was even less than one percent.

3. The New Weekly Dataset

From 2018 to 2020, we watched these company’s data quarterly and after 2021, we changed the frequency to monthly. However, during the pandemic period in 2020, we requested weekly position from companies from June to December to follow the covid-19 effects on FX positions. Due to the postponed financial reporting of companies within the framework of the coronavirus (Covid-19) epidemic, there were delays in the data reported to the CBRT.

Weekly data forms were also requested in accordance with Article 44 of the Central Bank of the Republic of Turkey Law No. 1211 in order to monitor the effect of monetary policies on the balance sheets in a timelier manner. Due to the current economic conditions, it was beneficial to collect the financial statements of the companies at a higher frequency in order to effective evaluation of the financial supports applied to the real sector. Companies that were in the scope of the system, required to report the Weekly Data Form in addition to their quarterly notifications.

As a result of the evaluation made within the scope of the developments regarding the normalization process, it is expected that the companies will complete their reporting for the first quarter of 2020 until 15 June 2020 and continue in their usual schedule, and weekly reports will start to be made from 16 June 2020 in the format determined by our Bank.

Weekly reportings were made in the format determined by our institution, using the values that were the basis of the legal financial statements prepared in accordance with the Accounting System Implementation General Communiqués ("MSUGT") and the Uniform Chart of Accounts (THP) requirements published by the Ministry of
Finance. The document was sent to companies through e-mails containing the form to be filled on a weekly basis, explanations about the form and answers to frequently asked questions from companies. Companies were not expected to make any calculations other than existing accounting records.

We requested from companies to send information and notifications regarding the weekly accounting period to the official e-mail address by filling out the form, until Tuesday evening every week, starting from June 16th 2020.

In the start of the process, negative comments were made that the data form requested to be filled was technically impossible for the company, that filling out the form would require additional work and this would be a waste of time for the company. Some of the companies that commented in this way stated that they could not fill out the form, and some of them stated the biggest difficulties in filling out the form in addition to their negative comments and wanted to get opinions on how to overcome these problems. The most frequently asked questions regarding the forms were about the weekly nature of the form and therefore the difficulty of updating the inventory items on a weekly basis. Especially in the questions sent from audit firms, the authorities stated that the data to be reported would be inconsistent and incorrect due to the time inconsistency between reporting and accounting realizations, and therefore weekly reporting would cause an unhealthy data flow.

Unlike the quarterly forms, the preparation of the form was made according to the Uniform Chart of Accounts and Tax Procedure Law, not according to TFRS and BOBI FRS.

Although there were limited number of balance sheet items (mostly financial, 13 assets and 12 liabilities) in the new form, there were no items that require additional calculations such as the balance sheet main total (for example, total assets and liabilities) or subtotals (for example, total assets). For this reason, companies were not expected to prepare weekly balance sheets or trial balances. In addition, if there were items that could not be calculated (for example, inventories, trade receivables) due to timing inconsistency in the reporting week, it would be sufficient to include the most up-to-date balance of the relevant items in the form.

In the weekly data form, the current balance of the requested item as of the last day of the reporting week (01.01.2020 – last day of the reporting week) should be written, not the change amounts of the relevant week for the requested items.

Since the account balances in the form would be updated on a weekly basis, changes resulting from such timing mismatches would be reflected in the account balances in subsequent reporting. Therefore, no retrospective adjustment was required.

If companies had balance other than USD and Euro for the items requested in the weekly form, the TL equivalent of the relevant amount should be written in the "TL Equivalent of Other Foreign Currency" column in the form. Weekly forms should be sent to official e-mail address until the end of the working day on the Tuesday following the end of the reporting week.

The recording or calculation period of some transactions might exceed one week. In these cases, the most recent balance available in the relevant item as of the last day of each reporting week should be written on the form. When the relevant transaction was recorded or calculated, the relevant amount should be included in
the form of the week containing this date (the day of recording or calculation). Also, there was no need to update the previous forms. Because certain balance sheet items were included in the weekly data form, and amounts that require additional calculations such as balance sheet main total (asset-liability total) or subtotal (asset-liability totals) were not requested, it was not necessary to prepare a weekly balance sheet or trial balance, as only the balance information for the requested items was requested.

The weekly form must be prepared within the scope of the Uniform Chart of Accounts and the Tax Procedure Law. So, no rediscount or accrual transactions were requested for the financial items included in the form. The weekly data sheet was not subject to independent audit. After completing the weekly data form, the reporting week information in the form was selected from the menu, the form would be saved and an e-mail would be sent to on the Tuesday following the reporting week. Based on the e-mail, tax identification number, title and reporting week was examined automatically and reporting controls was carried out. In the “Exchange Rate Value to be Used for Conversion to TL” field in the weekly data form, the USD and EUR rate valid on the last day of the reporting week should be entered. Since the last day of the reporting week is Sunday, the rates announced at 15:30 on the previous Friday should be taken into account. Companies that were liable within the scope of the Regulation based on Article 44 of the Central Bank of the Republic of Turkey Law No. 1211 will prepare the weekly data form.

It is stated in Article 68 of the aforementioned law that a judicial fine from one thousand days to two thousand days will be imposed on real persons and officials and related persons of legal persons who do not provide the information and documents required within the scope of the aforementioned article, or give falsely or in violation of the determined procedures and principles. The weekly data form was a form that was prepared on different principles from the quarterly data forms and contained limited items and aimed to provide information on the current balance sheets of companies. Due to these differences, there might be differences between other forms and tables. Since the controls in the form were a warning against incomplete and erroneous reporting, they can ignore the warning of the relevant control as long as companies were sure that the written amounts were correct.

Although some liability items could be tracked in time from administrative records, the only source of most of the data belonging to the companies is the notifications of the companies themselves. Although the Regulation is referred to as the scope of the firm, the said regulation aims to regulate the quarterly notifications. The weekly form was a complementary work with the quarterly statements of the companies. The “Current Date of Amount” field was added to the form. Thus, companies were able to enter the most up-to-date amount of the account items expected to be reported and notify us of the date.

With the weekly data form, 13 active and 12 passive items in the balance sheet of the companies were collected in USD, EUR, Other FX and TRY currencies. Data were collected via e-mails sent to official e-mail adress. The data collection process was carried out in the following steps:

First of all, an excel file template to be sent weekly was sent to all companies that have a reporting obligation. The Excel template was protected with a password, and companies were only allowed to enter data in the relevant fields. Firms were
requested to fill in the form with their current account balances as of Sunday every week and send it by the following Wednesday at the latest. E-mails sent by companies were read with a developed python script, error checks are passed and tables suitable for analysis were created. The process of the Python script was as follows:

Separating the mails that had excel that we sent to the companies as an attachment to the mail that comes to the mailbox or that were categorized as data forms by the business unit.

Checking whether the excel files attached to the parsed e-mails were compatible with the template we sent. If the template was corrupted, automatic e-mail reply containing this error was returned to the relevant companies.

If the template was correct, checking the basic information in the form (TAC, Title, reporting week, writing 0 in all fields) and returning an automatic mail reply suitable for errors (for example, missing vkn or reporting for future dates).

Reading the error-free form with the python openpyxl module using the excel sheet-excel cell and the account code matches (eg Data Form-C25 = Cashier) and converting it to tabular data format

Aggregating and converting the data collected for the relevant week into a single data file after completing the previous step for all forms. Combining the single data form created for the relevant week with the single data forms created for the previous weeks. At this stage, automatic confirmation e-mails were sent to the companies in order to prevent the companies from repetitive reporting for the same dates.

Updating the old data with the allowed data after the manual approval process (reading the mail content) (such as the company’s request to correct the old data).

Converting the data created in wide format into a structure suitable for analysis and transferring it to the big data environment.

The following data quality studies were carried out for the weekly collected data. Similar to the reading process in these studies, python scripts were used and the business process was automated.

**Significant changes in the same items between weeks.** With this control, significant changes in a reporting item above a certain amount and above the average of the item were detected, and confirmation e-mails were sent to the companies regarding the accuracy of the amounts.

**Not reporting foreign currency assets and liabilities.** With this control, confirmation e-mails were sent to companies that did not report any active or passive balances in foreign currency regarding the accuracy of the amounts.

**Inconsistency of loan balances with administrative records.** With this control, confirmation e-mails regarding the accuracy of the amounts were sent to the companies whose credit balance in the form differed significantly from the credit balance in the administrative records.

In the analyzes we made on a weekly basis, we were able to monitor the change in the general currency position and the weekly course of items such as safes, banks, receivables and payables on a currency basis. We were also able to combine this weekly dataset with other high-frequency loan and market currency-trading data. In addition to the analysis, the presence of a timely data set belonging to the companies also offered the opportunity to check the data quality with other high-frequency data.
4. The Post-Pandemic Adjustments

After the pandemic conditions relaxed, CBRT changed the reporting from weekly to monthly by integrating more complex reporting system to replace previous one. Thus, we reduced the reporting obligation on our companies from weekly to monthly and established a more sustainable data infrastructure. However, we did not want to return to quarterly reporting again, because the analyses arising from the combination of timely company data with other datasets were used extensively in upper management decisions.

In the monthly data form, different from the weekly data form, flow variables such as monthly domestic and international sales, stock purchases, tangible and financial fixed asset purchases are also included. Thus, it has made it possible to make monthly cash flow analysis of the companies. We can also provide ease of analysis by visualizing these cash flows with sankey charts.

The new monthly reporting system, which was also developed in house within the bank, was designed as 3 phases. The first phase includes companies accessing the system by performing security checks with e-government integration and entering data into the system within data confidentiality. In the first phase, basic controls such as entering the data with negative characters, subtotal control, grouping by thousands, decimal fraction control, minimum and maximum value are performed. In the second phase, the entry and exit transactions of the companies within the scope of the regulation and the corrections of the previous period are monitored. In addition, company communication, which was previously done via e-mail, is now done through a defined management module. Thus, important notes can be followed without getting lost. The third phase is devoted to the integration of secure data entry, user identification, deactivation, sending and recording of controls, which are desired to be done in the other two phases.

The addition of other high frequency data such as quarterly temporary income statements and monthly VAT returns is now carried out within the scope of post-pandemic studies. It was a win to integrate company data with other high frequency databases, now researchers can work on this integrated dataset. Seeing the benefits of timely analysis during the pandemic period has increased the data sharing efficiency of institutions with each other.

The monthly turnover table calculated within our Directorate every month; the main table of Revenue Administration, Company notifications transferred from Turkish Statistical Institute on a monthly basis pursuant to the Revenue Administration and CBRT protocol is obtained by compiling the "declaration year" and VAT and SCT declaration tables. The workflow, in which the monthly turnover calculation is carried out, includes the declarations of companies operating in the real sector. The aforementioned monthly turnover table has been prepared with the motivation to enable real sector companies to be monitored on a monthly basis and to provide input for other studies to be carried out within our Bank.

VAT declarations declared by the taxpayers until the evening of the 24th day of each month are transferred to our Bank through Turkish Statistical Institute approximately 45 days after they are notified to the Revenue Administration. Afterwards, the monthly turnover table is created by our Bank in its most up-to-date version. If more than one valid VAT declaration of a company for the same period is
Post pandemic adjustments CBT

detected during the calculation process, data cleaning processes are carried out to avoid duplication in the system. The monthly turnover table, which does not interfere with administrative records in any way, has been opened to users, except for the correction of previously determined error types that may occur due to the nature of the data. In addition, it was possible to expand the analysis studies carried out by using our large data set such as monthly domestic turnover, foreign turnover, VAT discounts, calculated VAT amounts and other tables of companies.

Temporary income statements are submitted to Revenue Administration once every 3 months, until the evening of the 17th day of the 2nd month following the quarterly period. The relevant data is transferred to the Turkish Statistical Institute by the Revenue Administration within 15 days, and transferred to our Bank from the Turkish Statistical Institute on the 15th of the next month. Therefore, realizations, which can be followed up to 6 months after 31 December within the scope of Sector Balance Sheets, can be followed after approximately 75 days in 3 months, thanks to the provision and processing of temporary income statements.

Temporary income statements consist of cumulative data. 1st period data covers 0-3 months of the relevant accounting period; 2nd period data covers 0-6 months; 3rd period data covers 0-9 months; The 4th period data represents the realizations between 0-12 months. Therefore, due to the fact that the accounting records cannot be deleted, it is expected that all account items in the table will increase in the following periods. At this point, when the 1st period temporary income statement data comes in, it is left fixed as there is no possibility of comparison; however, when the 2nd period data comes in, the 1st period data is adjusted. The latest period data is always accepted as correct and the previous period data is corrected by accepting the last period as correct. Within the scope of the corrections, in companies whose data has decreased recently, it is tried to provide a cumulative increase on an item basis for each company by transferring the decrease amount to another item that has increased more than this amount. Within the scope of the corrections, in companies whose data has decreased recently, it is tried to provide a cumulative increase on an item basis for each company by transferring the decrease amount to another item that has increased more than this amount. In addition, the data of the previous period is printed for the relevant period to the companies that do not submit a period declaration as of the 2nd period. Thus, it is aimed to prevent serial deterioration caused by incomplete declaration submission.

The data of the provisional income statements, both in their raw form and with the corrections were made available to the users of our Bank in series starting from 2010. Corrections regarding the completion of the missing declarations are made on the Hive platform using SQL, and the correction of the floating data is made in Jupyter Notebooks using the Python Pandas and Numpy libraries.

Perhaps the most important contribution of the pandemic to the world of statistics was to remind us how important a role it played. As statisticians, we made an effort to bring the most timely data possible to the database and make it available to researchers. Currently, we can see the benefits of establishing a timely and integrated database when we can perform timely impact analyses. This motivation brought to us by the pandemic process has positively affected the future of our statistical studies.
The effect of pandemic on Central Bank of Türkiye (CBRT) real sector database and post-pandemic adjustments

11th IFC Biennial Conference

25-26 August 2022

Basel
High frequency data for financials but what about company accounts?

- FX open position of NFC
- EFT, Money order and Swift information
- High Frequency credit and FX transaction on spot market
- High frequency derivative information

Big Data Platform

- Trade channel
- Liquidity channel
- FX channel

Firm behavior
The Motivation for New Datasets

- Timely monitoring of the effects of the regulations has become very important. However, delays in company reporting did not make this possible.

- Some changes had to be made in frequency and structure of reporting in order to obtain timely

- This situation required CBRT to fully adapt the reporting system to pandemic conditions. So, CBRT has changed the reporting standard from IFRS to national accounting standards, reduced the frequency to weekly, and compiled some part of reporting from administrative records (like derivatives)

- Our most prominent motivation is to prove more complete, timely and granular information as a complement to traditional macroeconomic indicators. It is important to strengthen analysis for decision-making and data is valuable only if we can manage to extract value from it
From 2018 to 2020, we watched these company’s data quarterly and after 2021, we changed the frequency to monthly. However, during the pandemic period in 2020, we requested weekly position from companies from June to December to follow the covid-19 effects on FX positions.

Weekly reportings were made in the format determined by our institution, using the values that were the basis of the legal financial statements prepared in accordance with the Accounting System Implementation General Communiqués (“MSUGT”) and the Uniform Chart of Accounts (THP) requirements published by the Ministry of Finance. CBRT had made some legal arrangements to establish the system.

Since the account balances in the form would be updated on a weekly basis, changes resulting from such timing mismatches would be reflected in the account balances in subsequent reporting. Therefore, no retrospective adjustment was required.

The weekly data sheet was not subject to independent audit.
With the weekly data form, 13 active and 12 passive items in the balance sheet of the companies were collected in USD, EUR, Other FX and TRY currencies. Data were collected via e-mails sent to official e-mail addresses. **The data collection process was carried out in the following steps:**

- First of all, an excel file template to be sent weekly was sent to all companies that have a reporting obligation. The Excel template was protected with a password, and companies were only allowed to enter data in the relevant fields. Firms were requested to fill in the form with their current account balances as of Sunday every week and send it by the following Wednesday at the latest. E-mails sent by companies were read with a developed python script, error checks are passed and tables suitable for analysis were created. The process of the Python script was as follows:
  - Separating the mails that had excel that we sent to the companies as an attachment to the mail that comes to the mailbox or that were categorized as data forms by the business unit.
  - Checking whether the excel files attached to the parsed e-mails were compatible with the template we sent. If the template was corrupted, automatic e-mail reply containing this error was returned to the relevant companies.
The New Data set- In-house data process

- If the template was correct, checking the basic information in the form (TAC, Title, reporting week, writing 0 in all fields) and returning an automatic mail reply suitable for errors (for example, missing vkn or reporting for future dates).

- Reading the error-free form with the python openpyxl module using the excel sheet-excel cell and the account code matches (eg Data Form-C25 = Cashier) and converting it to tabular data format

- Aggregating and converting the data collected for the relevant week into a single data file after completing the previous step for all forms. Combining the single data form created for the relevant week with the single data forms created for the previous weeks. At this stage, automatic confirmation e-mails were sent to the companies in order to prevent the companies from repetitive reporting for the same dates. Updating the old data with the allowed data after the manual approval process (reading the mail content) (such as the company's request to correct the old data).

- Converting the data created in wide format into a structure suitable for analysis and transferring it to the big data environment.
After the pandemic conditions relaxed, CBRT changed the reporting from weekly to monthly by integrating more complex reporting system to replace previous one. Thus, we reduced the reporting obligation on our companies from weekly to monthly and established a more sustainable data infrastructure. However, we did not want to return to quarterly reporting again, because the analyses arising from the combination of timely company data with other datasets were used extensively in upper management decisions.

In the monthly data form, different from the weekly data form, flow variables such as monthly domestic and international sales, stock purchases, tangible and financial fixed asset purchases are also included. Thus, it has made it possible to make monthly cash flow analysis of the companies. We can also provide ease of analysis by visualizing these cash flows with sankey charts.

The addition of other high frequency data such as quarterly temporary income statements and monthly VAT returns is now carried out within the scope of post-pandemic studies. It was a win to integrate company data with other high frequency databases, now researchers can work on this integrated dataset.
The monthly turnover table calculated within our Directorate every month; the main table of Revenue Administration, Company notifications transferred from Turkish Statistical Institute on a monthly basis pursuant to the Revenue Administration.

Temporary income statements are submitted to Revenue Administration once every 3 months, until the evening of the 17th day of the 2nd month following the quarterly period. The relevant data is transferred to the Turkish Statistical Institute by the Revenue Administration within 15 days, and transferred to our Bank from the Turkish Statistical Institute on the 15th of the next month.

Shopping center review with VAT and quarterly income statements
Cash Flow Analysis

Firm and sector level cash flow analysis

Cash flow analysis for an sector example

Trade channel cash flow

FX channel cash out flow
Nationality vs residency approach: measuring the impact of MNEs production structure on corporate financial statements statistics

Klaus Gerstner and Susanne Walter, Deutsche Bundesbank,
Javier Gonzalez, Santiago Martinez and Maria Garcia Riego, Bank of Spain,
Florian Resch and Stefan Kischner, Austrian Central Bank,
Alexandre Neves and Diogo Silva, Banco de Portugal

1 This presentation was prepared for the conference. The views expressed are those of the authors and do not necessarily reflect the views of the BIS, the IFC or the central banks and other institutions represented at the event.
Abstract

The organisation of global value chains of multinational enterprise groups (MNEs) poses a challenge to official statistics and the underlying concepts of presenting economic reality. At least since the „Irish Case“ the awareness of evaluating and addressing the effects of globalisation on generating official statistics has increased. Complex boundary crossing economic transactions and complex ownership structures contravene the concept of measuring national economic production. By combining information from individual and consolidated corporate financial statements and data about global group structures, we attempt to evaluate the impact and importance of globalisation on national statistics capturing economic activity of non-financial companies in selected European countries. We estimate the part of production that is attributed to the country of interest by applying traditional statistical concepts (residency approach based on country of incorporation) and compare it with the overall global production of MNEs (nationality approach based on country of control). Exploiting a newly established dataset that combines financial indicators (especially employment, value added, turnover, total assets) and information on national and international subsidiaries of European MNEs, we find differences in statistical key figures between the two approaches in all countries. We find that statistics based on nationality approach can enrich existing economic key figures but that there ongoing and consistent calculation would require harmonized microdatasets on an international level.

Keywords: globalisation; official statistics; corporate financial statements; nationality concept.
Introduction

One of the most pertinent phenomena in statistics on non-financial companies is the increasing complexity of economic transactions in the context of globalisation. In particular, the complex global ownership structures and boundary crossing intragroup transactions of multinational enterprise groups (MNE) blur the picture of national production as measured by national statistics. Statistical concepts were developed for closed economies in order to measure production that happens within the domestic boundaries (Eurostat 2017).

In particular, because of the so-called Irish Case\(^1\), official statistics have raised the awareness of effects resulting from globally active MNEs and has set up several initiatives to monitor the activities of MNEs and their impact on statistics\(^2\). In a vision paper of the committee on monetary, financial and balance of payment statistics (CMFB), the members of both systems, European Statistical System (ESS) and European System of Central Banks (ESCB), agreed on a common goal to measure and increase the visibility of MNE activities in statistics. One strategy to reach that goal is to enhance data collection, integration and usage of existing data sources. At the same time, the prevailing residency principle shall be complemented by the nationality approach as additional source of information in the statistical framework (CMFB 2020 p.7). Statistics based on the residency approach consider domestic enterprises with their individual financial statement. The nationality approach attributes the economic production and value added to the country of control rather than to the country of residence and is based on all enterprises controlled by domestic enterprises at the highest level of consolidation. Philipp Lane (2021) pinpointed it as the dichotomy between residency principle with national data collection and the global economic activities of MNEs.

Sturgeon (2013) identified in his report on measuring the global value chains the need to combine and integrate data sources to identify cross-border ownerships and the respective employment, investment and economic performance (Sturgeon 2013, p. 6) to differentiate MNE from domestic firms that do not operate globally. Although Europe hosts some of the biggest MNEs worldwide, an integrative and comprehensive database that combines the information on the group structure of MNE and the financial information of resident companies as well as of foreign subsidiaries does not exist. The data on MNEs is highly fragmented across different institutions and statistics. The divergent measurement of MNEs in these statistics can result in inconsistencies. An integration of the underlying microdata can help to understand and resolve these inconsistencies. Moreover, the overall effect of MNEs on several statistics in the area of non-financial companies has hardly been explored.

Following other initiatives that focus on understanding the statistical footprint of MNEs in national business statistics, the European Committee of Central Balance Sheet Data Offices (ECCBSO\(^3\)) installed an international Task Force at the end of 2020\(^4\). The aim of the Task Force was to analyse the impact of globalised enterprise groups on national financial statements statistics of non-financial companies. Set up as feasibility study, the member countries (Portugal, Italy, Spain, Austria, Greece, and Germany) evaluated whether the approach can be applied to the existing data and how the different measurement concepts affect the results. The works of the Task Force were finalised in October 2021. This paper presents the results of the cross-country analyses and data description.

In a first step, it was necessary to integrate different data sources on MNEs and to estimate their impact as well as their economic integration in the national and global economy. The integration of data sources on the structure of MNEs, their subsidiaries as well as their economic performance allows us to calculate

---

\(^1\) In the course of the reallocation of intellectual property within a multinational enterprise group from one economy to another, the Irish GDP has experienced a drastic shift upwards in 2016 that has attracted the attention of international economists and the press (Allafi et a. 2017).

\(^2\) Setting up Large Cases Units that check for consistency in statistical reporting with regard to MNEs (Connolly 2011, Ahlborn et al. 2021) or analysing relevant cases for the calculation of National Accounts and setting up an Early Warning System (Alajäskö et al. 2018).

\(^3\) European Committee of Central Balance Sheet Data Offices (www.eccbso.org)

\(^4\) The Task Force acted as a subgroup of the European Records of IFRS Consolidated Accounts working group (ERICA WG) (https://www.eccbso.org/wba/working-groups/erica-working-group).
statistics based on the nationality approach and compare it to the status quo statistics based on the residency principle. We combine data on consolidated financial statements (FSs) with individual FSs and data on the group structure to identify all legal entities lying within the perimeter of the group. In doing so, we are able to analyse and remove several effects that MNE reporting has on the statistics in our study: the effect of intragroup flows (consolidated vs individual FSs), the effect of missing reporting on the level of the individual FSs as well as the pure economic effect of applying a different approach to capture and attribute economic performance (concept of residence vs concept of ownership).

We combine micro data sources from the ESCB, namely FDI statistics, statistics on consolidated and individual FSs of non-financial companies with data from commercial registries, national business registers and the Euro Groups Register (EGR) from the ESS. We analyse four selected economic indicators (employment, sales, total assets and value added) and compute them according to the nationality approach. We find that the net effect of applying the nationality principle in almost all indicators is positive for the majority of the participating countries. Thus, the economic performance of domestic MNEs abroad outweighs in most cases the contribution of foreign resident companies as well as effects of consolidation of intragroup flows. These findings go in hand with other globalisation indicators from official statistics. We infer that statistics based on the nationality approach can enrich existing economic key figures.

The structure of the paper is as follows. The methodological section provides an overview on the transforming steps from the resident to the nationality approach. After describing the steps on data integration, we present the results and discuss them. The paper concludes with a summary, puts the results into a broader context, discusses limitations of the study and provides an outlook on further research avenues.

Methodology

Several institutions and different departments in the same institution prepare statistics that cover various aspects of an enterprise group. We understand enterprise groups according to the Council Regulation (EEC) 696/1993, as an association of enterprises bound together by legal and/or financial links. A multinational enterprise group characterizes as an association of enterprises that spans the borders of more than one nation. As a result, each single statistic captures different characteristics of the enterprise group (EG). Moreover, there is neither a common unique company identifier nor a unique identifier for an enterprise group that connects all these datasets. This concept of group is thus wider than the concept of groups typically applied in generally accepted accounting principles and thus not for every enterprise group a consolidated FS is prepared.

To gain a holistic view on enterprise groups and their effects on business statistics (here financial reporting), all relevant data sources have to be connected and compiled. To run the desired comparative analyses, the information on individual and consolidated balance sheets and income statements, on relationships between enterprises and their inclusion in the consolidated FSs need to be linked.

The participating countries face different starting conditions and heterogeneous data landscapes to conduct this study. While some countries could build on already integrated datasets with almost full coverage of the non-financial company sector (e.g., Portugal\(^5\)), other countries had to access and integrate several formerly unlinked datasets (e.g., Germany).

Since there is no single European dataset that captures all aspects of MNEs, each member had to implement the analyses for their own country. Aggregates and indices were then compared given the heterogeneous data and the varying coverage. For group structures, the EuroGroupsRegister (EGR) from Eurostat was used as the only European harmonized source. Independently of the data situation in each country, merging datasets was still necessary in every case because data from the ESS had to be linked to the data already available in the participating central banks\(^6\).

\(^5\) See Pinto et al. (2018) for comprehensive description of the Portuguese database.

\(^6\) Detailed information on matching quotas and quality can be provided upon request.
In this paper, we analyse a cross section of companies and combine the datasets for the reporting year 2018. This avoids time variations in the key financial indicators that are solely due to dynamic changes in the perimeter of the enterprise groups. Our analytical framework covers the steps depicted in Table 1. Steps 1 to 4 will be addressed in section 2 while step 5 is covered in section 3.

### Analytical steps

<table>
<thead>
<tr>
<th>1) Identify datasets</th>
<th>2) Merge datasets</th>
<th>3) Classify enterprises</th>
<th>4) Calculate results</th>
<th>5) Compare approaches</th>
</tr>
</thead>
<tbody>
<tr>
<td>Access relevant datasets, as information is allocated in numerous datasets.</td>
<td>Match datasets using common identifiers or, if not available, string based matching.</td>
<td>Identify enterprises that belong to a group and their country of control</td>
<td>Use all inputs to calculate the nationality approach.</td>
<td>Investigate the differences between nationality and residency approach.</td>
</tr>
</tbody>
</table>

#### Step 1: Identify datasets

Table 2 provides an overview over the relevant datasets used in every country for this study. It shows that several data sources had to be accessed and interlinked. The EGR contains multinational enterprise groups with at least one subsidiary abroad whereas the national business register covers all group companies with domestic residence. FDI data provides information on foreign controlled companies residing in the national territory and domestically controlled subsidiaries abroad. Analytical datasets used to calculate indicators according to the nationality approach are the data on consolidated financial FSs, that group heads resident domestically are publishing. Financial information of the individual company is extracted from the statistics on individual financial FSs.

### Relevant data used

<table>
<thead>
<tr>
<th>Country</th>
<th>Group data</th>
<th>Financial data</th>
<th>Other Sources</th>
</tr>
</thead>
<tbody>
<tr>
<td>Austria</td>
<td>FDI X, National Business Register X</td>
<td>EGR Individual Financial Statements X, Consolidated Financial Statements X</td>
<td>Internal dataset for group structure</td>
</tr>
<tr>
<td>Germany</td>
<td>FDI X&lt;sup&gt;9&lt;/sup&gt;, National Business Register X&lt;sup&gt;10&lt;/sup&gt;</td>
<td>EGR Individual Financial Statements X&lt;sup&gt;11&lt;/sup&gt;, Consolidated Financial Statements X&lt;sup&gt;12&lt;/sup&gt;</td>
<td>Commercial data</td>
</tr>
<tr>
<td>Portugal</td>
<td>FDI X, National Business Register X</td>
<td>EGR Individual Financial Statements X, Consolidated Financial Statements X</td>
<td>National groups database</td>
</tr>
<tr>
<td>Spain</td>
<td>FDI X&lt;sup&gt;13&lt;/sup&gt;, National Business Register X</td>
<td>EGR Individual Financial Statements X, Consolidated Financial Statements X</td>
<td>Internal dataset for group structure</td>
</tr>
</tbody>
</table>

---

7 In the Portuguese case the consolidated FSs were corrected for temporal changes in the perimeter of the group.
8 Also subgroup heads may publish a consolidated FS when they are active on a capital market. They were identified and excluded from the analyses.
9 Microdatabase Direct investment (MiDi – “Mikrodatenbank Direktinvestitionen”) | Deutsche Bundesbank (DOI:10.12757/BBk.MiDi.9919.07.08)
10 Dataset is not available to externals yet, ERICA database on microlevel available for participating NCBs.
13 In the Spanish case EGR was used only for comparison as the own internal group structure dataset was prioritized. It is a process that is controlled and goes hand in hand with the consolidated and individual financial accounts received.
Step 2: Merge datasets

Even though there is no common unique company identifier across all datasets, the existence of national identifiers such as the number of the commercial register or the VAT or tax identification number facilitates merging as datasets from different sources. On this basis, it was feasible to link the EGR with balance sheet data from the central banks. In some cases, it was necessary to link the EGR to the national business register and then link it to the balance sheet data with the help of VAT number or number from the commercial register. Due to inconsistencies in the reported commercial register number, we use company-linking tables provided by the data labs as an additional source. In case no common identifier was available, string-based matching was applied.

Step 3: Classify enterprises

After having identified and merged the relevant datasets, the various types of EGs had to be identified as a foundation for excluding economic activity stemming from foreign controlled MNEs and including activity from domestically controlled enterprises resident abroad. Table 3 shows the four types of EGs we use in this study. Each enterprise was allocated to one type.

<table>
<thead>
<tr>
<th>Types of enterprise groups</th>
<th>Non-group enterprise</th>
<th>All resident EG</th>
<th>Domestic MNE</th>
<th>Foreign MNE</th>
</tr>
</thead>
<tbody>
<tr>
<td>Enterprise does not belong to any enterprise group</td>
<td>All group enterprises reside in the domestic country, including the group head</td>
<td>At least one subsidiary resides in a foreign country, but the group head is resident in domestic country</td>
<td>At least one subsidiary resides in domestic country, but the group head is resident in a foreign country</td>
<td></td>
</tr>
</tbody>
</table>

To classify enterprises into the group types, we applied a top-down approach starting from the consolidated FSs of the group head, enriched with either information from internal group structure databases, EGR and data from national business registers. These datasets include information on the residency of the global group head (GGH). Domestic MNEs are defined as EGs with a GGH resident in domestic country and foreign MNEs as EG with a GGH resident outside. EGs not covered by EGR, with no foreign GGH or not marked as non-group enterprise in the national business register are identified as all-resident EGs. Enterprises that have not matched with EGR and are not attributed to any group in the national business register or internal group database are labelled as non-group enterprises.

Step 4: Calculate results

For the residency approach the sample comprises all enterprises which are resident domestically, i.e. all non-group EG, enterprises belonging to all-resident EG and the domestic subsidiaries of domestic MNEs and of foreign MNEs. We do not account for consolidation effects and consider each enterprise with its individual financial statement.

For the nationality approach the sample comprises all enterprises which are controlled domestically, i.e. all non-group EG, all resident EG and all domestic MNEs including their foreign group members. All enterprises are considered only once at the highest level of consolidation, i.e., sub-consolidated FSs and individual FSs of enterprises belonging to a consolidated group are not included.

---

14 For instance, for Deutsche Bundesbank the linktables from the Data Service Centre were used. See Data Report on Record Linkage p.7 (https://www.bundesbank.de/resource/blob/624432/207c774d468e82d76ec19ef6bfa1c8a7/ml/2021-05-company-data.pdf)
Results

The aim of the study is to analyse the relevance of multinational enterprise groups on the statistics on corporate financial reports. Before drawing inferences on the impact of MNEs, it is necessary to explore the

The relevance of enterprise groups for individual financial accounts statistics per country

<table>
<thead>
<tr>
<th>Country</th>
<th>N° of companies</th>
<th>Employees</th>
<th>Turnover</th>
</tr>
</thead>
<tbody>
<tr>
<td>Austria</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Domestic</td>
<td>14,0%</td>
<td>49,0%</td>
<td>49,0%</td>
</tr>
<tr>
<td>Foreign</td>
<td>7,0%</td>
<td>17,0%</td>
<td>27,0%</td>
</tr>
<tr>
<td>No group</td>
<td>79,0%</td>
<td>34,0%</td>
<td>24,0%</td>
</tr>
<tr>
<td>Germany</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Domestic</td>
<td>10,5%</td>
<td>46,7%</td>
<td>42,8%</td>
</tr>
<tr>
<td>Foreign</td>
<td>1,3%</td>
<td>4,1%</td>
<td>7,1%</td>
</tr>
<tr>
<td>No group</td>
<td>88,3%</td>
<td>49,3%</td>
<td>50,1%</td>
</tr>
<tr>
<td>Portugal</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Domestic</td>
<td>1.2%</td>
<td>15.5%</td>
<td>32.1%</td>
</tr>
<tr>
<td>Foreign</td>
<td>1.4%</td>
<td>13.6%</td>
<td>25.7%</td>
</tr>
<tr>
<td>No group</td>
<td>97.4%</td>
<td>70.9%</td>
<td>42.2%</td>
</tr>
<tr>
<td>Spain</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Domestic</td>
<td>5.9%</td>
<td>29.3%</td>
<td>37.2%</td>
</tr>
<tr>
<td>Foreign</td>
<td>2.3%</td>
<td>17.8%</td>
<td>28.6%</td>
</tr>
<tr>
<td>No group</td>
<td>91.8%</td>
<td>52.9%</td>
<td>34.2%</td>
</tr>
</tbody>
</table>
number and characteristics of the enterprises that are covered by the statistics on corporate financial reports. Table 4 shows the relevance of enterprise groups for statistics based on individual FSs by summarizing the main economic variables such as employees, turnover and number of companies per group type. Across all countries, only a small proportion of FSs (3% to 21%) are prepared by enterprises belonging to a group. For all countries, enterprises belonging to domestically controlled groups make up a larger and more relevant part of the enterprise population as compared to enterprises belonging to foreign MNEs. However, enterprises belonging to groups account for more than half of the turnover and employees in each country. In Portugal and Austria, foreign MNEs generate 1/3 of the turnover indicated in the financial statements. Despite their small number, group enterprises play an important role for the national economies and their official statistical representation. In the case of Germany, the sample of the individual FSs data seems to be representative for the total population as the distribution of companies over group types is similar to the distribution in the national business register.

With a focus on coverage, Portugal and Spain have an almost complete sample, while Austria and Germany in comparison cover fewer companies. Table 5 shows the data coverage for each country with respect to employees and turnover (compared to the overall company population). However, the samples are upward biased as the relative number of included companies is low but they make up for a significant amount of value added (as measured by employees and turnover, 1/3 in the case of Austria and 1/3 of employees and more than 2/3 of the turnover in Germany). Due to this heterogeneity in the data coverage and access, the comparative analyses were carried out using indices instead of absolute numbers.

Data Coverage

<table>
<thead>
<tr>
<th>Country</th>
<th>Coverage (Financial reports compared to overall population)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Employees (2018) in %</td>
</tr>
<tr>
<td>Austria</td>
<td>19</td>
</tr>
<tr>
<td>Germany</td>
<td>34</td>
</tr>
<tr>
<td>Portugal</td>
<td>100</td>
</tr>
<tr>
<td>Spain</td>
<td>76</td>
</tr>
</tbody>
</table>

The comparison of the nationality and residency approach is done by an index representing the ratio between the nationality and the residency aggregate in percent. An index above 100 reveals that the global activity of domestic MNEs outweighs the activity measured by the residency statistics. Figure 2 shows the results for four main economic indicators based on the residency and the nationality approach. The number of FSs, the number of employees, the turnover, total assets and the value added were chosen as these indicators are widely used (e.g. for size classifications) and were available in different dataset for comparison. The starting point on the left hand side is the residency approach, which is always 100 by definition. The right hand side shows the index value for each country under the nationality approach.

---

15 The Bundesbank to this point only publishes statistics on yearly financial statements of individual companies as well as half-yearly consolidated financial reports of enterprise groups that are active on the capital market. The yearly consolidated FSs included in these analyses are so far not published by Bundesbank.

16 Value added was not available in the FDI dataset. In case of missing data, value added was approximated by profit or loss.
Overall, for three countries (Spain, Austria and Germany) the net effect of applying the nationality approach is positive, indicating that residency statistics do not cover the full picture of the activity of MNEs. The observed effect was strongest for the number of employees, turnover and value added.

For Portugal, the estimates for total turnover, value added, assets and employees of non-financial corporations are lower when we compare the nationality approach with the residency approach. This dynamic applies to the four indicators that we have computed, which signals consistency. Statistics based on nationality approach have 88.6%-92.0% of the magnitude of the statistics in accordance with the residency approach. The main explanation for this outcome is that foreign MNEs invest more in Portugal than domestics MNEs invest in the rest of the world. This is a long-term feature of the Portuguese economy. An independent statistic that displays this characteristic is foreign direct investment, since foreign direct investment liabilities are systematically higher than foreign direct investment assets. In the nationality approach, we exclude domestic enterprises controlled by foreign MNEs. The foreign MNEs with the highest contribution to the statistics based on residency approach (excluded from the nationality approach) are domiciled in Spain and France. This is in line with foreign direct investment statistics, which indicate that the main ultimate investors in the Portuguese economy are from Spain and France. In addition, there is full coverage of Portuguese non-financial corporations, which ensures that the results are not affected by selection bias.

Compared to Portugal with the best coverage of the enterprise sample, the analyses in the other countries can only be interpreted against the background of the specific data availabilities and limitations.

The Austrian results do not cover all solo financial statements (FSs) nor do all consolidate FSs as well as not all (foreign) group structures. The results for all three indicators suggest that Austrian groups are more engaged abroad than foreign groups in Austria. However, given the data gaps (especially for subsidiaries of foreign groups in Austria); this effect might be overestimated (compared to Portugal and Spain).
For Germany, the number of employees triples in the nationality approach compared to the residency approach. Multiple effects drive this result: First, due to the legal publication obligation the number of individual FSs is relatively low. The missing data could be imputed by consolidated FSs. The sharp increase in employees is also partially an artefact from the moderate share of missing data in employee numbers. However, the employee numbers were also corrected for double counting (in cases when an individual and a consolidated FS were existent) and employees of foreign controlled groups were deducted. Even so, the inclusion of consolidated FSs and the employees of domestic MNEs working in foreign subsidiaries outweighed this deduction of employees.

Overall, German data faces similar limitations as Austrian data. However, the results of this study are in line with a case study conducted by the authors for the 30 largest enterprise groups active on the capital market. The financial reports of the companies were analysed to explore the integration of MNE into the German economy. On average 30% of the activity of the MNEs (with respect to employee share, number of subsidiaries, turnover) took place on the German territory, while the remaining 70% was undertaken abroad. The net effect for total assets, turnover and value added was also positive, even though to a lesser extent. One explanation for a lower outcome regarding these variables is that these numbers are non-additive, meaning that inter-group transactions are not included while the number of employees is additive. Overall, the results point to the same direction similar to other published indicators on the global integration of the German economy: the German business structure is characterised by a small share of MNEs that are of critical importance for national statistics and that are only partially integrated in the national economy and generate an enormous amount of production globally.

In Spain employing the nationality approach leads to an increase in employees, turnover and value added. These findings correspond with the fact that there are very important Spanish MNE with subsidiaries abroad (particularly relevant on the American continent). The importance of foreign subsidiaries of domestic MNEs outweighs the economic contribution of domestic subsidiaries of foreign MNEs. However, the amount of total assets drops in the Spanish case when the nationality approach is applied. This is because eliminations (particularly in investment-equity and intercompany transactions) have taken place during the consolidation process in the nationality approach. The Spanish results need to be interpreted with respect to the limitations of the Spanish data sample. The data are not a representation of the full population and might be biased. In particular, the sample may not be representative of the breakdown of all the sectors of the population.

Conclusion

The aim of this study was to shed light on the rarely explored effect of measuring the impact of MNEs production structure on corporate financial statement statistics. This project brings together numerous official initiatives that strive to understand and depict the impact of globalisation on statistics that are mainly rooted in residency-based concepts. By combining heterogeneous datasets that, we calculated primary economic aggregates based on the nationality principle that attributes economic outcomes to the country of control rather than the country of residence while also considering consolidation effects. We compared the effects of applying this alternative approach to the currently used residency approach which is based on individual FSs of domestically resident enterprises.

We find that the overall net effect of calculating the figures according to the nationality principle is positive for countries that are highly embedded in global transactions and negative for countries with a more inward oriented economy. It becomes apparent that for the majority of countries in this study a non-negligible amount of economic activity of domestically controlled enterprise groups is carried out outside the national borders and not captured by statistics using only data from domestically resident enterprises. Thus, the residency approach leads to an underestimation of economic activity in those countries as domestic MNEs are generating a high proportion of value added abroad.

Companies that are already included in a published consolidated FS and/or fall below certain size thresholds are generally exempted from the obligation to publish their individual FS.
However, both approaches have their legitimization, and the usefulness of their application depends highly on the desired purpose and claims of the statistics under exploration. As statistics based on the nationality approach are not widely published by official authorities\textsuperscript{18}, the users of statistical products could benefit from expanding the published indicators by alternative indicators calculated based on nationality principles (as proposed by the CMFB 2020).

The results of this study represent a first step towards capturing the impact of complex globally structured enterprise groups. As a starting point for further research, a repetition of our analysis with a more comprehensive data sample and thus higher or full coverage of companies of FSs might underpin the findings in this paper. Next, a generalization of our results (to other years, countries, indicators) could be explored by further studies. Furthermore, there is a whole set of business statistics, especially trade statistics and statistics in ESS, that might be combined with group data in an equal fashion to link information on ownership with patterns of international trade.

Besides the analytical lessons learned, our study also revealed room for improvement in the European organisation of microdata on MNEs. The accessible data is very heterogeneous among member countries in terms of coverage (population and variables), application of reporting standards and definition of variables and concepts. Furthermore, the cross-country comparison has revealed that the full data integration required substantial effort to conduct analyses with sufficient quality. Throughout the study important differences between EGR and the internal group data bases became apparent. These caveats limit the interpretability and reliability our results.

A successful application of the nationality approach depends strongly on the available data, especially individual FSs. A harmonized European micro dataset that combines comprehensive financial information of companies and groups with the respective information on the group structure would help reducing inconsistencies and keeping better track of the overall footprint of MNEs. A shared methodology to make datasets more coherent (harmonized) among all the directories/datasets (groups, companies and group structure) could tackle this problem. Also, the quality of the group structure’s depiction would benefit if all groups reported the direct control of one company over another on a digital basis. Furthermore, a common unique international company and enterprise group identifier would allow to also connect this harmonized microdatabase with other databases necessary for specific analytical goals and would reduce the effort of interlinking datasets that contain the same entities. Additionally, initiatives for connecting registers (like BRIS) and data sharing (G20 DGI) increase the potential to grasp the overall footprint of multinationals in European/ global statistics and to reduce inconsistencies in data.

\textsuperscript{18} The BIS is one of the institutions that publishes statistics based on the nationality approach.
References


Lane, P. (2021), Maximizing the user value of statistics: lessons from globalisation and the pandemic, Speech at the European Statistical Forum.


Nationality vs. Residency Approach

Measuring the impact of MNEs production structure on corporate financial statements statistics

25th August 2022

BIS – IFC 11th biennial conference
Motivation

- The “Irish Case” has raised awareness on the “dichotomy” of the residency principle of national statistics and the global economic activities of mutinational enterprise groups (MNEs).
- Numerous initiatives have been set up to measure globalisation effects and the footprint of MNEs in Official Statistics.
- The committee on monetary, financial and balance of payment statistics (CMFB) has emphasized that traditional statistics (country of location) can be complemented by the Nationality approach (country of control) as additional source of information.
- An international task force was created to conduct a feasibility study to integrate microdata on corporate financial statements and group structures from ESCB and ESS (European Statistical System), to generate statistics according to the nationality approach and compare it with traditional statistics (residency approach)
Concepts

• **Residency approach:**
  • Comprises domestic entities with their individual financial statement
  • This approach does not consider whether an entity is under foreign control or part of a group.

• **Nationality approach:**
  • Comprises all entities controlled by resident entities at the highest level of consolidation
  • Differences compared to the residency approach may arise from consolidation effects and from foreign control.
Nationality vs. Residency Approach

Entities

- **Group affiliation**
  - Group enterprise
  - Non-group enterprise

- **Residency of group mother**
  - Domestic
  - Foreign

- **Residency of all group firms**
  - All-resident EG
  - Domestic MNE
  - Foreign MNE

**Control ≠ Consolidation**
### Concepts

<table>
<thead>
<tr>
<th>Entity is</th>
<th>controlled</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>domestic</td>
<td>foreign</td>
</tr>
<tr>
<td>resident</td>
<td>All-resident EG/ Domestic MNE</td>
<td>Foreign MNE</td>
</tr>
<tr>
<td>foreign</td>
<td>Domestic MNE</td>
<td>Foreign MNE</td>
</tr>
</tbody>
</table>

### Nationality approach:
Comprises all entities controlled by resident legal entities at the highest level of consolidation (no double counting)

### Residency approach:
Comprises domestic legal entities with their individual financial statement

### Differences:
May arise from consolidation effects and from foreign control
Method

(1) **Identify** relevant **datasets**

(2) **Merge** them

(3) **Identify** enterprises that belong to a **group** and their **country of control**

(4) **Correct for double counting** (consider an entity only once) and remove sub-consolidated accounts

(5) **Calculate** nationality approach

(6) **Compare** to residency approach
## Data

<table>
<thead>
<tr>
<th>Country</th>
<th>Group data</th>
<th>Financial data</th>
<th>Other sources</th>
<th>Coverage</th>
</tr>
</thead>
<tbody>
<tr>
<td>Austria</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
</tr>
<tr>
<td>Germany</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
</tr>
<tr>
<td>Portugal</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
</tr>
<tr>
<td>Spain</td>
<td>X (EGR – FATS)</td>
<td>X</td>
<td>X</td>
<td>X</td>
</tr>
</tbody>
</table>
## Results

### Nationality vs. Residency Approach

<table>
<thead>
<tr>
<th>Country</th>
<th>Austria</th>
<th>Germany</th>
<th>Portugal</th>
<th>Spain</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>No. of companies</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Domestic groups</td>
<td>Foreign groups</td>
<td>No group firms</td>
<td>Domestic groups</td>
<td>Foreign groups</td>
</tr>
<tr>
<td>No group firms 79.0%</td>
<td>Domestic groups 14.0%</td>
<td>Foreign groups 7.0%</td>
<td>Domestic groups 88.2%</td>
<td>Foreign groups 1.3%</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Group Type</th>
<th>Domestic groups</th>
<th>Foreign groups</th>
<th>No group firms</th>
<th>Domestic groups</th>
<th>Foreign groups</th>
<th>No group firms</th>
<th>Domestic groups</th>
<th>Foreign groups</th>
<th>No group firms</th>
<th>Domestic groups</th>
<th>Foreign groups</th>
<th>No group firms</th>
</tr>
</thead>
<tbody>
<tr>
<td>Employees</td>
<td>49.0%</td>
<td>17.0%</td>
<td>34.0%</td>
<td>46.7%</td>
<td>4.1%</td>
<td>49.3%</td>
<td>15.5%</td>
<td>13.6%</td>
<td>70.9%</td>
<td>29.3%</td>
<td>17.8%</td>
<td>52.9%</td>
</tr>
<tr>
<td>Turnover</td>
<td>49.0%</td>
<td>27.0%</td>
<td>24.0%</td>
<td>42.8%</td>
<td>7.1%</td>
<td>50.1%</td>
<td>32.1%</td>
<td>25.7%</td>
<td>42.2%</td>
<td>37.2%</td>
<td>28.6%</td>
<td>34.2%</td>
</tr>
</tbody>
</table>

- **Group Firms:** Few entities, although highly relevant
- **Country peculiarities**
Results

- Positive effect for Austria, Spain and Germany
- Negative effect for Portugal

<table>
<thead>
<tr>
<th>Nationality</th>
<th>Employees</th>
<th>Turnover</th>
<th>Value added</th>
<th>Total Assets</th>
</tr>
</thead>
<tbody>
<tr>
<td>Germany</td>
<td>118</td>
<td>146</td>
<td>159</td>
<td>104</td>
</tr>
<tr>
<td>Portugal</td>
<td>89</td>
<td>90</td>
<td>89</td>
<td>89</td>
</tr>
<tr>
<td>Austria</td>
<td>184</td>
<td>138</td>
<td>159</td>
<td>89</td>
</tr>
<tr>
<td>Spain</td>
<td>118</td>
<td>122</td>
<td>122</td>
<td>96</td>
</tr>
</tbody>
</table>

Bar charts showing:
- Employees: Germany 310, Portugal 92, Austria 184, Spain 118
- Turnover: Germany 118, Portugal 92, Austria 138, Spain 122
- Value added: Germany 104, Portugal 89, Austria 159, Spain 122
- Total Assets: Germany 104, Portugal 89, Austria 159, Spain 96
The Nationality and Residency Approach have their legitimation. The usefulness of application depends on the purpose.

Prerequisites for the calculation are interconnected standardized micro datasets and a harmonized methodology to obtain a comprehensive picture of MNEs’ activities.

A full data integration requires substantial efforts to conduct analyses with sufficient quality.

Globalised MNEs are small in number (as compared to overall population) but highly relevant for statistical aggregates.

Users of statistical products could benefit from expanding the published indicators by alternative indicators calculated based on the nationality approach.
THANK YOU FOR YOUR ATTENTION

Susanne Walter, Deutsche Bundesbank
susanne.walter@bundesbank.de

BIS – IFC 11th biennial conference
Nationality vs. Residency Approach

Backup
Nationality vs. Residency Approach

**Method**

**Residency approach**
- Individual accounts no group
- Individual accounts All-resident groups
- Individual accounts domestic MNE
- Individual accounts foreign MNE

**Consolidated accounts**
- (- subconsolidated, - foreign MNE)

**Nationality approach**
- Foreign subsidiaries domestic MNE (non cons)

**Consolidated accounts**
- All-resident domestic MNE
- Individual accounts no group
- Individual accounts All-resident and domestic MNE (non cons)

**Individual accounts**
- (- consolidated, - foreign MNE)
Challenges in the external statistics framework: how to register MNE financial restructuring operations\textsuperscript{1}

Nadia Accoto, Giuseppina Marocchi and Silvia Sabatini,
Bank of Italy

\textsuperscript{1} This presentation was prepared for the conference. The views expressed are those of the authors and do not necessarily reflect the views of the BIS, the IFC or the central banks and other institutions represented at the event.
Challenges in the external statistics framework: how to register MNE financial restructuring operations

Nadia Accoto, Giuseppina Marocchi and Silvia Sabatini

Abstract

The rapid development of a globalised world has drastically increased the number, the complexity and the variety of possible intercompany operations. More and more frequently companies are involved in corporate inversions or decide to move their domicile to other countries in order to benefit from more advantageous fiscal and company management conditions.

These operations imply relevant problems in compiling external statistics, particularly foreign direct investment flows and stocks but also other balance of payments items. The purpose of the work is to highlight these challenges through the analysis of some real cases observed in Italy.

Keywords: multinational enterprises, external statistics, foreign direct investment

1 The views expressed in the article are those of the authors and do not involve the responsibility of the Bank of Italy.
Contents

Challenges in the external statistics framework: how to register MNE financial restructuring operations ............................................................................................................... 1

1. Introduction ....................................................................................................................................... 3

The Italian BOP/IIP Direct Reporting system ................................................................................ 4

2. The re-domiciliation of the holding company of a group ............................................................. 5
   2.1 General description and impacts on external statistics of the financial restructuring of the Davide Campari group ............................................................... 5
   2.2 Another type of re-domiciliation: the case of Mediaset Spa ................................................ 8
   2.3 Main challenges ..................................................................................................................... 10

3. Transfer of participations within a group: the Enel case ............................................................. 11
   3.1 General description and impacts on external statistics ......................................................... 11
   3.2 Main challenges ..................................................................................................................... 13

4. Conclusions ......................................................................................................................................... 13

References ................................................................................................................................................ 15
1. Introduction

The rapid development of a globalised world has drastically increased the number, the complexity and the variety of possible corporate restructuring operations that are reflected on external statistics. How to correctly capture and include these processes in the Balance of Payments/International Investment Position (BOP/IIP) framework has become one of the most important subject in many international fora.

In the second half of 2020, the process of update of the IMF Balance of Payments and International Investment Position manual (BPM6) began. Particular attention was paid to the analysis of corporate restructuring operations, in order to provide compilers with detailed instructions on how to correctly record these processes in the external statistics framework with the aim to harmonize compilation practices among countries and favor the reduction of possible bilateral asymmetries.

The paper intends to continue the analysis of the challenges posed to the external statistics compilers by the increasingly frequent and complex corporate restructuring, firstly presented in Lisbon in February 2020 during the IFC Conference on external statistics “Bridging measurement challenges and analytical needs of external statistics: evolution or revolution”.

On that occasion, two significant real cases that had important effects on Italy’s BOP/IIP were presented: (1) a cross-border merger and the subsequent creation of a branch (located and resident) in Italy and (2) the acquisition of an entire big Italian group through the establishment of a chain of newly created companies.

This paper deals with some further cases of corporate restructuring, occurred in Italy in recent years, which involved important Italian groups and led to some further challenges in recording the effects in the official statistics. The paper also intends to provide new examples in order to contribute to the ongoing discussion on defining specific guidelines in the new manuals about how to correctly and uniquely register these operations in the context of external statistics.

Chapter 2 analyses two different cases of re-domiciliation in the Netherlands of the holding company of two relevant Italian groups (Davide Campari and Mediaset) highlighting the different features of the branches created after the change of residence. Chapter 3 describes another type of restructuring related to the transfer of participations and other assets within enterprises of a same group. Chapter 4 summarises the results of the analysis and tries to point out the main challenges and possible solutions.

For each case, the main events of the operation are described, underlying the impacts on external statistics and focusing on the difficulties faced by compilers. The discussion concentrates only on the financial account of the BOP and on the IIP, particularly on Foreign Direct Investment (FDI) and Portfolio Investment. Similar and

---

2 I.e. cross-border mergers, re-domiciliation of a company, transfer of participation between group entities

3 At the moment of writing, corporate restructuring is being analyzed in different contexts, but only one document is currently disclosed: “D4 Corporate Inversions” (https://www.imf.org/-/media/Files/Data/Statistics/BPM6/DITT/d4-corporate-inversions.ashx)

4 https://www.bis.org/ifc/publ/ifcb52_20.pdf (Lisbon paper)
related challenges, not examined in the paper, are faced also for the possible impact of the operations on the current account of the BOP (investment income item).

The circumstances described in these three cases are derived from public documents available on the websites of the involved companies. Reported amounts are fictitious, without any relationships with real data, in order to preserve their confidentiality.

The Italian BOP/IIP Direct reporting system

The Italian data collection system is mainly based on direct reporting by entities involved in cross-border transactions.

Direct reporting entities include all the resident banks, all other financial intermediaries, and the central bank, for their own transactions and positions, and a sample of non-financial and insurance companies.

With reference to resident banks and other financial intermediaries, data are collected, on a census basis, within the framework of supervisory reporting, through specific BOP statistical reports inserted ad hoc.

The non-financial and insurance company Direct reporting system consists of a set of sample surveys covering specific non-financial and financial transactions and positions and varying in frequency from monthly to annual, depending on the investigated phenomena.

The total sample covers about 7,000 enterprises; the sample selection makes use of a BOP-specific business register (with a flag for the presence of FDI relationships) and it is based on the size of enterprises and on their geographical location on the Italian territory. The most relevant enterprises in terms of dimension are included on a census basis.

This system offers, as the main advantage, the possibility of establishing a direct contact with the reporting agents that allows compilers to investigate the more complex aspects of some important transactions that involve reporting companies.

In fact, for instance, when a relevant event is reported by the media, the BOP/IIP compilers can:

- contact the units directly or indirectly involved for further explanations on the event, in order to: clearly outline the process, obtain the necessary details (including real amounts), correctly record in the external statistics all the relevant aspects (functional category and instruments such as equity, bonds and loans);\(^1\)

- evaluate the possibility to insert in the sample newly established companies.

The direct link with the reporting agent also gives the possibility to know in advance about relevant operations, even when they do not jump to the press headlines or are not reported by commercial data providers.

Another relevant advantage of Direct reporting is related to the flexibility in defining the requested information: the monthly questionnaire has been recently modified to introduce new questions allowing the reporting of participations/shareholders emerged/disappeared after financial restructuring operations and, in the annual questionnaire, it is now asked to separately report participation in branches.

\(^1\) In the cases analysed in this paper, the frequent informal contacts with the reporting agent gave a prompt and complete picture of all the most relevant aspects of the processes.
2. The re-domiciliation of the holding company of a group

2.1 General description and impacts on external statistics of the financial restructuring of the Davide Campari group

During the period 2019-2020 a corporate restructuring involved Davide Campari, one of the world most important food and beverage groups located in Italy. The process developed in several steps and ended up in the re-domiciliation in the Netherlands of the holding company of the group, Davide Campari-Milano S.p.A (Campari).

Before the restructuring, Campari was an Italian company, listed on Milan stock exchange, with two direct investors (Alicros Spa, an Italian unlisted enterprise, and Cedar Rock, an investment fund domiciled in the UK) and other floating shareholders (Figure 1).

The corporate restructuring developed in three stages:

- a **first phase**, occurred in February 2019, during which Lagfin SCA first purchased the remaining participation in Alicros Spa from the other shareholder and
subsequently incorporates Alicros itself, thus becoming the new foreign direct investor of Campari;\textsuperscript{5}

- a second phase, occurred in July 2020, in which Campari, the Italian holding company of the group, moved its domicile to the Netherlands changing its name in Davide Campari Milano NV;\textsuperscript{6}

- a third phase, simultaneous but logically distinct from the second one, in which a permanent establishment in Italy (a branch) was created to carry out, without interruption, the activities performed by the old enterprise Campari.

Figure 2 summarizes the ownership structure after the corporate restructuring phases.

The articulated process produced several effects, both on the BOP and on the IIP. In order to outline them, let us suppose (see Figure 3) that, before the phase 1 started, the total equity investment in Alicros Spa was equal to 370 Euros and the total equity investment in Campari was equal to 4250 Euros. Furthermore, the FDI assets of Campari were equal to 3600 Euros.

\textsuperscript{5} As stated in the Lagfin SCA website (https://www.lagfin.lu/#Structure)

\textsuperscript{6} More detailed information on the transaction aimed at the transfer of the registered office of Davide Campari-Milano S.p.A. can be found in the Campari website: https://www.camparigroup.com/en/page/investors/transfer-registered-office
### Equity investment in Alicros Spa and Campari before the merger

<table>
<thead>
<tr>
<th>Equity investment in Alicros SpA</th>
<th>Amount (Euros)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Total</td>
<td>370</td>
</tr>
<tr>
<td>Lagfin SCA (LU)</td>
<td>200</td>
</tr>
<tr>
<td>Resident shareholder</td>
<td>170</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Equity investment in Davide Campari SpA</th>
<th>Amount (Euros)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Total</td>
<td>4250</td>
</tr>
<tr>
<td>Alicros SpA (IT)</td>
<td>2000</td>
</tr>
<tr>
<td>Cedar Rock (UK)</td>
<td>300</td>
</tr>
<tr>
<td>Resident floating shareholders</td>
<td>450</td>
</tr>
<tr>
<td>Non resident floating shareholders</td>
<td>1500</td>
</tr>
</tbody>
</table>

Figure 4 summarizes the hypothetical values for holdings of debt securities issued by Campari before the re-domiciliation (holdings by non-resident investors, assumed to be 250 Euros).

### Campari debt securities holdings

<table>
<thead>
<tr>
<th>debt securities issued by Davide Campari-Milano SpA</th>
<th>Amount (Euros)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Total</td>
<td>450</td>
</tr>
<tr>
<td>Resident investors</td>
<td>200</td>
</tr>
<tr>
<td>Non resident investors</td>
<td>250</td>
</tr>
</tbody>
</table>

From the Italian perspective, in the first phase, the acquisition of the remaining part of Alicros Spa from the other Italian shareholder implies an increase in the FDI liabilities: Lagfin raises the value of the unlisted equity investment in Alicros of 170 Euros (according to our assumptions). Furthermore, the subsequent merger of the two companies implies the disappearance of the direct investment in unlisted equity (Lagfin in Alicros) and the appearance of the direct investment in listed equity (Lagfin in Campari), reconciled through the “other changes in volume” item. The treatment of all the aspects related to merger (as other changes in volume or as transactions) is still under discussion in the international fora.

The following figure summarizes the recording in the Italian external statistics before, during and after the first phase.

### Impact of Phase 1 on BOP/IIP

<table>
<thead>
<tr>
<th>Impact of the Merger - PHASE 1 on external statistics</th>
</tr>
</thead>
<tbody>
<tr>
<td>Financial account</td>
</tr>
<tr>
<td>Assets</td>
</tr>
<tr>
<td>Financial account</td>
</tr>
<tr>
<td>Direct investment - Equity</td>
</tr>
<tr>
<td>in listed companies</td>
</tr>
<tr>
<td>in unlisted companies and other equity</td>
</tr>
<tr>
<td>Portfolio investment - equity</td>
</tr>
<tr>
<td>- debt securities</td>
</tr>
<tr>
<td>Other investment</td>
</tr>
<tr>
<td>Net IIP/ Net e&amp;o</td>
</tr>
</tbody>
</table>

The settlement of this operation is recorded in the Other investment item.
During the second and the third phase, no transaction actually occurs: the events related to the re-domiciliation of Campari and the subsequent establishment of a branch in Italy do not give rise to BOP recordings; yet they only affected the IIP.\textsuperscript{8}

In the second phase, the re-domiciliation of Campari from Italy to the Netherlands produces, as an immediate effect, the elimination of the stock in the direct investment listed equity liabilities (2300 Euros in our numerical example) and in the direct investment unlisted equity assets (3600 Euros). Furthermore, it produces effects also on the portfolio investment (equity and debt securities). The phase 2 implies the elimination of the stock of portfolio liabilities (connected to the holding by non-resident investors) and the appearance of portfolio assets (holdings by residents) (see first and second part of Figure 6). In the third phase, FDI liabilities in other equity (500 Euros) emerge because of the creation of the Italian branch (as well as FDI assets).

All these changes in the IIP are reconciled through an entry in the "other changes in volume” account.

### Impact of Phases 2 and 3 on IIP

<table>
<thead>
<tr>
<th>Financial account</th>
<th>Opening position (end of Phase 1)</th>
<th>Intermediate position (end of Phase 2)</th>
<th>Closing position (end of Phase 3)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Financial account</td>
<td>Assets</td>
<td>Liabilities</td>
<td>Assets</td>
</tr>
<tr>
<td>Direct investment - Equity</td>
<td>4050</td>
<td>650</td>
<td>0</td>
</tr>
<tr>
<td>in listed companies</td>
<td>2300</td>
<td>0</td>
<td>3600</td>
</tr>
<tr>
<td>in unlisted companies and other equity</td>
<td>3600</td>
<td>0</td>
<td>3600</td>
</tr>
<tr>
<td>Portfolio investment - equity</td>
<td>1500</td>
<td>450</td>
<td>0</td>
</tr>
<tr>
<td>- debt securities</td>
<td>250</td>
<td>200</td>
<td>0</td>
</tr>
<tr>
<td>Net IIP</td>
<td>-4090</td>
<td>650</td>
<td>3750</td>
</tr>
</tbody>
</table>

#### 2.2 Another type of re-domiciliation: the case of Mediaset Spa

In September 2021, Mediaset Spa, the holding company of an Italian group involved in media and communications, transferred its registered office to the Netherlands.

Before the transfer, Mediaset Spa was an Italian company listed on Milan stock exchange, with two direct investors (Fininvest Spa, an Italian enterprise, and Vivendi SA, a company resident in France) and other floating shareholders (Figure 7).

\textsuperscript{8} See BPM6 § 4.167 and § 9.21.
When moving its domicile to the Netherlands, Mediaset Spa changed its name first to Mediaset NV\(^9\) and then to Mediaforeurope NV; the re-domiciled company remained listed on Milan stock exchange.

Simultaneously with the re-domiciliation, a branch was created in Italy and provided with all the Mediaset's existing assets and liabilities even though it was considered as an indistinct part of the foreign parent company and therefore resident in the Netherlands.

Figure 8 summarizes the ownership structure after the re-domiciliation.

---

The main difference with respect to the Campari re-domiciliation (Figure 2) is the residence of the branch located in Italy: Italian in the Campari case and Dutch in the Mediaset case.

Impacts on IIP are similar to what illustrated in the Campari case and are not here reported and analysed. All changes have been treated as other changes in volume.

2.3 Main challenges

The two cases are typical examples of the complexity of financial restructuring operations. Even if enterprises involved in such processes supplied detailed information on the process on their website well in advance, there were still several challenges for BOP/IIP compilers.

As already deeply described in Lisbon paper, the main challenges regard:

- the availability of detailed information on the operation, structured for external statistics purposes;
- the identification and the correct treatment of the effects on BOP/IIP items other than FDI equity;
- the coherent interpretation of cases among compilers of the countries involved in the restructuring process in order to minimise the asymmetries and BOP errors and omissions.

The re-domiciliation in the Netherlands of Mediaset Spa posed further challenges with reference to the availability of a detailed description of the operation.

In fact, in this case, the document describing the preliminary re-domiciliation plan of the company, distributed well before the date of effectiveness of the operation, foresaw a tripartite common cross-border merger among Mediaset Spa and its subsidiaries in Spain and the Netherlands (Mediaset Espana Communication SA and Mediaset Investment NV). The recording of the effects of the operation in Italian external statistics was consequently analysed and defined on the basis of this plan but, in the end, due to legal disputes, the merger did not take place as described in the document and, at last, Mediaset Spa simply transferred its registered office to the Netherlands. This change had less prominence in the press and on the company’s website than the previous re-domiciliation plan, resulting in difficulties for the compiler in adapting the registration criteria in the external statistics.

However, the main challenge related to Campari and Mediaset re-domiciliation refers to the identification of the residency and the features of the branch located in Italy created at the same time as the change of residence. In the Campari case, it was clearly stated in the company plans that the branch was a separate unit from the Dutch mother company. On the contrary, in the Mediaset case, no clear indication on the status of the branch was available on the website or in the official documents. Only through bilateral contacts with the reporting agent, it was possible to understand that the branch was an indistinct part of Mediaforeurope NV and, thus, non-resident in Italy.

A detailed description of the tripartite common cross-border merger is available on MFE website: https://www.mfemediaforeurope.com/binary/documentRepository/13/1.%20Common%20merger%20plan%20SITO%20COMPLETO_531.pdf
The ascertainment of the residency and the accounting autonomy of the branch is a relevant element for the coherent statistical reporting in different domains (BOP/IIP and National Accounts at least) and across countries. In case of a distinct entity, in fact, the operation of the branch should be excluded from the accounts of the economy of its head office and included in the statistics of the economy in which it is located. Otherwise, all assets and liabilities should be allocated to the mother company.\textsuperscript{11} Unfortunately, it is not always possible to have a direct contact with the enterprise involved in this kind of operation in order to solve this dilemma. Furthermore, even in this most optimistic case, indications supplied by the enterprise are not always implemented in the same way in all statistical domains. In the case of Mediaset, in fact, further insights pointed out inconsistent classification across different registers (business registers, BOP register, RIAD, ...), which may result in inconsistent statistical treatment. Consequently, the following issues remain unsolved:

- the difficulty of unambiguously establishing whether a branch is resident or not (e.g. who should verify if the requested conditions are satisfied? how should be assured that the assigned residency is uniformly acquired by all registers?)
- the heterogeneity of residency attribution across countries, statistical domains and reporters.

To our knowledge, in the last year, at least other two large Italian companies moved to the Netherlands, following the same pattern of Mediaset and implying the same difficulties and possible asymmetries across domains due to this kind of misinterpretation.

3. Transfer of participations within a group: the Enel case

3.1 General description and impacts on external statistics

During the last few years, a wide reorganization involved some companies of the Enel group, one of the Italian largest private energy operator in the world. The process aims for the simplification and optimization of the group’s corporate structure and is still under way in some countries in which the group operates.

In 2020, the reorganization took place in Italy and was carried out in two stages.\textsuperscript{12} A first one concerned the partial demerger of Enel Green Power Spa (EGP) with the transfer to Enel Spa (i) of the 100% equity held by EGP in North America (in the companies Enel North America Inc. and Enel Green Power Development North America LLC) and (ii) of an intercompany loan agreement with Enel Finance International NV. With the completion of this phase, EGP Italian participations were transferred to a newly established company denominated Enel Green Power Italia Srl (Figure 9).

\textsuperscript{11} See BPM6 § 4.26 to § 4.28.
The second phase concerned the transfer of the Italian business unit from Enel Spa to an Italian sub-holding. More specifically, Enel transferred to Enel Italia Spa (the company chosen for the role of sub-holding) the equity participation in: Enel Energia Spa, Servizio Elettrico Nazionale Spa, e-distribuzione Spa, Enel Produzione Spa and Enel Green Power Italia Srl (the last one being the new company established after the spin-off occurred at the first stage). The business transferred to Enel Italia also included: (i) a portion of the debt of Enel in respect of Enel Finance International NV, (ii) the contracts, receivables and payables of Enel with the components of the above business unit and (iii) the debt of the transferred companies in respect of Enel Finance International NV. With this last point, in particular, Enel Italia Spa took over in the loan agreements of the transferred Italian companies with Enel Finance International NV (Figure 10).
From the BOP perspective, this reorganization did have very limited impacts. Indeed, all the operations refer to transfers between Italian companies of assets and/or liabilities, both cross-border and domestic. They may only affect the IIP in case of significant differences (e.g. listed/not listed, non-financial/financial) between the Italian companies exchanging cross-border positions. If any, those impacts would be treated as other changes in volume.

Effects on the BOP/IIP would have been much more considerable if assets/liabilities transfers had also involved non-resident companies in the group, an eventuality theoretically possible.

3.2 Main challenges

The main challenges of this type of restructuring concern the BOP/IIP compilation, especially in a system that uses direct reporting on a sample basis, as the Italian one.

A first issue is related to the coverage, as it is necessary to ascertain if all the Italian companies involved in the reorganization are directly included in the sample or are included only indirectly, by grossing-up techniques. In the second case, particular caution is needed in order to avoid the arising of inconsistencies and/or errors and omissions.

Furthermore, relevant efforts are necessary to verify that data reported by all different parties exchanging assets/liabilities are consistent, both in terms of substance (same financial instruments, same amounts, same counterparts have to be indicated by all the respondents) and in terms of timing (the moment of the exchange of assets/liabilities has to be the same for all parties).

In cases like the ENEL restructuring, another relevant challenge for the BOP compilers is related to the lack of information. Usually, a clear and exhaustive picture of complex reorganizations is available only when they are completed, long after the single operations (spin-offs, transfers of assets and liabilities, etc.) took place.

Until then, partial and preliminary data have to be collected, with a lot of effort, picking up information in the media, in the commercial data providers and from the direct reporters. This is particularly difficult for assets and liabilities other than equity: events such loans and credits transfers, like those occurred in the Enel restructuring, are usually not covered by the media; furthermore, the value of the assets/liabilities exchanged can be very difficult to estimate, as it is generally not available in specialized database nor in the companies’ balance sheet. Thus, compilation issues related to debt instrument are even more challenging.

4. Conclusions

The paper aims to highlight how the work of the BOP/IIP compilers is increasingly challenging due to the more and more complex internationalization and corporate restructuring strategies adopted by multinational companies.

The real-world cases analyzed in this paper offer a fairly varied overview of these challenges that can be summarized as follows:
− **Availability and timeliness of information.** The more timely and accurate the details provided by the media or in the companies’ balance sheets, the fewer checks to be made through additional sources/information. In any case, the available information is rarely structured for BOP/IIP purposes and concerns only a detailed description of the transaction for shareholders and investors, with few numerical details limited to some balance sheet items without the breakdowns necessary to make them usable for BOP/IIP purposes.

− **Consistent interpretation of facts and identification and allocation of the effects on all BOP/IIP items** among all the external statistics compilers from countries involved in the cross border operations, in order to avoid potential asymmetries and reduce BOP errors and omissions. Not all compilers of the involved countries have the same available information and, if there are no unique and internationally agreed guidelines for the registration of all possible cases (i.e. the treatment of mergers as other changes in volume or transaction), each country can interpret and register them in a different way.

− **Necessity to deal with peculiar cases**, such as the establishment of branches (the ascertainment of their residency and their accounting autonomy) or the creation of temporary ad-hoc units whose short life prevent from the possibility to collect the necessary information from them.

− **Lack of flexibility of data collection system**, not always structured to immediately include new entities as reporting agents and to collect all necessary information on time (sometimes figures outside the scope of BOP are essential to complete the framework, for example position vs. residents).

The work of external statistics compilers can be improved, first, through the outlining of univocal definitions, classifications and registration practices to be implemented in the current revision process of international manuals. This will surely reduce possible asymmetries between countries and the weight of the errors and omissions item. The new manuals should contain detailed guidelines enriched with concrete examples in order to help in assuring homogeneous recording among countries and statistical domains.

In this context, the exchange of views and current practices among compilers should be supported in all international fora dealing with this subject.

Lastly, the use of flexible data collection systems, in terms of structure of the questions and reporting agents (see Box1 for details), is essential to collect information directly from companies as much as possible.

---

13 BPM6 as stated in the Introduction, but also the OECD Benchmark Definition of Foreign Direct Investment and the System of National Accounts.
References


International Monetary Fund, Direct Investment Task Team (DITT): *D4 Corporate Inversions*.


Mediaset Spa (2021): *Completion of the transfer of the registered office to the Netherlands*.

Challenges in the external statistics framework: how to register MNE financial restructuring operations

Nadia Accoto, Giuseppina Marocchi and Silvia Sabatini
Bank of Italy - Statistical Analysis Directorate

11th IFC Biennial Conference “Post-pandemic landscape for central bank statistics”
(Basel, 25-26 August 2022)
Introduction

In recent years, the number, the complexity and the variety of possible corporate financial restructuring operations have been increasing.

The process of revision of international manuals is working on a precise definition of these operations and their impacts on external statistics (harmonization of compilation practices, reduction of possible bilateral asymmetries).

Two kind of operations, and related challenges, are here presented:

• **Re-domiciliation** of the holding company of a group

• **Transfer of participation** between group entities

...in continuity with what presented in Lisbon in February 2020 during the IFC Conference on external statistics
During the period 2019-2020 a corporate restructuring involved Davide Campari, one of the world most important food and beverage groups located in Italy. The process developed in several steps and ended up in the re-domiciliation in the Netherlands of the holding company of the group, Davide Campari-Milano S.p.A.

In September 2021, Mediaset Spa, the holding company of an Italian group involved in media and communications, transferred its registered office to the Netherlands.

In both cases, a permanent establishment in Italy (a branch) was created to carry out, without interruption, the activities performed by the “old” enterprise.
The Davide Campari group after the restructuring

The Mediaset (Mediaforeurope) group after the restructuring

Green (blue) boxes refer to Italian (non resident) enterprises
The main difference in the two cases is the **residence of the branch located in Italy**: Italian in the Campari case (as stated on the official documents) and Dutch in the Mediaset case (ascertained only through bilateral contacts with the reporting agent).

In the Mediaset case, further insights put in evidence inconsistent classification across different registers (business registers, BOP register, RIAD, ...); it could imply **incoherent statistical reporting in different domains** (BOP/IIP and national account at least).

Unresolved issues:

- the difficulty of establishing whether a branch is resident or not (e.g. who should verify if the requested conditions are satisfied? how should be assured that the assigned residency is uniformly acquired by all registries?)
- the heterogeneity of residency attribution across countries, statistical domains and reporters.
During the last few years, a wide reorganization involved some companies of the Enel group, one of the Italian largest private energy operator in the world. The process aims for the simplification and optimization of the group’s corporate structure and is still under way in some countries in which the group operates.

In Italy, the reorganization took place in 2020 and was carried out in two stages:

1) Partial demerger of Enel Green Power SpA with the transfer to Enel SpA of North-American participations and intercompany loan agreement with Enel Finance International NV.

2) the transfer of the Italian business unit from Enel SpA to the Italian sub-holding Enel Italia SpA which also took over in the loan agreements of the transferred Italian companies with Enel Finance International NV.
Stage 1)

Re-domiciliation

Transfer of participations

Conclusions

crossborder loans

ENEL SpA

foreign companies

Enel Italia

Enel Energia

distribuzione

Enel Produzione

Enel North America Inc

EGP Dev. North Am. LLC

Enel Green Power Italia Srl

Enel Finance International NV

Italian companies

Conclusions
Stage 2)

ENEL SpA

- foreign companies
  - Enel North America Inc
  - EGP Dev. North Am. LLC

Enel Italia

- Enel Energia
- e-distribuzione
- Enel Produzione
- Enel Green Power Italia Srl

Crossborder loans

Enel Finance International NV

Re-domiciliation

Transfer of participations

Conclusions

Italian companies
In order to correctly represent in BOP/IIP the effects of the reorganization, all the Italian companies involved need to be directly included in the sample; if some of them is indirectly included, by grossing-up techniques, then particular caution is needed in order to avoid inconsistencies and/or errors and omissions.

Relevant efforts are necessary to verify that data reported by all different parties exchanging assets/liabilities are consistent (in terms of amounts, financial instruments...)

Compilation issues related to debt instrument are even more challenging: loans and credits transfers are usually not covered by the media; furthermore, the value of the assets/liabilities exchanged can be very difficult to estimate, as it is generally not available in specialized database nor in the companies’ balance sheet.
### Challenges

Peculiar cases, such as **branches** or temporary ad-hoc units and related problems

Availability and timeliness of **information**

Consistent **interpretation** of facts among compilers

Identification and allocation of the effects on BoP/IIP items other than FDI equity

Lack of flexibility of **data collection system**

### Hints

Widespread sharing of the main definitions, classifications and **methodologies**

**Detailed guidelines** enriched with concrete examples

**Exchange views** among compilers

**Flexible data collection systems**
Thank you for your attention!

Nadia Accoto
nadia.accoto@bancaditalia.it
Does firm size predict the residency status of the final investor? Evidence from Romanian FDI enterprises

Catalina-Florentina Pricope, National Bank of Romania

---

1 This presentation was prepared for the conference. The views expressed are those of the author and do not necessarily reflect the views of the BIS, the IFC or the central banks and other institutions represented at the event.
Does firm size predict the residency status of the final investor? Evidence from Romanian FDI enterprises

Author: Cătălina-Florentina Pricope

Abstract
The present study examines the extent to which the residency status of FDI enterprises is related to different firm-size measurements. Specifically, the analysis aims at revealing whether firm-size characteristics can predict an overlap (or lack thereof) between the Ultimate Investing Country (UIC) and the Immediate Investor Country (IIC). The research hypotheses are tested on a sample of 7,311 Romanian FDI entities, using data collected via the National Bank of Romania (NBR) and the National Institute of Statistics (NIS) 2020 FDI survey. Results indicate that enterprises with higher turnover, profits and FDI positions are more likely to be finally controlled by investors whose residence does not coincide with that of any immediate investor, while, surprisingly, FDI firms with more employees are ultimately controlled by entities resident in the IIC.

Key words: ultimate investing country, immediate investor country, final investor, foreign direct investment, firm level analysis, ultimate controlling parent

JEL classification: F21, F23
Introduction

During the last two decades, the acceleration of business and capital markets globalisation has increased demand for high-quality foreign direct investment (FDI) statistics. Given their various analytical applications, data users and compilers are interested both in aggregated FDI figures included in the Balance of Payments (BoP) and in the International Investment Position (IIP) as well as statistics broken down by partner country and industry (OECD, 2020). Demand for reliable FDI statistics was further enhanced by the SARS-CoV-2 pandemic, as the crisis revealed the need to have more timely, frequent and well-documented indicators to guide policy (Tissot & De Beer, 2020).

Financing arrangements undertaken by multinational enterprises (MNEs) have become more complex over time due to a variety of causes, including the requirement to manage worldwide production networks and the desire to decrease tax and regulatory costs (OECD, 2015a). Layers of equity ownership linkages connect a parent corporation to its subsidiaries, determining whether it has direct or indirect control over them. The growing complexity of corporate structures raises concerns about the efficacy of national and international investment regulations based on investor residency. In addition to their ultimate investor, affiliates can have one or more direct shareholders and multiple indirect shareholders, all of whom may be residents of various countries (Alabrese & Casella, 2020).

MNEs are important players in the globalization process, as their activities involve production, trade, direct investment and technology transfer aimed at maximizing global profitability. Transnational corporations strategically distribute worldwide production among their cross-border affiliates, in order to take advantage of reduced labour costs, enhanced market access, more favourable regulatory frameworks, tax benefits and higher skilled work force. MNEs are increasingly setting up special purpose entities (SPEs) to channel their financial assets abroad and gain access to a wider range of financial markets.

For international business conglomerates, the choice of entrance method is critical within the internationalization decision-making process, since it defines the degree of control over an enterprise’s activities in foreign markets. A suitable market entry mode is able to assist MNEs in gaining competitive advantages and even decide the investment’s efficiency and development.

In this context, economic researchers have been constantly concerned with the study of the investment channels chosen by MNEs, as well as with the analysis and understanding of the factors that determine their strategic choices.

The purpose of the present study is to investigate the extent to which firm-size characteristics differ between entities for which the UIIC coincides with at least one IIC and entities for which the ultimate investor is not a resident of any IIC. Although there is a plethora of literature on FDI, most studies focus mainly on the direction and driving forces of FDI flows as well as their impact on home and host economies (Ngoasong, Wang, Amdam & Bjarnar, 2021; Ahmed, Jones & Temouri, 2020; Li, Liao & Sun, 2018). The contribution that this study makes to prior literature is the use of a comparative analysis of firm-level characteristics based on the ultimate investors’ residence.

The paper is structured as follows: the first section presents a review of the literature, indicating relevant studies that focus on FDI investment chains. The literature review is followed by an empirical study in which the dependencies between foreign investors’ residency status and entity-size are identified and analysed. Following the discussion of the results, new research topics are proposed in pursuit of an even more comprehensive analysis of cross-border capital transactions.

Literature review

FDI has long been recognized as a significant financial resource. This type of investment enables capital, know-how and technological transfers, resulting in enhanced economic growth, higher productivity and stronger trade ties. From an analytical standpoint, it is important to determine the direct investment’s source (European Comission, 2019).
Traditional FDI datasets are compiled based on the country of residence of the immediate investor (IIC principle). This, in turn, enables MNEs complicated structures to disguise the ultimate source of FDI into a country as the UIC and the associated controlling parent, in cases where investors use chains of investment entities in economies other than their own, cannot be identified (European Commission, 2019).

In order to provide compilers and users of FDI data with more meaningful metrics, the OECD produced the 4th edition of its Benchmark Definition of FDI (BD4). According to BD4, countries should compile and disclose information on inward FDI positions by UIC (OECD, 2015b). This representation enables statisticians to identify the residence of the entity which ultimately controls the investment. Additionally, it facilitates the production of more nuanced statistics that provide deeper insights into economic relationships and can improve the traceability of funds (European Commission, 2019).

Data on UIC can be sourced from direct reporting or from already existing sources. While direct reporting through surveys enables data collection in line with international statistical standards, it also represents an additional response burden for reporters. Thus, compilers aim to manage this burden whenever feasible, through a combination of survey with other available data sources such as business registers or administrative sources. According to the European Statistical System’s (ESS) and European System of Central Banks’ (ESCB) joint Task Force on FDI (TF-FDI) (2020) certain countries have already implemented enterprise group registers that can be used to identify the UIC, while others use private sources on MNEs. Annual reports are another major source of information on foreign control or ownership. However, these often lag behind the reference year.

Significant differences can exist between the two representations (i.e. UIC and IIC) of the FDI position (OECD, 2015b). Identifying the underlying economic meaning of these differences as well as their economic, social and political impact on both investing and recipient countries has become a growing area of interest for governments, international organisations and academia.

There is a consensus in the literature that MNEs employ complicated ownership structures in order to manage their global operations, finances and intellectual property, as well as to decrease their tax and regulatory responsibilities (Ngoasong et al., 2021; Ahmed et al., 2020; Bankman, Kane & Sykes, 2018). The UIC compilation of inward FDI statistics enables users to look through these complicated systems to the ultimate source of investment in the country of interest. Additionally, FDI UIC figures supply analysts and policymakers with important information about who ultimately owns, reaps the benefits and bears the risks of certain investments (OECD, 2015b).

One of the early attempts to describe ownership patterns of significant enterprises across countries was made by La Porta, Lopez-De-Silanes and Shleifer (1999). The study examines the ownership chain of a sample of significant firms in developed economies to determine their level of concentration, as well as who controls them and how. The presence of pyramidal control structures and rare examples of cross-shareholding are documented by the authors.

A later line of academic inquiry looks into the specific factors that influence MNEs’ financial and investment decisions, which may have an impact on the structure of ownership chains. Several studies investigate MNEs’ choices as a result of tax consideration.

Altshuler and Grubert (2003) examine how transnational corporations employ affiliates to carry out investment repatriation schemes. The authors acknowledge that despite broad interest in globalization, research on MNEs tends to focus on a narrow spectrum of financial flows between overseas affiliates and their parent companies. MNEs are able to choose between direct dividend repatriation to the parent and continued real investment in the foreign affiliate. Dividend payments are a costly tax option since these transfers are taxed at the higher home country rate when received by the parent. Real investment in the overseas affiliate, which may yield lower returns than domestic investment, is one of many alternatives to direct dividend repatriation. MNEs can employ various tactics to achieve the equivalent of repatriation without having to pay the home country tax on direct low-tax income repatriation.
Ahmed et al. (2020) examine the complementary relationship between tax haven use and FDI in the developing world. The authors conduct their inquiry on a firm-level dataset and show a robust positive relationship between tax haven use and FDI into countries with low economic development and high levels of capital flight.

Mintz and Weichenrieder (2010) findings suggest that the likelihood of group consolidation and the capital exporting country’s credit system influence the presence of convoluted ownership arrangements. Concerns about political and expropriation risks also motivate investors to seek investment protection through international agreements, although financial exposure, financing tactics and the host country’s institutional framework may also be factors (Lewellen & Robinson, 2013).

These studies have attempted to provide insights into current FDI research questions, however, the relationship between FDI enterprise size and the overlap between UIC and IIC may generate interesting answers for future developments, which can be of significant importance for analysts, policymakers and academia. It is reasonable to expect that MNEs with more complex chains of ownership are able to take better advantage of all the benefits associated to the channelling of funds through one or even more intermediary countries before allocating them to the ultimate host country. Thus, final investors of such MNEs might be able to ultimately control larger entities than those who invest directly in the host economy.

Over 3,000 bilateral international tax treaties have been signed by countries throughout the world. As a result, the tax environment is continually shifting and MNEs can make use of transfer pricing tactics to redirect earnings from high-tax jurisdictions to low-tax ones (Kleist, 2018; Eden & Kudrle, 2005). These tactics, which are facilitated by gaps and mismatches in fiscal legislation, enable domestic tax base erosion and profit shifting (BEPS). Tax avoidance practices cost governments an estimate of about USD 100 to 240 billion in income, yearly. Because developing countries rely to a greater extent on corporate income tax than developed ones, they are disproportionately affected by BEPS (OECD, 2022).

The OECD and G20 countries have taken action towards rectifying the flaws in the international tax system that allow BEPS to thrive. They have devised a comprehensive package of measures to combat BEPS (the BEPS package) (OECD, 2017). Members of the framework (141 nations and jurisdictions) collaborate in order to combat tax evasion, increase international tax coherence and create a more transparent tax environment (OECD, 2022).

MNEs are able to channel funds and employ BEPS strategies through setting up SPEs. Given their growing importance, the International Monetary Fund’s (IMF) Committee on Balance of Payments Statistics endorsed the creation of a Task Force on SPEs (TFSPE), in charge with formulating the definition of SPEs, which states that:

- “An SPE, resident in an economy, is a formally registered and/or incorporated legal entity recognized as an institutional unit, with no or little employment up to maximum of five employees, no or little physical presence and no or little physical production in the host economy;
- SPEs are directly or indirectly controlled by non-residents;
- SPEs are established to obtain specific advantages provided by the host jurisdiction with an objective to (i) grant its owner(s) access to capital markets or sophisticated financial services; and/or (ii) isolate owner(s) from financial risks; and/or (iii) reduce regulatory and tax burden; and/or (iv) safeguard confidentiality of their transactions and owner(s);
- SPEs transact almost entirely with non-residents and a large part of their financial balance sheet typically consists of cross-border claims and liabilities.” (IMF, 2018).

In light of previous research and international political and economic initiatives targeting MNEs, the present study aims to investigate whether there are significant differences in terms of size between enterprises that are ultimately controlled by an investor resident in an IIC compared to those that have
no immediate investment relationship with the UIC. Four explorative hypotheses on the relationship between firm size and final investors’ residency were developed, namely:

**H1:** FDI enterprises with higher turnover levels are more likely ultimately controlled by entities which are not residents in the IIC.

**H2:** FDI enterprises with higher profit levels are more likely ultimately controlled by entities which are not residents in the IIC.

**H3:** FDI enterprises with higher employment levels are more likely ultimately controlled by entities which are not residents in the IIC.

**H4:** FDI enterprises with higher FDI positions are more likely ultimately controlled by entities which are not residents in the IIC.

Data and research method

The study was conducted on a sub-sample of 7,311 FDI enterprises that were subject to the National Bank of Romania (NBR) and the National Institute of Statistics (NIS) 2020 FDI survey. Out of the entire survey sample, only legal entities that submitted a complete set of survey forms or that allowed data imputation from administrative sources were kept in the analysis.

Romania’s 2020 FDI statistics were computed in compliance with the methodology set forth in the IMF Balance of Payments and International Investment Position Manual, Sixth Edition (BPM6), based on data collected mainly through a direct statistical survey.

NBR employs the *winner takes it all* (WTA) approach to allocate FDI positions to the ultimate investor instead of the proportional approach (PA), which accounts for direct minority investment. The fundamental difference between the two approaches is that while the WTA allocates all inward FDI to where the MNEs decisions are made, PA allocation emphasises the true ultimate country portfolio of inward FDI, including the allocation of risk exposure. However, due to the scarcity of data and high additional costs for both compilers and reporters, members of the IMF Committee on Balance of Payments Statistics expressed their preference towards the use of the WTA approach for UIC FDI allocation (TF-FDI, 2020).

At end-2020, Romania’s total FDI stock stood at EUR 90,773 million. Equity positions amounted to EUR 63,952 million, while debt positions totalled EUR 26,821 million (NBR, 2020). Romania’s top five investing countries (IIC and UIC) and their corresponding FDI positions are presented in Table 1.

<table>
<thead>
<tr>
<th>IIC</th>
<th>FDI position (EUR mil.)</th>
<th>% total FDI position</th>
<th>UIC</th>
<th>FDI position (EUR mil.)</th>
<th>% total FDI position</th>
</tr>
</thead>
<tbody>
<tr>
<td>The Netherlands</td>
<td>19,994</td>
<td>22.0</td>
<td>Germany</td>
<td>13,792</td>
<td>15.2</td>
</tr>
<tr>
<td>Germany</td>
<td>11,070</td>
<td>12.2</td>
<td>Austria</td>
<td>10,153</td>
<td>11.2</td>
</tr>
<tr>
<td>Austria</td>
<td>10,858</td>
<td>12.0</td>
<td>France¹</td>
<td>8,733</td>
<td>9.6</td>
</tr>
<tr>
<td>Italy</td>
<td>7,652</td>
<td>8.4</td>
<td>Italy</td>
<td>7,556</td>
<td>8.3</td>
</tr>
<tr>
<td>France¹</td>
<td>5,642</td>
<td>6.2</td>
<td>United States</td>
<td>6,167</td>
<td>6.8</td>
</tr>
</tbody>
</table>

Source: NBR, 2020

In 2020, Romania’s FDI net flows dropped by 41.9 percent from the previous year, to EUR 3,005 million. This decline was mainly caused by the outbreak of the SARS-CoV-2 pandemic and the subsequent declared state of emergency in the country, in March 2020. During this atypical year, a large

¹ including the Principality of Monaco
number of FDI enterprises incurred losses when faced with supply chain disruptions and temporary business suspensions (NBR, 2020).

Firm-level (granular) data were imported into Stata 15.1 for statistical and econometric interpretation.

A multivariate approach was employed for empirical testing of the research hypotheses. The study variables are presented in Table 2.

<table>
<thead>
<tr>
<th>Variables</th>
<th>Measures/concepts</th>
</tr>
</thead>
<tbody>
<tr>
<td>Residency status (RSTAT)</td>
<td>dummy dependent variable; takes value 0 if, in 2020, the firms’ ultimate investor was resident in (one of) the IIC and 1 otherwise</td>
</tr>
<tr>
<td>Turnover (TURN)</td>
<td>turnover (EUR thousand) in 2020</td>
</tr>
<tr>
<td>Profit/loss (PL)</td>
<td>profit or loss (EUR thousand) in 2020</td>
</tr>
<tr>
<td>Number of employees (EMPL)</td>
<td>average number of people employed on a full time basis in 2020</td>
</tr>
<tr>
<td>FDI position (FDI)</td>
<td>FDI position (EUR thousand) at end-2020</td>
</tr>
</tbody>
</table>

The research hypotheses were tested based the following equation:

\[
\text{Logit (RSTAT)} = \alpha_0 + \alpha_1 \text{TURN} + \alpha_2 \text{PL} + \alpha_3 \text{EMPL} + \alpha_4 \text{FDI} + \varepsilon
\]

\(\alpha_i\) are the regression's coefficients and \(\varepsilon\) represents the residuals.

Several tests were carried out using the functions implemented in Stata 15.1:

- t-test and Mann-Whitney test to determine the significance of the difference between turnover, profit/loss, number of employees and FDI positions of entities whose UIC coincides with at least one IIC compared to those for which there is no overlap between the UIC and the IIC;
- Pearson-R and Spearman-R to determine the correlation between the dependent and the independent variables.

Data analysis and results

The main descriptive statistics of the dataset are present in Table 3. In order to ensure the confidentiality of the microdata used, minimum and maximum values corresponding to each variable were not included in the analysis.

<table>
<thead>
<tr>
<th>Variable</th>
<th>Mean</th>
<th>Median</th>
<th>Standard deviation</th>
<th>Skewness</th>
<th>Kurtosis</th>
</tr>
</thead>
<tbody>
<tr>
<td>RSTAT</td>
<td>0.22</td>
<td>0</td>
<td>0.41</td>
<td>1.37</td>
<td>2.87</td>
</tr>
<tr>
<td>TURN</td>
<td>23,816</td>
<td>2,957</td>
<td>127,810</td>
<td>15.47</td>
<td>315.24</td>
</tr>
<tr>
<td>PL</td>
<td>936</td>
<td>65</td>
<td>9,564</td>
<td>16.28</td>
<td>436.46</td>
</tr>
<tr>
<td>EMPL</td>
<td>165</td>
<td>37</td>
<td>614</td>
<td>13.33</td>
<td>257.61</td>
</tr>
<tr>
<td>FDI</td>
<td>12,294</td>
<td>1,407</td>
<td>62,298</td>
<td>18.45</td>
<td>497.16</td>
</tr>
</tbody>
</table>

A preliminary analysis of the data was carried out in order to find potential errors and outliers. The mean value of the RSTAT variable suggests that for approximately 22% of the sampled entities (i.e.
1,592 observations) the immediate investor is not resident in the same country as the ultimate investor. The descriptive statistics presented in Table 3, particularly skewness and kurtosis, suggest that the independent variables are not normally distributed and record extreme values. In order to minimise the potential adverse effects on results’ analysis and interpretation, data were winsorized at the 10% and the 90% percentile. Table 4 presents the descriptive statistics of the independent variables post-winsorization.

Table 4. Descriptive statistics after data winsorization

<table>
<thead>
<tr>
<th>Variable</th>
<th>Mean</th>
<th>Median</th>
<th>Standard deviation</th>
<th>Minimum</th>
<th>Maximum</th>
<th>Skewness</th>
<th>Kurtosis</th>
</tr>
</thead>
<tbody>
<tr>
<td>TURN</td>
<td>8,486</td>
<td>2,957</td>
<td>11,406</td>
<td>116</td>
<td>35,509</td>
<td>1.48</td>
<td>3.80</td>
</tr>
<tr>
<td>PL</td>
<td>331</td>
<td>65</td>
<td>723</td>
<td>-552</td>
<td>1,953</td>
<td>1.14</td>
<td>3.34</td>
</tr>
<tr>
<td>EMPL</td>
<td>82</td>
<td>37</td>
<td>99</td>
<td>1</td>
<td>315</td>
<td>1.39</td>
<td>3.60</td>
</tr>
<tr>
<td>FDI</td>
<td>4,986</td>
<td>1,408</td>
<td>7,065</td>
<td>0</td>
<td>21,744</td>
<td>1.49</td>
<td>3.82</td>
</tr>
</tbody>
</table>

Although standard deviation values remain high, the skewness and kurtosis of the four variables point to a slightly more normal distribution of the data. This is also shown by the comparative analysis of the normal probability plots in Figures 1-8.
Comparative descriptive statistics (Table 5) were computed in order to determine if any preliminary differences exist between the two groups of observations, classified based on the dummy dependent variable of the study.

<table>
<thead>
<tr>
<th>Variable</th>
<th>RSTAT = 1; UIC &lt;&gt; IIC</th>
<th>RSTAT = 0; UIC = IIC</th>
<th>t-test</th>
<th>p-value</th>
<th>Mann Whitney test</th>
<th>p-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>TURN</td>
<td>Mean = 10,776; Median = 4,824; Standard deviation = 12,519</td>
<td>Mean = 7,849; Median = 2,593; Standard deviation = 10,994</td>
<td>-9.11</td>
<td>0.00</td>
<td>-9.38</td>
<td>0.00</td>
</tr>
<tr>
<td>PL</td>
<td>Mean = 459; Median = 157; Standard deviation = 826</td>
<td>Mean = 295; Median = 51; Standard deviation = 688</td>
<td>-8.03</td>
<td>0.00</td>
<td>-6.02</td>
<td>0.00</td>
</tr>
<tr>
<td>EMPL</td>
<td>Mean = 92; Median = 42; Standard deviation = 109</td>
<td>Mean = 80; Median = 37; Standard deviation = 97</td>
<td>-4.42</td>
<td>0.00</td>
<td>-2.10</td>
<td>0.00</td>
</tr>
<tr>
<td>FDI</td>
<td>Mean = 6,631; Median = 2,395; Standard deviation = 8,079</td>
<td>Mean = 4,259; Median = 1,256; Standard deviation = 6,685</td>
<td>-10.57</td>
<td>0.00</td>
<td>-8.70</td>
<td>0.00</td>
</tr>
</tbody>
</table>

Findings indicate that FDI entities that have their final investor resident in one of the IIC recorded on average lower turnover and profit levels, while also employing less people during 2020 than FDI entities for which the UIC and IIC differ. Central trend indicators also suggest that entities controlled by an IIC resident final investor exhibit smaller FDI positions at end-2020 than their counterparts. Furthermore, the t-test revealed that on average all indicators differ significantly between the two groups of observations. Results of the Mann-Whitney test are consistent with those of the t-test.

Pearson R and Spearman R coefficients (Table 6 and Table 7) suggest positive correlations between all the variables included in the analysis. However, the intensity of the correlations between the dependent variable RSTAT and the independent variables is low to moderate and it is not likely to reduce the precision of the model’s estimated coefficients.
Table 7. Spearman coefficients

<table>
<thead>
<tr>
<th>Variable</th>
<th>RSTAT</th>
<th>TURN</th>
<th>PL</th>
<th>EMPL</th>
<th>FDI</th>
</tr>
</thead>
<tbody>
<tr>
<td>RSTAT</td>
<td>1.0000</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>TURN</td>
<td>0.1098</td>
<td>1.0000</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>PL</td>
<td>0.0704</td>
<td>0.4400</td>
<td>1.0000</td>
<td></td>
<td></td>
</tr>
<tr>
<td>EMPL</td>
<td>0.0245</td>
<td>0.7087</td>
<td>0.2662</td>
<td>1.0000</td>
<td></td>
</tr>
<tr>
<td>FDI</td>
<td>0.1018</td>
<td>0.4853</td>
<td>0.3214</td>
<td>0.2053</td>
<td>1.0000</td>
</tr>
</tbody>
</table>

Due to the fact that the dependent variable of the model is binary, a logistic model is applied to the dataset in order to estimate the probability that an observation belongs to one of the two categories. After running the logistic regression, the following equation was obtained:

\[
\text{Logit}(\text{RSTAT}) = -1.519226 + 0.0000122 \times \text{TURN} + 0.0001231 \times \text{PL} - 0.0007465 \times \text{EMPL} + 0.0000266 \times \text{FDI}
\]

The model revealed that all the independent variables of the study are predictive of the residency status of the final investor at a confidence level of 90%. Results of the Hosmer–Lemeshow test show that the overall model fit is 45.48%.

Table 8. Stata regression output

<table>
<thead>
<tr>
<th>Variable</th>
<th>Expected result</th>
<th>Coefficient</th>
<th>Odds ratio</th>
<th>p-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>TURN</td>
<td>+</td>
<td>0.000122</td>
<td>1.000012</td>
<td>0.002</td>
</tr>
<tr>
<td>PL</td>
<td>+</td>
<td>0.0001231</td>
<td>1.000123</td>
<td>0.005</td>
</tr>
<tr>
<td>EMPL</td>
<td>+</td>
<td>-0.0007465</td>
<td>0.9992538</td>
<td>0.053</td>
</tr>
<tr>
<td>FDI</td>
<td>+</td>
<td>0.0000266</td>
<td>1.000027</td>
<td>0.000</td>
</tr>
<tr>
<td>Const.</td>
<td></td>
<td>-1.519226</td>
<td>0.2188813</td>
<td>0.000</td>
</tr>
</tbody>
</table>

\[
\chi^2 = 130.64
\]

The multivariate analysis revealed that higher levels of turnover, profit and FDI increase the likelihood of an enterprise being controlled by an ultimate investor that is not resident in an IIC. According to the odds ratios obtained:

- a turnover increase of EUR 1,000 is associated with an increase of 0.0012% in the odds that the ultimate investor is not resident in the IIC;
- a profit increase of EUR 1,000 is associated with an increase of 0.0123% in the odds that the ultimate investor is not resident in the IIC;
- an FDI position increase of EUR 1,000 is associated with an increase of 0.0027% in the odds that the ultimate investor is not resident in the IIC.

We can be more than 95% confident that these relationships did not occur by chance and that they will be reflected in the population. There is, thus, strong empirical evidence to support hypotheses H1, H2 and H4.
Additionally, findings indicated that the more people an entity employs, the more likely it is to be ultimately controlled by an investor resident in the IIC. Specifically, an additional employee per entity decreases the odds that the UIC is different from the IIC by 0.07462%. This result is statistically significant at a 90% confidence level and contradicts hypothesis $H3$ of the present study.

The value of $\chi^2$ is 130.64 and has a null associated probability ($p$-value = 0.0000), suggesting that the research model is valid. However, caution is warranted when interpreting the results as the study is subject to certain limitations. Firstly, a number of factors that may have a significant influence on whether an FDI entity is likely to be controlled by a final investor resident in an IIC have not been included in the model due to the difficulty to operationalise them at this stage. Secondly, the hypotheses were tested on NBR’s FDI dataset for 2020. This particular year was marked by the outbreak SARS-CoV-2 pandemic that caused FDI flows to drop, determined FDI enterprises to suspend their business operations and prompted the government to implement financial aid schemes that may distort firm-level indicators. Expanding the analysis for a longer period of time might yield different results.

The main findings of the present inquiry could inform academia, as it contributes to prior literature by approaching a new and topical subject, namely FDI-UIC statistics. Scarcity of such data makes research on this topic challenging. However, further debates could shed light into previously sporadically studied economic relationships established between entities and geographical regions.

Furthermore, the business environment could benefit from knowing the characteristics of companies, which increase their probability of being controlled by a certain category of final investor. For instance, entrepreneurs that develop start-ups with the intention of selling them or expanding their operations could design underling business strategies that enable them to become attractive investment opportunities for particular investors.

**Conclusions**

Crises are frequently an opportunity to learn. The SARS-CoV-2 pandemic highlighted the need to access timely and high-quality statistical data for a wide variety of purposes. It represents thus an opportunity to consider the further development of the infrastructure that underpins official statistics, in particular by identifying and filling key data gaps and reorganizing processes (Rosolia et al., 2021). The health crisis also highlighted the need to undertake initiatives that enable production of new datasets, which could, in turn, facilitate a better understanding of how international finances are channelled.

The purpose of the present study was to investigate the relationship between FDI entities' size and the residency status of their final investors. Returning to the hypotheses posed at the beginning of the study, findings seem to indicate that fewer employees and higher levels of turnover, profit and FDI increase the likelihood of an enterprise being controlled by an ultimate investor that is not resident in an IIC. These results may suggest that the ultimate controlling parents of FDI enterprises might be able to benefit from certain strategic advantages when they decide to invest in host countries through intermediate ones, which in turn could enable them to set up or take over better performing projects. In turn, the aforementioned findings may be of significant importance for: (i) academia, by enabling researchers to better understand the dynamics of the business environment and the economic relationships that form between certain geographical regions; (ii) private companies interested in becoming attractive investment avenues for certain MNEs, by informing them which firm-characteristics are of interest to the final business-conglomerate owner. The analysis focused exclusively on Romanian resident FDI entities and used data collected via the 2020 NBR and NIS statistical survey. It thus makes a noteworthy contribution to the literature by testing the research hypotheses on firm-level microdata. Findings revealed that specific firm-size features are predictive of the residency status of their final investors. However, it is important to note that these results were obtained and are statistically significant for Romania’s particular case. To date, no entities that meet the criteria stated in the SPEs definition, operationalized by the IMF, have been identified in Romania, as the national fiscal legislation does not
encourage their establishment. Testing of the four research hypotheses on a sample of firms resident in other economies (particularly those with significant SPE presence) might yield different results.

The international flow of capital and its key drivers represent a fruitful area for further work. More research is needed to better understand the reasoning behind investors' decision to choose one particular investment strategy over another. Subsequent inquiries could assess whether final investors, which are not residents of an IIC, would rather take over existing entities or choose to start new businesses from scratch.
Bibliography


Does firm size predict the residency status of the final investor? Evidence from Romanian FDI enterprises

Cătălina-Florentina Pricope
Statistics Department

11th Biennial IFC Conference on “Post-pandemic landscape for central bank statistics”

Basel, August 25, 2022
State of knowledge

• The SARS-CoV-2 pandemic revealed the need to have more timely, frequent and accurate statistics to guide policy (Tissot & De Beer, 2020).

• Central banks are the forefront of both the production and use of economic and financial data → hold a unique viewpoint on official statistics (Rosolia, Stapel-Weber & Tissot, 2021).

• The acceleration of business and capital markets globalisation has increased demand for high-quality FDI statistics (OECD, 2020).

• As key-players in the globalization process, MNEs:
  ➢ engage in production, trade, direct investment and technology transfer aimed at maximizing global profitability;
  ➢ strategically distribute worldwide production among their cross-border affiliates (Ngoasong et al., 2021; Ahmed et al., 2020; Bankman, Kane & Sykes, 2018).
• Complex corporate structures enable MNEs to disguise their final investor. Traditional FDI datasets compiled based on the *Immediate Investor Country (IIC)* principle do not capture the ultimate source of investment.

• OECD’s Benchmark Definition of FDI 4th ed. (BD4) recommends countries to compile and disclose inward FDI positions by *Ultimate Investing Country (UIC)*, as this representation:
  - enables statisticians to identify the residence of the entity which ultimately controls the investment;
  - facilitates the production of more nuanced statistics that provide deeper insights into economic relationships;
  - can improve the traceability of funds (European Commission, 2019).

• It is reasonable to expect that MNEs with more complex chains of ownership are able to take better advantage of all the benefits associated to the channelling of funds through intermediary countries before allocating them to the ultimate host country.

• Final investors of such MNEs might be able to ultimately control larger entities than those who invest directly in the host economy.
Purpose of the study and research hypotheses

• **Purpose**: to investigate the extent to which firm-size characteristics differ between entities for which the UIC coincides with at least one IIC and entities for which the ultimate investor is not a resident of any IIC.

• **Contribution to prior literature**: use of a comparative analysis of firm-level characteristics based on the ultimate investors’ residence.

• **Research hypotheses**:

  
  ![FDI enterprises with higher turnover levels](image1)
  ![FDI enterprises with higher profit levels](image2)
  ![FDI enterprises with higher employment levels](image3)
  ![FDI enterprises with higher FDI positions](image4)

  are more likely ultimately controlled by entities which are not residents in the IIC.
Data and research method

- The study was conducted on a sub-sample of 7,311 FDI enterprises that were subject to the National Bank of Romania (NBR) and the National Institute of Statistics (NIS) 2020 FDI survey.

- Out of the entire survey sample, only legal entities that submitted a complete set of survey forms or that allowed data imputation from administrative sources were kept in the sub-sample analysed.

- NBR employs the *winner takes it all* (WTA) approach to allocate FDI positions to the ultimate investor and ensures the voluntary transmission of the data to Eurostat.

- A multivariate approach was employed for empirical testing of the research hypotheses:

  \[
  \text{Logit} \ (RSTAT) = \alpha_0 + \alpha_1 \text{TURN} + \alpha_2 \text{PL} + \alpha_3 \text{EMPL} + \alpha_4 \text{FDI} + \varepsilon
  \]

<table>
<thead>
<tr>
<th>Variables</th>
<th>Measures/concepts</th>
</tr>
</thead>
<tbody>
<tr>
<td>Residency status (RSTAT)</td>
<td>dummy dependent variable; takes value 0 if, in 2020, the firms’ ultimate investor was resident in (one of) the IIC and 1 otherwise</td>
</tr>
<tr>
<td>Turnover (TURN)</td>
<td>turnover (EUR thousand) in 2020</td>
</tr>
<tr>
<td>Profit/loss (PL)</td>
<td>profit or loss (EUR thousand) in 2020</td>
</tr>
<tr>
<td>Number of employees (EMPL)</td>
<td>average number of people employed on a full time basis in 2020</td>
</tr>
<tr>
<td>FDI position (FDI)</td>
<td>FDI position (EUR thousand) at end-2020</td>
</tr>
</tbody>
</table>
Romania’s FDI figures - 2020

flows € 3 bln.

position € 91 bln.

FDI positions by top investing countries at end-2020 (bln. €)

Data analysis

- ≈22% of the sampled entities UIC ≠ IIC (i.e. the immediate investor is not resident in the same country as the ultimate investor)

- the independent variables are not normally distributed and record extreme values → data were winsorized at the 10% and the 90% percentile.
On average, FDI entities that have their final investor resident in one of the IIC recorded on average lower turnover and profit levels, while also employing less people during 2020 than FDI entities for which the UIC and IIC differ.

Central trend indicators also suggest that entities controlled by an IIC resident final investor exhibit smaller FDI positions at end-2020 than their counterparts.

The t-test and Mann-Whitney test revealed that on average all indicators differ significantly between the two groups of observations (p-value = 0.0000).

Pearson R and Spearman R coefficients suggest positive correlations between all the variables included in the analysis, ranging between 0.0245 and 0.7087.

<table>
<thead>
<tr>
<th>Variable</th>
<th>RSTAT = 1; UIC &lt;&gt; IIC</th>
<th>RSTAT = 0; UIC = IIC</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Mean</td>
<td>Median</td>
</tr>
<tr>
<td>TURN</td>
<td>10,776</td>
<td>4,824</td>
</tr>
<tr>
<td>PL</td>
<td>459</td>
<td>157</td>
</tr>
<tr>
<td>EMPL</td>
<td>92</td>
<td>42</td>
</tr>
<tr>
<td>FDI</td>
<td>6,631</td>
<td>2,395</td>
</tr>
</tbody>
</table>
Results

• The model revealed that all the independent variables of the study are predictive of the residency status of the final investor at a confidence level of 90%.

• Results of the Hosmer–Lemeshow test show that the overall model fit is 45.48%.

• Key findings:
  ➢ higher levels of turnover, profit and FDI increase the likelihood of an enterprise being controlled by an ultimate investor that is not resident in an IIC → H1 ✓ H2 ✓ H4 ✓
  ➢ contrary to what central trend indicators initially suggested, the logit model revealed that the more people an entity employs, the more likely it is to be ultimately controlled by an investor resident in the IIC → H3 ✗

• These results may be of significant importance for:
  ➢ private companies interested in becoming attractive investment avenues for certain MNEs, by informing them which firm-characteristics are of interest to the final business-conglomerate owner;
  ➢ academia, by enabling researchers to better understand the dynamics of the business environment and the economic relationships that form between certain geographical regions.
Conclusions

• Findings suggest that specific firm-size features are predictive of the residency status of their final investors.

• The analysis focused exclusively on Romanian resident FDI entities and used firm level data collected via the 2020 NBR and NIS statistical survey.

• Limitations:
  ➢ results were obtained and are statistically significant for Romania’s particular case;
  ➢ factors that may have a significant influence on whether an FDI entity is likely to be controlled by a final investor resident in an IIC have not been included in the model due to the difficulty to operationalise them at this stage;
  ➢ to date, no entities that meet the criteria stated in the SPE definition have been identified in Romanian economy;
  ➢ hypotheses were tested on NBR’s FDI dataset for 2020. This particular year was marked by the outbreak SARS-CoV-2 pandemic that caused FDI flows to drop, determined FDI enterprises to suspend their business operations and prompted the government to implement financial aid schemes that may distort firm-level indicators. Expanding the analysis for a longer period of time might yield different results.
Thank you for your attention!

Cătălina-Florentina Pricope
Statistics Department
25 Lipscani street, Bucharest, Romania, postcode 030031
T: +4031 132 2154  E: catalina.pricope@bnro.ro
www.bnro.ro
Who stands behind European FDI investors?
A novel characterisation of pass-through within the EU

Carles Gómez-Llabrés, Fausto Pastoris and Martin Schmitz,
European Central Bank

---

1 This presentation was prepared for the conference. The views expressed are those of the authors and do not necessarily reflect the views of the BIS, the IFC or the central banks and other institutions represented at the event.
Who stands behind European FDI investors? A novel characterisation of pass-through within the EU

Fausto Pastoris, Martin Schmitz, Carles Gómez Llabrés

Abstract

Macro-statistics on foreign direct investment (FDI) provide an indication of bilateral direct investment linkages between country pairs. However, with a large part of FDI being intermediated by financial centres, official FDI data often depict oversized linkages involving these intermediate countries. We exploit Orbis company-level data on corporate ownership to break down who is the ultimate owner behind bilateral investing relationships. We thus add an additional dimension to standard FDI data, namely the shares of a country’s FDI activity which can be attributed to domestic corporate groups and to foreign-controlled corporations, respectively. Focusing on intra-EU FDI, we show that for the group of EU FDI hubs (Belgium, Cyprus, Hungary, Ireland, Luxembourg, Malta and the Netherlands), foreign corporations control most of the outward foreign investment activity. We also show that more than a quarter of intra-EU FDI is controlled by non-EU corporations, with the US standing behind, as the ultimate owner, a share even larger than that of any individual EU country. We also provide evidence on the sectoral characterisation of intra-EU FDI, with a large share of FDI ultimately originating from manufacturing and ICT groups being intermediated by EU FDI investors belonging to sectors associated with holding companies and other special purpose entities.

Keywords: foreign direct investment, financial globalisation, external statistics

JEL classification: C82, F21, F62

Contents

1. Introduction ....................................................................................................................................... 2

2. Methodology ..................................................................................................................................... 4

2.1 Decomposing outward FDI by ultimate controlling parent ........................................ 4

2.2 Building up intra-EU FDI triads from Orbis firm-level data ........................................... 7

3. Main findings ................................................................................................................................... 10

4. Conclusion ........................................................................................................................................ 17

References ................................................................................................................................................ 18

1 European Central Bank, External Statistics and Sector Accounts Division. Carles Gómez Llabrés was associated with the ECB when the work for this project was conducted. The views expressed in this paper are those of the authors and do not necessarily reflect those of the European Central Bank.
1. Introduction

Foreign Direct Investment (FDI) data play a key role in analysing economic interconnectedness and integration. FDI is generally assumed to create stable and long-lasting connections between economies, reflecting investing decisions based on long-run factors. FDI statistics are key metrics used to investigate the attractiveness and competitiveness of economies, their role in global production networks and the economic impact generated by foreign investors. Bilateral FDI data are important indicators to understand countries’ bilateral economic ties and financial exposures.

With large part of global FDI nowadays being intermediated by financial centres, FDI data have become an imperfect indicator for analysis of economic linkages and bilateral exposures. Complex corporate structures, financial engineering, and fiscal optimisation strategies by multinationals (MNEs) often involve channelling large amounts of equity and intra-group debt through subsidiaries located in financial centres, before reaching their final investing destinations. Official FDI statistics only capture the investment links between immediate partner countries, to identify where financial claims and liabilities are created and held; however, they do not provide information with respect to the ultimate source of the investment and thus on who takes the underlying investment decisions or bears the ultimate risks. FDI datasets nowadays show outsized amounts of FDI recorded vis-à-vis financial centres and a distortion in the geography of major countries’ bilateral exposures.2

Interest is high among researchers and policy-makers to re-map the size and geography of global cross-border financial links filtering out the role of intermediate financial centres and providing a better representation of the underlying economic connections. Damgaard et al. (2019) show that investment into corporate structures with no economic substance (“phantom” FDI) may account for almost 40% of global FDI and, only if removing the role of special purpose entities (SPEs), the true structure of the global FDI network can be uncovered. Borga and Caliandro (2018) provide a methodology to estimate consolidated FDI statistics removing pass-through capital and detail how the amount of pass-through capital in operating affiliates is quite extensive, accounting for around 25% of inward FDI positions – excluding resident SPEs – in a sample of European countries. Gregori et al (2019) build up a database on foreign ownership of EU firms, which singles out the importance of non-European ultimate owners of companies active in the EU. Bertaut and Curcu (2019) and Coppola et al. (2020) focus on the large size of portfolio securities issuance by MNEs’ subsidiaries resident in offshore financial centres and their estimates uncover a larger exposure of advanced economies to emerging markets than visible in official residency-based data.

The international statistical community is undergoing an overall review of the methodology underpinning the compilation of macroeconomic statistics and ensuring a harmonised approach in the development of supplementary FDI indicators

2 See Blanchard and Acalin (2016) and Lane and Milesi-Ferretti (2017) for an overview on the main issues associated to the interpretation of the recent developments in global FDI.
is high on the agenda. Several countries have recently extended their regular FDI publications by introducing supplementary indicators providing additional information on e.g. the breakdown of inward FDI positions by ultimate investing economy, the breakdown of outward FDI positions by ultimate host economy, the breakdown of FDI associated to the activities of special purpose entities (SPEs).

Our paper contributes to the ongoing discussions on providing supplementary FDI indicators. We use a simple framework that decomposes countries’ outward FDI by the residency of the ultimate controlling investor to shed additional light on the structure of a country’s foreign investment activity. Decomposing outward FDI by ultimate controlling parent, the amount of a country’s investment abroad ultimately due to decisions of domestic companies (i.e. the ultimate controlling parent is a domestic company) can be separated from that ultimately controlled by foreign companies (i.e. the ultimate controlling parent is foreign). This analytical presentation of FDI is a useful complement to the standard FDI presentations as it allows to disentangle the part - if any - of foreign investment activities which are beyond the control of the investing country and thus may be driven by different motives and intentions compared with the part of FDI abroad which is domestically controlled.

Focusing on intra-EU FDI we show that more than a quarter of these linkages is actually controlled by corporate groups from outside the EU, with a prominent role of US corporations which ultimately stand behind a major part of the investment activities of large EU FDI hubs, such as Ireland, Luxembourg and the Netherlands. While countries within the EU share strong commercial, economic and financial ties, bilateral FDI among European countries does not only reflect the strong economic integration within the block, but strongly depends on the organisational choices of large multinationals which make use of EU financial centres as hubs for their global operations. Standard FDI data would show that a small number of EU FDI hubs (Belgium, Cyprus, Hungary, Ireland, Luxembourg, Malta, Netherlands), which combine to only around of 10% EU GDP, are responsible for almost 60% of intra-EU FDI as immediate FDI investors. We make use of information on corporate ownership structures and balance sheet information at the company-level to go beyond the standard FDI presentation and understand the residency of the ultimate corporate owner of EU direct investors. We can thus provide a novel characterisation of intra-EU FDI patterns, breaking down each investing position by the residency of the ultimate corporate owner.

---

3 See, for example, the Direct Investment Task Team guidance note “D6 Ultimate Investing Economy/Ultimate Host Economy and Pass-through Funds” at https://www.imf.org/en/Data/Statistics/BPM/DITT

4 See, for example, the OECD FDI dataset for an overview of data availability on these supplementary indicators for OECD countries. Eurostat is planning to publish EU countries’ FDI based on ultimate investing economy concept in the course of 2022. The IMF has recently published a new dataset on SPEs: https://www.imf.org/en/News/Articles/2022/03/02/pr2253-imf-announces-the-release-of-a-new-database-on-special-purpose-entities

5 Average over the period 2013-2018 for intra-EU FDI assets, according to data collected by the ECB in the context of quarterly Balance of Payments dataset. In this paper, given our latest data point is for 2018, we consider the EU in its EU28 composition, thus still including the United Kingdom.
2. Methodology

2.1 Decomposing outward FDI by ultimate controlling parent

The standard presentation of the geography of FDI is on an immediate counterparty basis, allowing to understand where the funds invested in the domestic economy come from or where the domestic funds are invested: for inward FDI, this reflects the country from which the domestic company directly receives foreign funds; for outward FDI, the country to which the domestic direct investor provides funds. Given that, in the presence of complex corporate structures by large multinationals, direct investment relationships may span through long chains of ownership across multiple countries, alternative analytical presentations have emerged, moving from an immediate towards an ultimate counterparty basis. For inward FDI, the ultimate basis shows the country of the ultimate direct investor (ultimate investing country, or UIC). This presentation allows to look-through the intermediate ownership chain of owners to understand who is ultimately controlling the foreign investment in the domestic enterprise.

For outward FDI, the presentation by ultimate basis looks-through the intermediate holding-companies in foreign countries to identify the location of the final investment (ultimate host country, or UHC). An additional dimension of the geography of FDI on an ultimate basis, namely the decomposition of outward FDI by ultimate controlling parent, distinguishes the amount of a country’s investment abroad ultimately due to decisions of domestic companies (i.e. the ultimate controlling parent of the domestic direct investor abroad is a domestic entity) from that which is ultimately controlled by foreign companies (i.e. the ultimate controlling parent of the domestic direct investor abroad is foreign). This presentation is a useful complement to inward FDI by UIC and outward FDI by UHC, as it shows, for each country engaging in direct investment abroad, the ultimate control on its foreign investing activity, thus allowing to disentangle the part of foreign investment genuinely depending on ultimate decisions by domestic firms. This presentation thus provides a broad definition of “pass-through” FDI, singling-out the amounts of FDI invested abroad, which are under ultimate control of non-domestic corporations.6

Following this approach, outward foreign direct investment by country \( i \) (investor country) into country \( h \) (host country) on an immediate basis can be decomposed, on an ultimate controlling parent basis, in three components:

\[
FDI_{i,h} = FDI^i_{i,h} + FDI^h_{i,h} + \sum_{f \in F} FDI^f_{i,h} \tag{1}
\]

- a domestic-controlled component, \( FDI^i_{i,h} \), which reflects investment by country \( i \) into country \( h \), controlled by ultimate domestic investors based in country \( i \);

---

6 For an overview on the current discussions in the statistical community on the topics of FDI by UIC/UHC and pass-through FDI, see the Direct Investment Task Team guidance note “D6 Ultimate Investing Economy/Ultimate Host Economy and Pass-through Funds” at https://www.imf.org/en/Data/Statistics/BPM/DITT
• a host country controlled (or round-tripping) component, \( FDI_{i,h} \), which reflects investment by country \( i \) into country \( h \), controlled by ultimate investors from host country \( h \);

• a foreign-controlled (or third country) component, \( \sum_{f \in F} FDI_{i,h} \), which reflects investment by country \( i \) into country \( h \), controlled by ultimate investors from the set of foreign (third) countries \( F \).

The round-tripping component and the foreign-controlled-component represent the part of outward FDI outside the ultimate control of the domestic economy and whose dynamics and underlying motivation may differ from those of domestic-controlled outward FDI. This decomposition of outward FDI is particularly relevant for analysing and understanding investment patterns by countries which host a large proportion of companies engaged in direct investment abroad which are at an intermediate level in a corporate ownership chain, as their outward FDI positions may be strongly inflated by the round-tripping and foreign-controlled components, in addition to the genuine domestic-controlled investment abroad.

Aggregating investor country’s bilateral investment links \( FDI_{i,h} \) across the full set \( W_1 \) of possible host countries \( h \), one obtains the investment of country \( i \) in the rest of the world: \( FDI_{i,W_1} \). This overall investment abroad by country \( i \) can be decomposed showing which part of a country’s investment abroad is domestic-controlled, host country-controlled and third country-controlled:

\[
FDI_{i,W_1} = FDI_{i,W_1}^{i} + FDI_{i,W_1}^{h} + FDI_{i,W_1}^{f} \tag{2}
\]

If we define \( \theta_i = FDI_{i,W_1}^i / FDI_{i,W_1} \) as the proportion of a country’s investment abroad controlled by domestic enterprises, \( (1 - \theta_i) \in [0,1] \) describes country’s propensity of being an intermediate FDI investing country. In the extreme case where \( (1 - \theta_i) = 1 \), country’s investment abroad can be fully ultimately attributed to non-resident owners and thus country \( i \) acts purely as an intermediate FDI investing country.

Aggregating all outward investment positions \( FDI_{i,W_1} \) across the full set \( W_0 \) of possible investor countries \( i \), one obtains the overall global FDI: \( FDI_{i,W_0} \). The overall global FDI position can be decomposed showing how much of the global FDI is ultimately controlled by each country \( i \):

\[
FDI_{i,W_0} = \sum_{i \in W_0} FDI_{i,W_0}^i \tag{3}
\]

Where, for each country \( i \), \( FDI_{i,W_0}^i = FDI_{i,W_1}^i + \sum_{j \neq i} FDI_{j,W_1}^i \) measures the amount of global FDI country \( i \) ultimately controls, either as immediate direct investor or, indirectly, by controlling intermediate investing links in countries \( J \), thus showing the overall ultimate share of country \( i \) on global FDI amounts.

This type of decomposition of global FDI does not aim at directly removing any pass-through investment from the official figures: its purpose is to distribute the total amounts of FDI to the countries which are ultimately controlling the activities of intermediate FDI investors. Intermediate outward FDI linkages are not removed from the data, but they are singled-out as being controlled by a country different from that of the intermediate FDI investor.

We can illustrate the added value of this supplementary FDI framework with a real-life example. Figure 1 shows a (simplified) version of the corporate structure of the Mexican group Cemex at end-2017. The foreign operations of the Mexican group
Cemex are held by a (holding) company in the Netherlands (New Sunward Holding B.V.), which controls a Spanish subsidiary (CEMEX Espana S.A.), which in turn controls, directly or indirectly through other subsidiaries, most of the group activities in the rest of the world, including all the other group’s subsidiaries in other European countries.

Within the standard FDI framework, only bilateral links between the immediate investor and host country are depicted and, in this case the following bilateral links would emerge from the initial layers\(^7\) of the Cemex FDI structure:

- investment by Mexico into Netherlands (contributing to the aggregate Mexican FDI assets)
- investment by Netherlands into Spain (contributing to the aggregate Dutch FDI assets)
- investment by Spain into France and the UK (contributing to the aggregate Spanish FDI assets)

In the standard FDI framework there would be no trace of Mexican (equity) FDI investment in any other foreign country, apart from the FDI direct link with the Netherlands, which sits at the top of Cemex’s structure; there would be no way to disentangle that e.g. part of Spanish overall FDI assets are ultimately controlled by a Mexican group.

Within the supplementary framework suggested in this paper, we would be able to complement the information on the immediate bilateral links with additional information on the nationality ultimately controlling the investment chain. What would emerge would be the following decomposition of bilateral FDI links:

- investment by Mexico into Netherlands, ultimately controlled by Mexico (contributing to the aggregate Mexican FDI assets, controlled by domestic Mexican companies)
- investment by Netherlands into Spain, ultimately controlled by Mexico (contributing to the aggregate Dutch FDI assets, controlled by foreign companies)
- investment by Spain into France and the UK, ultimately controlled by Mexico (contributing to the aggregate Spanish FDI assets, controlled by foreign companies)

The suggested framework, fully consistent with the standard FDI presentation, provides useful information on who ultimately controls intermediate FDI linkages worldwide.

\(^7\) For ease of exposure we consider here only the initial layers of the structure, as encircled with red rectangles in Figure 1.
2.2 Building up intra-EU FDI triads from Orbis firm-level data

Our aim is to provide an enhanced dataset on intra-EU FDI positions where, for each bilateral FDI link between EU countries, the breakdown by the geography of the ultimate controlling investor is also included. Since detailed information on the geography of the ultimate controlling parent (UCP) of FDI abroad is not available from standard macroeconomic data publications, we build-up the details of our enhanced FDI dataset from Bureau van Dijk’s Orbis microdata on companies’ balance sheets.

---

8 We focus on intra-EU positions mainly for two reasons: first, as we are interested in the extent to which the recent surge in intra-EU FDI originates from genuine FDI activities of EU companies. Second, due to our micro-data limitations: the growing literature using Orbis microdata highlights how the coverage of information needed to build-up representative FDI linkages is rather poor for non-European countries.

9 One notable exception is represented by Ireland, where the CSO publishes annual FDI data, with a breakdown of the outward position between domestic and foreign controlled corporations. Comparable information is also present in topical statistical publications for Austria and Switzerland.

As a first step, we identify in Orbis all the ownership relationships giving rise to an FDI link between EU companies, according to the standard FDI definition of intra-EU linkages, that is we extract information on all active ownership linkages between EU companies, where at least 10% of the equity capital of a (EU-resident) company is directly owned by a foreign (EU-resident) company. For each of these ownership relationships we calculate the corresponding FDI amounts, multiplying the value of the equity capital on the balance sheet of the subsidiary by the relative ownership share of each direct investor. In this way we obtain a fully-fledged micro-level FDI dataset consisting of subsidiary-by-investor FDI links. These micro-level linkages are in line with the standard definition of intra-EU FDI, as they show only the direct linkage between a subsidiary and its FDI immediate shareholder(s). We then transform the subsidiary-by-investor dataset into a subsidiary-by-investor-by-ultimate controlling parent dataset, adding to each micro-level FDI link additional information based on the residency of the ultimate controlling parent (UCP), that is the company ultimately controlling the intra-EU direct investor. While the subsidiaries and direct investors in our dataset are limited to residents within the EU, the residency of the UCP is left unrestricted, so that the country of ultimate control can be both an EU or extra-EU country. At this point, we can simply aggregate the company-level FDI triads (i.e. the subsidiary-by-investor-by-ultimate controlling parent company linkages) across their country dimension, to obtain a macro-level dataset of intra-EU FDI triads: in this way, we obtain a dataset where, for each EU FDI investing country, we can identify not only the set of host EU countries, but also the set of countries (including the domestic country) ultimately controlling these bilateral links – it is the implementation of the decomposition of bilateral outward FDI as shown in equation (1).

Several important choices and assumptions had to be made when building the extended FDI dataset from the available corporate information from Orbis.

First, identification of FDI links in Orbis is possible only for equity assets, while there is no information on intra-group debt. In practice, estimation of FDI equity linkages from Orbis is done through combining information on subsidiaries’ total shareholders’ funds and information of direct investors ownership percentages and residency. Once all the EU enterprises with foreign subsidiaries in other EU countries are identified, their ownership percentage in their foreign subsidiaries is multiplied by the shareholders’ funds to estimate the equity claim the direct investors has on the subsidiaries. While Orbis also presents balance sheet information on debt liabilities for each company, there is no immediate way to understand who the holder of this debt is. Similarly to Borga and Caliandro (2019) we assume debt financing follows the same pattern as equity financing in terms of the country-by-country distribution of ultimate controlling parents, thus we do not engage in any explicit estimation of FDI debt from Orbis company-links – as instead done in Damgaard et al (2019) on the basis of estimates from Orbis information on overall companies’ debt and equity-debt proportions in macro-FDI datasets.

10 If the company identified as immediate direct investor is not controlled by any further corporation, there will be a complete equivalence between the residency of the immediate investor and ultimate controlling parent. The same happens if the two companies differ, but they are resident in the same country.
Second, in the identification of the country of the ultimate controlling parent, we assume each immediate FDI investor is fully controlled by its ultimate controlling parent. Focusing on identifying the country of the entity that exerts control over the direct investor is fundamental in our approach since we want to understand if immediate direct investors are ultimately controlled from foreign countries or domestically. As controls assumes that an entity owns more than 50% (directly or indirectly) of the voting power in a subsidiary, ultimate control can be exercised even if the ultimate controlling parent controls (directly or indirectly) only 51% of the equity of the immediate direct investor. In this case, we attribute the full 100% of the immediate FDI relationship to the ultimate controlling parent, even if the effective control is lower.\textsuperscript{11} This approach is referred to as the winner-takes-all methodology.\textsuperscript{12}

Third, in the identification of the country of the ultimate controlling parent we focus on ultimate corporate control, thus not investigating if there is any further cross-border control relationship on top of the last corporate layer. While it is normally the case that the residency of the controlling entity at the top of the corporate group is the same as the natural persons who may in turn hold control of this entity, there are some particular constructs for which this does not hold true. This would be the case when a natural person from country X controls the corporate group Y, through a holding company established in country Z: our methodology would show country Z as the country of the ultimate controlling parent, while a further cross-border layer exists involving control by natural persons. Given the predominance of these holding structures in some limited jurisdictions (e.g. Luxembourg and the Netherlands), our baseline estimates may provide some over-estimation of ultimate control attributed to countries hosting these holding companies against countries where the natural persons behind the holding companies reside.

Fourth, some of the company-level triads identified in Orbis miss information on the residency of the ultimate controlling parent. In this case, for the largest missing bilateral intra-EU links (top 100 missing links for each investor country) we impute manually information from external data-sources\textsuperscript{13}, while for the remaining links we assume that the missing information is distributed in the same proportion as the existing ultimate controlling parent information for that investor country. Overall, for each EU investor country less than 5% of value of FDI needs this additional imputation.

Finally, while our framework is based on the residency of the ultimate corporate layer as identified in Orbis, our baseline estimation provides a specific treatment for US corporate inversions and for Chinese offshore listed abroad registered companies. Information on the ultimate corporate controlling parent in Orbis reflects the place of incorporation of the ultimate controlling corporate layer in the group structure: while this information normally should provide sufficient information to establish which country hosts the company controlling EU FDI investors, in the above mentioned cases there is an evident mismatch between the place of incorporation of

\textsuperscript{11} Please note that, instead, to build the initial immediate investor-immediate subsidiary relationship we follow the usual proportional approach where each immediate investor is attributed an FDI share in each subsidiary proportionally to the effective capital share invested in the subsidiary. This is consistent to the usual FDI definition.

\textsuperscript{12} For further information see the Direct Investment Task Team guidance note “D6 Ultimate Investing Economy/Ultimate Host Economy and Pass-through Funds” at https://www.imf.org/en/Data/Statistics/BPM/DITT

\textsuperscript{13} We mainly use information from companies’ annual reports and statistical business registers.
the ultimate corporate controlling company in Orbis and the country where corporate control is actually exercised. Recent studies by Tabova (2020) and Coppola et al. (2020) show how these corporate constructions have a strong distortionary impact on understanding the geography of global corporate control. Thus, we provide an exception to our framework based on the residency of ultimate corporate parents as identified in Orbis and we do the following: 1) US corporations, which had inverted to Ireland and are identified in Orbis as having an Irish ultimate controlling parent, are re-allocated to the US as the country of ultimate control; 2) Chinese corporations listed offshore and incorporated outside mainland China are re-assigned to China as the country of ultimate control. Overall, the impact is rather relevant in our sample for US-inverted companies – around 30% of Irish intra-EU FDI is reassigned from being domestically controlled to being foreign controlled from the US - while it is rather insignificant for Chinese offshore companies (only around 1bn EUR is reallocated for 2018) – due to the fact that these companies do not appear as ultimate owners behind large intra-EU FDI linkages in our dataset.

Overall, the coverage rate of our dataset compared with standard macroeconomic statistics is encouraging and in line with previous studies on constructing representative macroeconomic datasets using Orbis data: specifically, our dataset covers around 65% of intra-EU FDI positions in 2018 and provides a similar picture as official statistics with regards to the country distribution of EU immediate FDI investors.

3. Main findings

Breaking down FDI linkages between EU countries following the approach shown in Section 2, we provide a novel characterization of the intra-EU FDI network, disentangling how much of outward FDI from each EU country is ultimately controlled by domestic or foreign companies.

European FDI statistics have been heavily influenced by multinational groups setting-up part of their corporate structures in several EU FDI hubs, such as the Netherlands, Luxembourg or Ireland, to benefit from tax advantages and other incentives, such as enhanced ownership rights and easy access to capital markets or sophisticated financial services. A sizeable proportion of subsidiaries set-up in these EU FDI hubs consist of holding entities, which sit at the top of other foreign activities of the group and thus have disproportionately large foreign holdings on their balance sheet, generating large amounts of FDI assets for the host country, without correspondingly generating much employment nor having any significant physical presence in the host country. When this type of corporate ownership structures emerges, countries hosting holding companies emerge as large FDI investors, even though they are mostly re-investing funds which are ultimately controlled from somewhere else. Official statistics show (see Chart 1) that a small number of EU FDI hubs (Belgium, Cyprus, Hungary, Ireland, Luxembourg, Malta, Netherlands), which combine to only around of 10% EU GDP, acts as (immediate) investor in around 60% of intra-EU FDI activity in 2013-2018, a larger investing role than the largest 5 EU economies (which contribute to around 70% of EU GDP).
Who stands behind European FDI investors?

11

**Chart 1: Main intra-EU investors in official FDI statistics**

Proportion of overall intra-EU FDI assets

<table>
<thead>
<tr>
<th>Year</th>
<th>EU FDI hubs</th>
<th>EU Big 5</th>
<th>Other EU</th>
</tr>
</thead>
<tbody>
<tr>
<td>2013</td>
<td>0.60</td>
<td>0.40</td>
<td>0.00</td>
</tr>
<tr>
<td>2014</td>
<td>0.60</td>
<td>0.40</td>
<td>0.00</td>
</tr>
<tr>
<td>2015</td>
<td>0.60</td>
<td>0.40</td>
<td>0.00</td>
</tr>
<tr>
<td>2016</td>
<td>0.60</td>
<td>0.40</td>
<td>0.00</td>
</tr>
<tr>
<td>2017</td>
<td>0.60</td>
<td>0.40</td>
<td>0.00</td>
</tr>
<tr>
<td>2018</td>
<td>0.60</td>
<td>0.40</td>
<td>0.00</td>
</tr>
</tbody>
</table>

Source: Authors’ calculations based on quarterly balance of payments data, ECB.
Note: “EU FDI hubs” include BE, CY, IE, LU, MT, NL, HU. “EU Big 5” includes GB, DE, FR, IT, ES. “Other EU” includes the remaining EU countries not singled-out in the chart. EU is intended as EU-28, thus including the current EU-27 composition plus GB. FDI includes both equity and debt assets.

However, these figures from official statistics do not allow us to investigate what is behind the oversized role these EU FDI hubs retain within intra-EU FDI linkages: are they ultimately controlling the large FDI positions they originate as immediate investors? Or are they simply intermediating FDI linkages ultimately controlled by other large EU economies or by extra-EU advanced or emerging economies? Results from our enhanced FDI dataset, which is based on aggregating micro-level information from Orbis on immediate FDI linkages and ultimate corporate control, help us putting this disproportionate role of EU FDI hubs within the EU FDI network into context.

While Luxembourg, the Netherlands and Ireland, figure prominently among the main FDI investors within EU, accounting for around 50% of the intra-EU FDI investing activity as immediate investors, their role as ultimate investors is much more limited, as they ultimately control only around 10% of this investment activity (see Chart 2). The same phenomenon appears for the remaining FDI hubs (Belgium, Cyprus, Malta, Hungary) which have much smaller contributions as ultimate controlling investors of intra-EU FDI compared with those as immediate FDI investors. The opposite – however not fully compensating – result is observed for several large EU countries (the UK, Germany, and France) where our dataset shows that they are behind a larger share of intra-EU FDI as ultimate investors than they are as immediate investors, providing some insights that part of the investment activity they ultimately control spans through several EU countries. For the remaining EU countries our data show that their FDI positions do not substantially differ when considering them as immediate or ultimate investors behind intra-EU FDI linkages.14

---

14 This may come from two situations: the first option is that a country’s immediate FDI activity is fully domestically controlled and its domestic FDI groups do not create investment chains in other EU countries, thus having an equal value within the intra-EU FDI network as immediate and ultimate investor; but this situation of similarity of immediate and ultimate share in the EU FDI network can also materialise if a country’s immediate FDI activity is only partly domestically controlled, but its domestic groups create investment chains in other EU countries.
Who stands behind European FDI investors?

Chart 2: EU countries’ contribution to intra-EU FDI: role as immediate vs ultimate investor

Country’s proportion of intra-EU FDI assets, 2018

Source: Authors’ calculations based Orbis data.
Note: Yellow bars represent EU countries’ contribution to intra-EU FDI as immediate source of FDI investment: they sum up to 1; blue bars represent EU countries’ contribution to intra-EU FDI as ultimate sources of FDI investment: they do not sum-up to 1 as some of the intra-EU FDI is ultimately controlled by extra-EU countries.

Overall, EU countries appear as ultimate controlling investors of less than 70% of the immediate intra-EU FDI links: the remaining part of the intra-EU immediate FDI links are thus ultimately controlled by companies from outside the EU. Our dataset shows how non-EU countries have a sizeable role as ultimate investors behind intra-EU FDI linkages, with an outsized role of the US. In fact, ranking individual countries by the share they ultimately control in intra-EU FDI (see Chart 3), the US turns out as the largest individual investor, ultimately controlling more than 20% of the FDI linkages between EU countries. This shows that a sizeable part of intra-EU FDI linkages ultimately depends on investing decisions by US corporate groups and thus may react to different incentives and conditions than corresponding intra-EU links controlled by EU groups. These differences in the ultimate source of control of intermediate FDI linkages are not visible in standard FDI datasets, while they can be disentangled in the enhanced framework we propose. In addition to the US, albeit with a minor role, also other advanced and emerging non-EU economies, such as Japan, China, Canada, Mexico and Switzerland, have some relevant role as ultimate investors behind intra-EU linkages: each of them ultimately controls around 1% of intra-EU FDI and combined control more than the intra-EU FDI ultimately controlled by Italy or Spain. With a minor share also Russia and Brazil appear as ultimate investors into intra-EU FDI links, with a similar importance (less than 1%) as some offshore financial centre such as Cayman Islands, Bermuda, Curacao, and Jersey.15

15 As mentioned in Section 2, the importance of these offshore centers as ultimate investors may be overstated in our dataset due to our assumption of stopping at the last corporate layer and not to look-through the nationality of natural persons behind this last corporate control layer. For example, if a British person has set-up a holding corporate structure for some FDI investments in Jersey, the ultimate controlling country identified in our procedure will be Jersey and not the UK.
Turning to the decomposition of each country’s intra-EU outward FDI, our dataset shows that for EU FDI hubs the majority of their FDI activity is ultimately foreign controlled (see Chart 4). For Ireland, Luxembourg, and the Netherlands, who stand among the largest intra-EU immediate FDI investors, only less than a quarter of their intra-EU FDI assets are actually controlled by domestic groups, thus considerably downplaying these countries’ outsized role in the intra-EU FDI network. While for Luxembourg and the Netherlands it is corporate groups from both other EU countries and the US controlling the largest share of their FDI activity in a similar proportion (around 30% each) – plus a smaller, but significant (more than 10%) share of ultimate control from other non-EU groups – it is US groups fully dominating Ireland’s FDI outward activity into other EU countries (controlling around 80% of Irish intra-EU FDI).16 US companies also have a disproportionate influence on intra-EU FDI investment by Hungary, where around 50% is ultimately controlled by US companies. US companies also control a sizeable share (more than 10%) of intra-EU FDI from Belgium, the UK and Denmark. Companies from other extra-EU countries have a sizeable control over intra-EU FDI by Malta and Cyprus, where they ultimately control more than 20% of intra-EU FDI activity. For these two FDI hubs, we see considerable ultimate control coming from offshore centres (e.g. Bermuda, Bahamas, British Virgin Islands), which may signal the presence of a chain of holding structures set-up with the ultimate corporate parent registered directly offshore – but the natural persons behind these investments structures are most likely nationals of third countries rather than nationals of offshore centres. For Cyprus we also see a non-negligible ultimate control from Russian ultimate controlling (corporate) parents.

16 Please note that, as mentioned in Section 2, we consider US companies which had gone through corporate inversions (i.e. redomiciled companies) to Ireland as ultimately controlled by the US in our dataset. These companies alone stand behind around 30% of Irish intra-EU FDI ultimately controlled by the US in our dataset. The outsized role of US companies for Irish FDI is also visible in the publication of annual Irish FDI by the CSO which shows around 40% of outward Irish FDI in 2018 is being controlled by US redomiciled companies: https://www.cso.ie/en/releasesandpublications/ep/p-ia/internationalaccountsq12021final/redomiciledplcs2020/.
For the remaining of the EU countries, domestic groups are controlling the majority of their outward FDI investment activity into other EU economies, with some heterogeneity on the share of FDI domestically controlled and the geography distribution of the residual part, which is foreign controlled. For Germany, France, Italy and Spain more than 80% of the investment into other EU countries is ultimately controlled by domestic groups, in line with the traditional view that outward FDI represents decision to invest abroad by the investing country and not simply pass-through of investment controlled from elsewhere. For the UK it is only around two thirds of their FDI activity being domestically controlled, with a sizeable role of US companies behind the remaining part. However, foreign control of outward FDI does not necessarily signal that the intermediate FDI investor is just a pure holding structure, with no economic impact on the resident economy or any autonomy of decision. In fact, a considerable part of UK FDI abroad which is US-controlled also simply mirrors the fact that several large UK companies with sizeable foreign FDI operations were acquired by US companies over the past decades. Another interesting case among EU countries is Austria: while not often considered among FDI pass-through centres within the EU, our dataset shows that only less than 60% of Austrian FDI into other EU countries is domestically controlled, while around a quarter is controlled by companies from outside the EU, with a strong role by Russian companies as ultimate controlling parents. These results from Austria are in line with their official FDI data, showing that a non-negligible part of their overall FDI assets (around 15%) are invested abroad by resident special purpose entities.\(^\text{17}\) Our approach can be used to enhance this information by providing the geography of the ultimate controlling parent behind these pass-through funds.

Extending the view to the bilateral intra-EU FDI network, the suggested enhanced presentation provides some additional insights on ultimate control behind bilateral

\(^{17}\) According to OECD data for 2018: https://stats.oecd.org/index.aspx?DataSetCode=FDI FLOW PARTNER#
FDI linkages and helps us understanding what is behind large bilateral position between EU countries. It is clear how the largest EU economies, even though their FDI assets in other EU countries are mostly controlled by domestic firms, are still contributing to overall pass-through of FDI funds as they are mostly directing their FDI activity to the largest EU FDI hubs (see Chart 5a). A considerable part of British, German and Italian (domestically-controlled) FDI is directed towards the Netherlands and Luxembourg, most of which will be then re-invested in third countries (both within and outside the EU) or possibly back into the original investing country (round-tripping). The importance of intermediate corporate holding structures within the EU emerges by focusing on the bilateral intra-EU links ultimately controlled by a single EU country (see Chart 5b for the UK - labelled as “GB”). It is evident how UK companies, ultimately controlling around two thirds of UK intra-EU FDI, direct the majority of their direct FDI assets in subsidiaries resident in Luxembourg and the Netherlands (see the large blue links between the UK as investor and Luxembourg and the Netherlands as hosts). However, our representation allows us to also see which part of Luxembourgish and Dutch investment into the EU is ultimately controlled by UK firms and where it is invested to. It is clear how several UK companies have set-up some rather complex corporate structures within the EU with an initial layer in Luxembourg or the Netherlands and then a second layer back in the UK, as shown by the large UK-controlled FDI assets by Luxembourg and the Netherlands in the UK.
Chart 5: Intra-EU FDI Network, enhanced breakdown
2018 data

a. By type of ultimate control

b. By country of ultimate control

Source: Authors’ calculations based on Orbis data.
Some additional evidence of the importance of pass-through for FDI within the EU can be seen by the sectoral classification of EU FDI investors (see Chart 6). Traditional FDI datasets, based on the immediate investor-host concepts, show a disproportionate share of investment abroad coming from economic sectors associated with the activity of holding companies and other special purpose entities. This phenomenon is visible in our dataset where immediate FDI investors from finance and insurance NACE sectors represent almost 40% of intra-EU FDI. These sectors are not traditionally linked to FDI activity. Our additional framework provides information on the sector of the ultimate corporate group behind the immediate FDI linkages and here we observe that the economic sectors traditionally associated with large multinationals, like manufacturing, mining and other industry as well as information and communication, account for around 70% of intra-EU FDI links. This additional sectoral view of FDI by sector of ultimate investing economy is thus very useful to disentangle the economic sectors ultimately behind FDI within the EU.

**Chart 6: Most relevant economic sectors of intra-EU FDI investors**

2018; proportion of total

Source: Authors’ calculations based on Orbis data.
Note: Only Orbis linkages with information on the NACE code for both immediate and ultimate investors are used. Only individual FDI links larger than 1 million EUR are taken into account.

4. Conclusion

Traditional FDI data do not only reflect linkages from production and trade networks, but also multinationals’ fiscal optimisation and financial engineering purposes.

When complex corporate structures exist, with the presence of several layers of intermediate holding companies, interpreting FDI data becomes challenging as positions from intermediate financial centres are inflated by their role as intermediate investors. As several FDI hubs appear as the largest FDI investors worldwide, traditional FDI datasets do not allow to disentangle which part of their investment activity is ultimately controlled by domestic companies and which part is instead...
ultimately controlled and funded by foreigners, thus providing evidence of FDI pass-through. This additional information is provided in the framework suggested in our paper, which suggests decomposing outward FDI by the residence of the ultimate controlling parent. Providing this additional dimension to standard FDI data helps putting in context the large role of financial centres as global FDI investors. When large share of FDI activity is generated by foreign-controlled corporate groups, the overall size of FDI assets does not (fully) reflect domestic-decision making and may react to different determinants.

Exploiting Orbis company-level data on corporate ownership structures, we break down each EU country’s contribution to the intra-EU FDI network into a domestic and foreign ultimately controlled part. We showed that for a group of EU FDI hubs (Ireland, the Netherlands, Luxembourg, Malta, Hungary, Belgium, and Cyprus), foreign groups ultimately control most of the outward foreign investment activity. We also showed that more than a quarter of intra-EU FDI is ultimately controlled by non-EU corporations, with the US as the largest ultimate owner of intra-EU FDI – even exceeding all individual EU countries. We also provided additional evidence on the sectoral characterisation of intra-EU FDI, with a large share of FDI ultimately originating from manufacturing and ICT groups, which is being intermediated in the EU by holding companies and other special purpose entities.

As also highlighted by Borga and Caliandro (2018), despite the relevance of pioneering studies and estimates of supplementary FDI breakdowns based on detailed corporate level microdata via Orbis, it would be optimal for additional breakdowns of FDI data between domestic and foreign controlled to be directly collected and published by the statistical authorities responsible for national FDI datasets. Our exercise thus wants to provide some initial, encouraging results on the level and amount of information which could be additionally obtained by researchers and users of FDI data if these estimates were to be made available for a large set of countries worldwide. The work of the statistical community for the next update of the international methodological manuals should provide an additional push in this direction.

References


Who stands behind European FDI investors?


IMF Direct Investment task Team (DITT) (2021), “D.6 Ultimate Investing Economy/Ultimate Host Economy and Pass-through Funds”, DITT guidance note


Tissot B. (2016), “Globalisation and financial stability risks: is the residency-based approach of the national accounts old-fashioned?”, BIS Working papers No 587
Who stands behind European FDI investors?
A novel characterisation of pass-through within the EU

11th Biennial IFC Conference

F. Pastoris, M. Schmitz, C. Gómez Llabrés
European Central Bank

The views expressed in this paper are those of the authors and do not necessarily reflect those of the European Central Bank.
1. Introduction

**Starting point: oversized importance of EU financial centers as FDI investors**

- **Traditional FDI dataset → EU FDI hubs responsible for almost 60% of intra-EU FDI activity as immediate investors in recent years**

- Not possible to distinguish who is ultimately behind these large FDI investments:
  - Is it EU FDI hubs themselves?
  - Is it other EU large countries?
  - Is it other extra-EU advanced or emerging markets?

---

**Chart 1: Main intra-EU investors in official FDI statistics**

Proportion of overall intra-EU FDI assets

Source: Authors’ calculations based on quarterly balance of payments data, ECB.
Note: “EU FDI hubs” include BE, CY, IE, LU, MT, NL, HU. “EU Big 5” includes GB, DE, FR, IT, ES. “Other EU” includes the remaining EU countries not singled-out in the chart. EU is intended as EU-28, thus including the current EU-27 composition plus GB. FDI includes both equity and debt assets.
2. Methodology

Decomposing standard outward FDI by global ultimate owner (GUO)

\[ FDI_{i,h} = FDI_{i,h}^i + FDI_{i,h}^h + \sum_{f \in F} FDI_{i,h}^f \]  \hspace{1cm} (1)

Standard bilateral FDI from country \( i \) to country \( h \) decomposed in:

- a **domestic-controlled component**, \( FDI_{i,h}^i \),
- a **host-country controlled** (or round-tripping) component, \( FDI_{i,h}^h \),
- a **foreign-controlled** (or third country) component, \( \sum_{f \in F} FDI_{i,h}^f \).

Aggregating country \( i \)'s investment over the set \( W1 \) of host countries, country \( i \)'s FDI position can then be decomposed into **domestic-controlled** and **foreign-controlled** (including round-tripping component).

\[ FDI_{i,W1} = FDI_{i,W1}^i + FDI_{i,W1}^F \]  \hspace{1cm} (2)
2. Methodology

A (simplified) real example of a global MNEs corporate structure and extended FDI recording

Standard FDI data: INV.HOST bilateral info
- MX.NL → MX.W1
- NL.ES → NL.W1
- ES.FR; ES.UK → ES.W1

Extended FDI data: (UCP)_INV.HOST triad info
- (MX)_MX.NL → (DOM)_MX.W1
- (MX)_NL.ES → (FOR)_NL.W1
- (MX)_ES.FR; (MX)_ES.UK → (FOR)_ES.W1

- Allows to separate countries’ FDI positions in domestically and foreign controlled
- Fully consistent with standard FDI presentation
- Provide info on who ultimately controls FDI links worldwide
3. Main findings

Stark discrepancy in EU countries’ relevance as immediate vs ultimate owners of intra-EU FDI

Chart 2: EU countries’ contribution to intra-EU FDI
Country’s proportion of intra-EU FDI assets, 2018

Source: Authors’ calculations based Orbis data.
Note: Yellow bars represent EU countries’ contribution to intra-EU FDI as immediate source of FDI investment: they sum up to 1; blue bars represent EU countries’ contribution to intra-EU FDI as ultimate sources of FDI investment: they do not sum-up to 1 as some of the intra-EU FDI is ultimately controlled by extra-EU countries.

Chart 3: Who is ultimately behind intra-EU FDI linkages?
Ultimate controlling parents of intra-EU FDI assets, proportion of overall intra-EU FDI, 2018

Source: Authors’ calculations based on Orbis data.
Note: Countries highlighted in red are extra-EU countries. Only countries which contribute individually (as ultimate corporate owners) to at least 0.2% of intra-EU FDI are shown. Overall, the countries shown in the chart contribute, as ultimate corporate owners, to around 98% of intra-EU FDI.
3. Main findings

For EU FDI hubs foreign companies ultimately control >50% of outward FDI

Chart 4: Domestic vs foreign-controlled intra-EU FDI assets
Domestic vs foreign-controlled, proportion of country’s intra-EU FDI, 2018

Source: Authors’ calculations based on Orbis data.
Note: The blue part of the stacked bat represents the percentage of EU countries’ intra-EU FDI which is controlled by domestic corporate groups; the yellow part shows its intra-EU investment ultimately attributable to corporate groups from other EU-countries; red bars for its intra-EU investment controlled by US corporate groups; green for its intra-EU investment controlled by corporate groups from outside the EU (and not US groups).

FDI hubs’ FDI largely controlled by foreign groups

**US groups** particularly relevant for FDI by IE, HU, LU and NL

**Other EU groups** particularly relevant for FDI by LU, NL, MT, BE

**Other extra-EU groups** particularly relevant for FDI by MT, CY, AT
3. Main findings

Decomposing the intra-EU bilateral FDI Network by country of ultimate owner

Chart 5: Intra-EU FDI linkages, by type of ultimate owner
EUR billion, 2018

Considerable part of domestic-originated FDI from large EU countries is invested to EU FDI hubs:

- Stays in the country
- Re-invested in other EU countries
- Round-tripping
- Re-invested outside the EU

Source: Authors’ calculations based on Orbis data.
Note: Only EU countries representing at least 1% of intra-EU FDI as immediate investors in our dataset are shown.
The size of the grey boxes (nodes) on the Y-axes represents the country’s proportion of intra-EU FDI as immediate
investor and immediate host. The size of the links shows the proportion of bilateral FDI linkages with respect to the
intra-EU FDI. The colouring of the links represents enhanced information on the ultimate controlling parent behind
otherwise standard bilateral FDI links.
3. Main findings

Stark discrepancy in sectoral relevance as immediate vs ultimate owner of intra-EU FDI

Chart 6: Most relevant economic sectors of intra-EU FDI investors
Proportion of total, 2018

FDI by sector of ultimate investors provides additional insights on sectors behind FDI.

Finance and insurance and Professional, scientific and administration see a reduced role (SPEs, holding/head office).

Manufacturing, mining and other industry and ICT see an increased role.

Source: Authors’ calculations based on Orbis data.
Note: Only Orbis linkages with information on the NACE code for both immediate and ultimate investors are used. Only individual FDI links larger than 1 million EUR are taken into account.
4. Conclusions

Main takeaways

• Providing an additional dimension to standard FDI data helps putting in context the large role of financial centers as global FDI investors.

• If a large share of FDI activity is generated by foreign-controlled corporate groups, the overall size of FDI assets reflects domestic-decision making only to a limited extent.

• Information on the ultimate owners of resident FDI investors could be rather easily available to FDI compilers and can thus be added as a supplementary dimension to standard FDI datasets.

• A sectoral breakdown of the ultimate owners behind intermediate FDI linkages improves the understanding of the global network of FDI linkages by industry.
A typology of captive financial institutions in Luxembourg: lessons from a new database

Gabriele Di Filippo and Frédéric Pierret,
Central Bank of Luxembourg

---

1 This presentation was prepared for the conference. The views expressed are those of the authors and do not necessarily reflect the views of the BIS, the IFC or the central banks and other institutions represented at the event.
A Typology of Captive Financial Institutions in Luxembourg: Lessons from a New Database

Gabriele Di Filippo  Frédéric Pierret
Department of Statistics  Department of Statistics
Banque centrale du Luxembourg  Banque centrale du Luxembourg

Abstract

The paper draws a typology of captive financial institutions and money lenders (CFIs, sector S127) in Luxembourg from a new database. The latter combines information from three sources: the EuroGroups Register managed by Eurostat, the Statistical Business Register managed by STATEC (the National Institute for Statistics and Economic Studies) and the Central Balance Sheet Register managed by STATEC. The new database enhances coverage of CFIs in Luxembourg. Indeed, it goes beyond the BCL reporting framework (BCL (2014)) by not only including CFIs with total assets larger or equal to EUR 500 million but also those with total assets of less than EUR 500 million. The period of analysis spans 2011 to 2019. Results show that CFIs present different characteristics depending on their balance sheet size. On the one hand, CFIs with total assets larger than EUR 100 million mainly regroup holding companies, intragroup lending corporations, mixed structures and conduits. On the other hand, CFIs with total assets lower than EUR 100 million feature mostly mixed structures. Overall, while holding corporations own the majority of total assets, the largest number of companies consists of mixed structures.

Keywords: Captive financial institutions and money lenders, Sector S127, Typology, EuroGroups Register, Statistical Business Register, Central Balance Sheet Register, Big Data

JEL classification: C80, C81, L22

Disclaimer: This paper should not be reported as representing the views of the Banque centrale du Luxembourg or the Eurosystem. The views expressed are those of the authors and may not be shared by other research staff or policymakers in the Banque centrale du Luxembourg or the Eurosystem.
Contents

1. Introduction ........................................................................................................................................... 3
2. Building of the new database ........................................................................................................ 4
   2.1 Selection of CFIs ...................................................................................................................... 4
   2.2 Accounting information on CFIs ............................................................................................ 4
3. Prototype balance sheets of CFIs ................................................................................................. 5
4. Qualitative approach ......................................................................................................................... 8
5. Results from the new database .................................................................................................... 10
   5.1 Typology of CFIs in Luxembourg .................................................................................... 10
   5.2 Typology of CFIs across balance sheet sizes ...................................................................... 11
       Typology of CFIs: a comparison between the new database and the BCL reporting framework .......................................................................................................................... 11
       Typology of CFIs across ranges of total assets .................................................................... 12
6. Conclusion ........................................................................................................................................... 15
References ................................................................................................................................................ 16

A Typology of Captive Financial Institutions in Luxembourg: Lessons from a New Database
1. Introduction

The paper draws a typology of captive financial institutions and money lenders (CFIs, sector S127) in Luxembourg from a new database. The latter combines information from three sources: the EuroGroups Register managed by Eurostat, the Statistical Business Register managed by STATEC (the National Institute for Statistics and Economic Studies) and the Central Balance Sheet Register managed by STATEC. The new database enhances coverage of CFIs in Luxembourg. Indeed, it goes beyond the BCL reporting framework (BCL (2014)) by not only including CFIs with total assets larger or equal to EUR 500 million but also those with total assets of less than EUR 500 million. This larger coverage enables a more comprehensive analysis of CFIs in Luxembourg compared to previous studies that used the BCL reporting framework (Di Filippo and Pierret (2020a, 2020b)).

However, differences prevail between the new database and the BCL reporting framework. Indeed, the new database provides annual data over the period 2011-2019, while the BCL reporting framework features monthly data over the period December 2014-September 2021. Hence, the BCL reporting framework provides data updates with shorter delays. In addition, contrary to the BCL reporting framework, the new database does not include a breakdown of balance sheet items by geographical counterpart, by maturity and by currency. In spite of these limits, the accounting information available in the new database allows implementing the method presented in Di Filippo and Pierret (2020a) to draw a typology of CFIs. In particular, the new database enables investigating whether small CFIs with total assets lower than EUR 500 million share similar features with larger CFIs whose total assets are at least equal to EUR 500 million.

The remainder of the paper is organised as follows. Section 2 presents the data sources used to build the new database of CFIs in Luxembourg. Section 3 defines the potential types of CFIs. Section 4 describes the method used to identify the prototype balance sheets of CFIs. Section 5 presents the results by exposing the typology for CFIs as a whole, and across various ranges of total assets. Section 6 concludes.
2. Building of the new database

The paper builds a new database of CFIs in Luxembourg by combining information from three datasets: the EuroGroups Register, the Statistical Business Register and the Central Balance Sheet Register.

2.1 Selection of CFIs

The selection of CFIs is based on vintage NACE codes from the EuroGroups Register and on current NACE codes from the Statistical Business Register. The paper pinpoints resident entities whose NACE codes fall under the categories [64.20, 64.305], in accordance with statistical standards. The NACE code 64.20 includes entities performing “activities of holding companies”. The NACE code 64.305 includes entities involved in the management of private wealth (“sociétés de gestion de patrimoine familial” (SPFs)).

The EuroGroups Register (EGR)\(^1\) provides annual data on the structure of multinational groups operating within the European Union (EU) member states and European Free Trade Association (EFTA) countries. To create the EGR frames, Eurostat collects data on enterprise groups from the national statistical business registers of EU and EFTA countries. The national statistical institutes provide micro data on the constituent units of the groups and on their relationships.\(^2\) In terms of data content, EGR notably provides vintage NACE codes concerning the economic activity undertaken by entities composing the structure of multinational groups.

The Statistical Business Register (Répertoire des entreprises, SBR) classifies resident companies according to their main economic activity by providing a current NACE code. The STATEC compiles this register and updates it on a monthly basis.

2.2 Accounting information on CFIs

The STATEC compiles and manages the accounting information of resident companies in the Central Balance Sheet Register (or centrale des bilans, henceforward CBSR). This register retrieves information from the annual accounts deposited by

---

\(^1\) Bikauskaite et al. (2019)

\(^2\) In Luxembourg, the STATEC performs this task. Eurostat is in charge of the centralization, compilation and management of the data.
resident companies on the electronic platform of the Luxembourg Business Register. These annual accounts comprise a balance sheet (*bilan comptable*), a profit and loss account (*compte de résultats*) and a standardised chart of accounts (*plan comptable*). While companies can deposit a complete or an abridged version of the balance sheet and of the profit and loss account, a complete version of the standardised chart of accounts is mandatory. As a result, the paper resorts to data from the standardised chart of accounts to build the prototype balance sheets of CFIs and identify a given type of CFI. Data are available over the period 2011-2019.

3. Prototype balance sheets of CFIs

Table 1 presents the potential types of CFIs listed in Di Filippo and Pierret (2020a) along with the structure of their respective prototype balance sheets.

Holding corporations can take the form of pure and mixed holdings. While the former confine their role to owning controlling-level amount of equity in one or more affiliates in a passive manner, without providing any other service to their affiliates, the latter can perform other ancillary activities.

Conduits raise or borrow funds from unrelated enterprises or the open market and remit those funds to their parent or to other affiliated enterprises. Conduits typically do not transact on the open markets on the assets side. A synonym for conduit is external financing.

Intragroup lending corporations perform lending from and to affiliated companies.

Captive factoring and invoicing corporations concentrate the accounts receivable (i.e. invoices or sales claims) of a group. According to ECB-Eurostat-OECD (2013), captive factoring should be classified under the NACE Rev. 2 section K6499 “Other financial service activities, except insurance and pension funding activities, n.e.c.”. The paper

Captive financial leasing corporations provide a loan agreement in which they, as lessor, purchase the assets on behalf of the lessee for economic use. Since captive financial leasing corporations feature a unique NACE code across the different types of CFIs, the paper identifies captive financial leasing corporations based on the NACE Rev. 2 classification code K6491, provided by STATEC.
Loan origination corporations finance companies external to the group to which they belong, based on funding obtained from the parent or from affiliated enterprises.

Securitisation vehicles carry out securitisation transactions. In Luxembourg, the Law of 22 March 2004 regulates securitisation vehicles. Given that the latter relate to sector S125 “Other financial intermediaries, except insurance corporations and pension funds”, they should not be part of the typology on CFIs.

Wealth-holding entities manage personal wealth for individuals or families by holding financial and non-financial assets. Since wealth-holding entities feature a unique NACE code across the different types of CFIs, the paper identifies wealth-holding entities based on the NACE Rev. 2 classification code K64.305, provided by the STATEC.

Predominant NFA corporations feature non-financial assets as the major item on the assets side of the balance sheet. No condition prevails on the liabilities side.

Extra-group loan origination corporations use loans obtained from third parties outside the group to finance specific assets, namely equity and debt securities (both as portfolio investment) and loans (as other investment).

Eventually, Di Filippo and Pierret (2020a) point to the relative importance of CFIs whose balance sheet features mixed activities in Luxembourg. This category labelled as mixed structures, does not cover borderline cases that are difficult to classify in the typology. Rather, it includes specific prototype balance sheets that are unique and distinct from the other types of CFIs. In particular, mixed structures bring together a mix of holding and intragroup lending corporations, companies declaring losses (negative capital) all over their living period and other mixed structures.
**Table 1: Prototype Balance Sheets of CFIs**

<table>
<thead>
<tr>
<th>Table 1.1: Holding corporations</th>
<th>Table 1.2: Conduit corporations</th>
<th>Table 1.3: Intragroup lending corporations</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Non-Financial Assets</strong></td>
<td><strong>Non-Financial Assets</strong></td>
<td><strong>Non-Financial Assets</strong></td>
</tr>
<tr>
<td>Direct investment Equity</td>
<td>Direct investment Equity</td>
<td>Direct investment Equity</td>
</tr>
<tr>
<td>Debt</td>
<td>Debt</td>
<td>Debt</td>
</tr>
<tr>
<td>Portfolio investment Equity</td>
<td>Portfolio investment Equity</td>
<td>Portfolio investment Equity</td>
</tr>
<tr>
<td>Debt</td>
<td>Debt</td>
<td>Debt</td>
</tr>
<tr>
<td>Other investment Loans</td>
<td>Other investment Loans</td>
<td>Other investment Loans</td>
</tr>
<tr>
<td>Currency &amp; Deposits</td>
<td>Currency &amp; Deposits</td>
<td>Currency &amp; Deposits</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th><strong>Table 1.4: Captive factoring and invoicing corporations</strong></th>
<th><strong>Table 1.5: Captive financial leasing corporations</strong></th>
<th><strong>Table 1.6: Loan origination corporations</strong></th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Non-Financial Asset</strong></td>
<td><strong>Non-Financial Asset</strong></td>
<td><strong>Non-Financial Asset</strong></td>
</tr>
<tr>
<td>Direct investment Equity</td>
<td>Direct investment Equity</td>
<td>Direct investment Equity</td>
</tr>
<tr>
<td>Yes/No</td>
<td>Yes/No</td>
<td>Yes/No</td>
</tr>
<tr>
<td>Portfolio investment Equity</td>
<td>Portfolio investment Equity</td>
<td>Portfolio investment Equity</td>
</tr>
<tr>
<td>Debt</td>
<td>Debt</td>
<td>Debt</td>
</tr>
<tr>
<td>Other investment Loans</td>
<td>Other investment Loans</td>
<td>Other investment Loans</td>
</tr>
<tr>
<td>Currency &amp; Deposits</td>
<td>Currency &amp; Deposits</td>
<td>Currency &amp; Deposits</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th><strong>Table 1.7: Securitisation vehicle corporations</strong></th>
<th><strong>Table 1.8: Personal and family wealth-holding entities</strong></th>
<th><strong>Table 1.9: Predominant non-financial assets</strong></th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Non-Financial Assets</strong></td>
<td><strong>Non-Financial Assets</strong></td>
<td><strong>Non-Financial Assets</strong></td>
</tr>
<tr>
<td>Direct investment Equity</td>
<td>Direct investment Equity</td>
<td>Direct investment Equity</td>
</tr>
<tr>
<td>Debt</td>
<td>Debt</td>
<td>Debt</td>
</tr>
<tr>
<td>Portfolio investment Equity</td>
<td>Portfolio investment Equity</td>
<td>Portfolio investment Equity</td>
</tr>
<tr>
<td>Debt</td>
<td>Debt</td>
<td>Debt</td>
</tr>
<tr>
<td>Other investment Loans</td>
<td>Other investment Loans</td>
<td>Other investment Loans</td>
</tr>
<tr>
<td>Currency &amp; Deposits</td>
<td>Currency &amp; Deposits</td>
<td>Currency &amp; Deposits</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th><strong>Table 1.10: Extra-group loan origination</strong></th>
<th><strong>This cell intentionally left blank</strong></th>
<th><strong>This cell intentionally left blank</strong></th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Non-Financial Assets</strong></td>
<td><strong>Non-Financial Assets</strong></td>
<td><strong>Non-Financial Assets</strong></td>
</tr>
<tr>
<td>Direct investment Equity</td>
<td>Direct investment Equity</td>
<td>Direct investment Equity</td>
</tr>
<tr>
<td>Debt</td>
<td>Debt</td>
<td>Debt</td>
</tr>
<tr>
<td>Portfolio investment Equity</td>
<td>Portfolio investment Equity</td>
<td>Portfolio investment Equity</td>
</tr>
<tr>
<td>Debt</td>
<td>Debt</td>
<td>Debt</td>
</tr>
<tr>
<td>Other investment Loans</td>
<td>Other investment Loans</td>
<td>Other investment Loans</td>
</tr>
<tr>
<td>Currency &amp; Deposits</td>
<td>Currency &amp; Deposits</td>
<td>Currency &amp; Deposits</td>
</tr>
</tbody>
</table>

Source: Di Filippo and Pierret (2020a), adapted from IMF (2018)
4. Qualitative approach

The identification of a given type of CFI relies on a qualitative approach. The latter analyses the relative predominance of a given balance sheet item over the others. This method considers three potential balance sheet layouts.

The first layout (Red layout) characterises a balance sheet where only one item predominates strongly over the others. The second layout (Red/Yellow layout) represents a balance sheet where one item (Red) predominates over the others but with a second item (Yellow) which features a relative importance compared to the remaining ones. The first item is thus larger than the second item. The third layout (Yellow/Yellow layout) features a balance sheet where no single item predominates over the others but where the sum of two items represents the majority of the balance sheet. The charts below present the three potential balance sheet layouts:

<table>
<thead>
<tr>
<th>Case 1: Red layout</th>
<th>Case 2: Red/Yellow layout</th>
<th>Case 3: Yellow/Yellow layout</th>
</tr>
</thead>
<tbody>
<tr>
<td>Assets / Liabilities</td>
<td>Assets / Liabilities</td>
<td>Assets / Liabilities</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Red</td>
<td>Yellow</td>
<td>Yellow</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

To distinguish between the three layouts, the paper applies qualitative criteria to the balance sheet of each CFI at each period. This approach features two steps.

A first step classifies the balance sheet items from the largest to the lowest in terms of proportion in the total balance sheet of a given CFI. Hence:

\[ S = \{s_1, s_2, \ldots, s_N\} \]

Where \( s_i > s_j \) for all \( i > j \), for \( i = 1, \ldots, N \) and \( j = 1, \ldots, N-1 \) and \( i \neq j \)

With \( s_i = \frac{\text{Balance sheet item}_i}{\text{Total Assets}} \) and \( s_j = \frac{\text{Balance sheet item}_j}{\text{Total Assets}} \)
A second step identifies the respective layouts based on three conditions. The existence criterion analyses whether a specific item features a positive amount and thus exists in the balance sheet. The predominance criterion tests whether a specific item predominates over the others in the balance sheet. The relative predominance criterion verifies whether any second item predominates over the remaining ones (excluding the first predominant item). To this end, this latter condition relies on an indicator of statistical dispersion applied to the distribution of the proportions of a company’s balance sheet items.

Hence, in the Red layout, the Red item $s_1$ must fulfil the following conditions:

$$\begin{align*}
\text{Red layout} & \quad \Leftrightarrow \\
\text{Existence criterion:} & \quad s_1 > 0 \\
\text{Predominance criterion:} & \quad s_1 > (1 - s_1) \\
\text{Strong predominance over the second largest item:} & \quad (s_1 - s_2) > \sigma^2/\mu
\end{align*}$$

In the Red/Yellow layout, the Red item $s_1$ and the Yellow item $s_2$ must respect the following conditions:

$$\begin{align*}
\text{Red / Yellow layout} & \quad \Leftrightarrow \\
\text{Existence criterion:} & \quad s_1 > 0 \text{ and } s_2 > 0 \\
\text{Predominance criterion:} & \quad s_1 > (1 - s_1) \\
\text{Weak predominance over the second largest item:} & \quad (s_1 - s_2) \leq \sigma^2/\mu
\end{align*}$$

In the Yellow/Yellow layout, the Yellow item $s_1$ and the Yellow item $s_2$ must fulfil the following conditions:

$$\begin{align*}
\text{Yellow / Yellow layout} & \quad \Leftrightarrow \\
\text{Existence criterion:} & \quad s_1 > 0 \text{ and } s_2 > 0 \\
\text{Non-predominance criterion:} & \quad s_1 \leq (1 - s_1) \\
\text{Non-predominance criterion:} & \quad s_2 \leq (1 - s_2)
\end{align*}$$

The paper thus identifies the balance sheet layout and links it to the associated prototype balance sheet (Table 1). This enables to draw a typology of CFIs. The advantage of the qualitative approach is to avoid the use of arbitrary quantitative thresholds when analysing the relative predominance of a given balance sheet item over the others.
5. Results from the new database

5.1 Typology of CFIs in Luxembourg

Chart 1.1 presents the typology of CFIs by total assets. On average over the period 2011-2019, the most important asset holders are holding corporations (52%), followed by intragroup lending companies (22%), mixed structures (15%), conduits (9%) and loan origination companies (1%). These corporations represent about 99% of the total assets held by CFIs. The remainder consists of companies with predominant non-financial assets, extra-group loan origination firms, wealth-holding entities, captive financial leasing corporations, and captive factoring and invoicing corporations.

Chart 1: Typology of CFIs with the new database

Chart 1.1 Total assets

Source: Authors' calculations. Units: EUR billion

Chart 1.2 Total number

Source: Authors' calculations. Units: Total number

Chart 1.2 presents the typology of CFIs by total number. On average over the period 2011-2019, CFIs regroup mixed structures (49%), holding corporations (20%), intragroup lending companies (13%), conduits (8%) and companies with predominant non-financial assets (5%). These corporations represent about 96% of the total number of CFIs. The remainder consists of loan origination companies, extra-group loan origination firms, wealth-holding entities, captive factoring and invoicing corporations, and captive financial leasing corporations.

Altogether, the relative proportions of the various types of CFIs differ depending on whether the typology is considered by total assets or by total number.
While holding corporations own the majority of total assets, the largest number of companies consists of mixed structures.

5.2 Typology of CFIs across balance sheet sizes

Typology of CFIs: a comparison between the new database and the BCL reporting framework

The paper compares the results of the typology between the new database and the BCL reporting framework (Di Filippo and Pierret (2020a)). As the BCL reporting framework considers only CFIs with total assets larger than EUR 500 million, the paper selects CFIs that suit this balance sheet size in the new database.

Charts 2 and 3 compare the results of the typology between the new database and the BCL reporting framework, by considering respectively the total number and the total assets of CFIs. While data are available in yearly frequency from 2011 to 2019 in the new database, the BCL reporting framework compiles data in monthly frequency from December 2014 to September 2021. For sake of comparison, the paper presents the typology based on the BCL reporting framework in annual frequency, over the period 2014 to 2019.
In terms of total number (Chart 2) and total assets (Chart 3), the typology of CFIs between the new database and the BCL reporting framework shares similar characteristics. Both charts point to the relative importance of holding corporations, intragroup lending companies, mixed structures and conduits.³

### Typology of CFIs across ranges of total assets

Altogether, for CFIs with total assets larger than EUR 500 million, the typology from the new database broadly shares similar characteristics with the typology of CFIs presented in Di Filippo and Pierret (2020a) based on the BCL reporting framework. Hence, a natural question that arises is in what range of total assets does the proportion of mixed structures increase?

To address this question, Charts 4 and 5 present the typology of CFIs across various ranges of balance sheet sizes, spanning [500;+∞[, [400;500[, [300;400[, [200;300[, [100;200] and [0;100].
The charts show that CFIs feature different types depending on their balance sheet size. On the one hand, CFIs with total assets larger than EUR 100 million mainly regroup holding companies, intragroup lending corporations, mixed structures and conduits. On the other hand, CFIs with total assets lower than EUR 100 million feature mostly mixed structures.
Mixed structures bring together a mix of holding and intragroup lending corporations, companies declaring losses (negative capital) all over their living period and other mixed structures. The prototype balance sheets of mixed structures are distinct from the prototype balance sheets of the other types of CFIs.

In terms of number and on average over the period 2011-2019, mixed structures bring together other mixed structures (49%), companies declaring losses (negative capital) all over their living period (44%) and a mix of holding and intragroup lending corporations (7%).

In terms of total assets and on average over the period 2011-2019, mixed structures represent a mix of holding and intragroup lending corporations (38%), other mixed structures (34%) and companies declaring losses (negative capital) all over their living period (28%).

A possible explanation relating to the use of mixed structures by MNEs is that they reduce costs and increase organisational efficiency. Indeed, mixed structures concentrate on different types of activities within a single structure, instead of resorting to multiple entities that perform a specific activity.4

6. Conclusion

The paper draws a typology of captive financial institutions and money lenders (CFIs, sector S127) in Luxembourg from a new database. The latter retrieves information from three sources: the EuroGroups Register managed by Eurostat, the Statistical Business Register managed by STATEC (the National Institute for Statistics and Economic Studies) and the Central Balance Sheet Register managed by STATEC. The new database enhances coverage of CFIs in Luxembourg. Indeed, it goes beyond the BCL reporting framework (BCL (2014)) by not only including CFIs with total assets larger or equal to EUR 500 million but also those with total assets of less than EUR 500 million. This larger coverage enables a more comprehensive analysis of CFIs in Luxembourg compared to previous studies that used the BCL reporting framework (Di Filippo and Pierret (2020a, 2020b)).

4 For example, a mixed structure could combine the activities of an intragroup lending corporation and a pure holding corporation.
Results show that CFIs present different characteristics depending on their balance sheet size. On the one hand, CFIs with total assets larger than EUR 100 million mainly regroup holding companies, intragroup lending corporations, mixed structures and conduits. On the other hand, CFIs with total assets lower than EUR 100 million feature mostly mixed structures. Overall, while holding corporations own the majority of total assets, the largest number of CFIs consists of mixed structures. Mixed structures bring together a mix of holding and intragroup lending corporations, companies declaring losses (negative capital) all over their living period and other mixed structures. The prototype balance sheets of mixed structures are distinct from the prototype balance sheets of the other types of CFIs.

References


https://www.czso.cz/documents/10180/88506450/32019719q1_069.pdf/22a9389c-27aa-4b23-b552-6fb75b9f7c6a?version=1.0


**ECB-Eurostat-OECD, 2013**, “Final Report by the Task Force on Head Offices, Holding Companies and Special Purpose Entities (SPEs)”, June 2013  

A Typology of Captive Financial Institutions in Luxembourg: Lessons from a New Database

Gabriele di Filippo & Frédéric Pierret
Statistics Department
Banque Centrale du Luxembourg

Disclaimer: This presentation should not be reported as representing the views of the BCL or the Eurosystem. The views expressed are those of the authors and may not be shared by other research staff or policy makers in the BCL or the Eurosystem.
1. Definition of Captive Financial Institutions (CFIs)

Captive financial institutions and money lenders (CFIs)
Classified in sector S127 (financial sector)

Can be found in MNEs’ structures and located
btw the headquarters and the operating affiliates

Meaning of term « captive »
« Captive » as owned and controlled by and typically for the sole use of an organisation: the parent

Main purpose
- Holding and finance activities
- e.g. holding of participations, treasury management, cash pooling, intragroup lending, etc.

Geographical location
- Financial centres
- access to various financing means, infrastructure, tax, regulatory and institutional advantages

Diagram 1: Example of a hypothetical complex structure of MNE
2. Importance of CFIs in Luxembourg

- CFIs are the major holder of inward and outward stocks of FDI in the international investment position of Luxembourg, compared to the other sectors.
- Their share represent 90% of the assets (outward FDI) and 88% of the liabilities (inward FDI) over the period Q4 2014 - Q4 2020.

Source: BCL. Unit: EUR billion. Foreign Direct Investment (FDI) stocks measure the total level of direct investment at the end of a quarter. The outward FDI stock is the value of the resident investors’ equity in and net loans to enterprises in foreign economies (hence residents’ assets). The inward FDI stock is the value of non-resident investors’ equity in and net loans to enterprises resident in the reporting economy (hence residents’ liabilities).
3. Typology of CFIs: Literature Review

- **Previous work:** Prototype balance sheets of CFIs by IMF (2018)
- **Contribution:** fine-tuned definitions of the potential types of CFIs
- **Data:** BCL reporting collects balance sheet data for resident CFIs with total assets ≥ EUR 500 million

Coverage rate of the population of CFIs in Di Filippo and Pierret (2020)
- **BCL reporting** covers a sub-sample of the whole population of CFIs
- Coverage rate in terms of **total assets:** 90% of the total assets held by the whole population of CFIs
- Coverage rate in terms of **number:** 5% of the whole population of CFIs

Main challenge
- How to improve the **coverage rate of CFIs** in Luxembourg?
4. New Database for CFIs in Luxembourg

STEP 1: Data providers gathering entities classified as CFIs
- EuroGroups Register (EGR) by Eurostat
- Statistical Business Register (SBR) by STATEC
  - CFIs with total assets ≥ EUR 500 million and ≤ EUR 500 million

STEP 2: Accounting data to build the prototype balance sheet of CFIs
- Central Balance Sheet Register (CBSR) by STATEC
  - Balance Sheet
  - Abridged Balance Sheet
  - Standardised Chart of Accounts
  - Resort to Chart of Accounts as available for more than 99% of CFIs

STEP 3: Build the prototype Balance Sheet of CFIs by matching the balance sheet items
- Matching btw Chart of Accounts items & Balance Sheet items in Central Balance Sheet Register
- Matching btw Balance Sheet items in Central Balance Sheet Register & BCL Balance Sheet items
- Matching btw BCL Balance Sheet items & Prototype Balance Sheet items (IMF (2018), BCL (2020))
  - Improve coverage of CFIs in LU by more than 900% in terms of number and by more than 5% in terms of total assets
  - Annual data over 2011-2019

STEP 4: Apply qualitative method as in Di Filippo and Pierret (2020)
- Screen the balance sheet of each CFI based on qualitative criteria (predominance of balance sheet items)
- Match with the pre-determined prototype balance sheet associated to a given type of CFI
- Advantage: avoid arbitrary quantitative thresholds
5. Typology of CFIs: New Database

Total assets - Main types of CFIs
- Holding corporations (52%)
- Intragroup lending companies (22%)
- Mixed structures (15%)
- Conduits (9%)
- Loan origination companies (1%)
⇒ These entities represent about 99% of the total assets held by CFIs

Total number - Main types of CFIs
- Mixed structures (49%)
- Holding corporations (20%)
- Intragroup lending companies (13%)
- Conduits (8%)
- Companies with predominant non-financial assets (5%)
⇒ These entities represent about 96% of the total assets held by CFIs

Proportions of the various types of CFIs differ when considering the typology by total assets or by total number
- Holding entities = largest holder of total assets
- Mixed structures = highest number of companies consists of mixed structures
  - In terms of number: other mixed structures (49%), companies declaring losses (negative capital) all over their living period (44%), mix of holding and intragroup lending corporations (7%);
  - In terms of total assets: mix of holding and intragroup lending corporations (38%), other mixed structures (34%), companies declaring losses (negative capital) all over their living period (28%)
6. Typology of CFIs: New Database versus BCL Reporting

Proportions of the various types of CFIs differ when considering the typology by total assets or by total number:

- Typology of CFIs btw new database and BCL reporting shares similar characteristics for CFIs with total assets ≥ EUR 500 million
- Relative importance of holding corporations, intragroup lending companies, mixed structures and conduits
- => In what range of total assets does the proportion of mixed structures increase?
- To address this question, the typology of CFIs must be split across various ranges of balance sheet sizes, spanning [500;+∞[, [400;500], [300;400], [200;300], [100;200] and [0;100]
7. Typology of CFIs: New Database
Decomposition by range of total assets

CFIs present different characteristics depending on their balance sheet size

- **CFIs with total assets ≥ EUR 100 million**: holding companies, intragroup lending corporations, mixed structures and conduits
- **CFIs with total assets < EUR 100 million**: mixed structures

=> The typology of CFIs differs, depending on the balance sheet size of CFIs

**Mixed structures**

- mix of holding and intragroup lending corporations, companies declaring losses (negative capital) all over their living period and other mixed structures

**Use of mixed structures by MNEs**

- Lower costs and higher organisational efficiency
- Mixed structures concentrate on different types of activities within a single structure, instead of resorting to multiple entities that perform a specific activity
Microdata base for sustainability indicators (ESG) developed at the Banco de España¹

Borja Fernández Rosillo San Isidro,
Bank of Spain

¹ This presentation was prepared for the conference. The views expressed are those of the author and do not necessarily reflect the views of the BIS, the IFC or the central banks and other institutions represented at the event.
Microdata base for sustainability indicators (ESG) developed at the Banco de España

Analysing climate change data gaps

Borja Fernández-Rosillo San Isidro

Abstract

This paper focuses on the various limitations encountered and achievements made in the process of developing a microdata base for sustainability indicators for non-financial companies. After carefully researching ESG standards, consulting ESG experts, analysing regulatory obligations in Europe and conducting practical research, a list of the 39 most relevant ESG indicators was selected from those normally reported by companies. We have currently gathered more than 15,000 indicators (2019-2020) using an internally developed semi-automatic search application. Numerous limitations were identified in the process, in terms of different metrics, lack of information and lack of digital support for downloading, comparability difficulties and regulatory restrictions.

Keywords: ESG (Environmental, Social, Governance), data gaps, standards, climate change, sustainability, databases, metrics, digitalised, regulation.

JEL classification: Q5 (environmental economics).
Index

Analysing climate change data gaps................................................................. 1

1. Introduction........................................................................................................ 3

2. Researching and establishing the relevant indicators .................................... 4

3. Extracting ESG information – transformation from unstructured to structured form ........................................................................................................ 5

4. Data gaps and limitations in current ESG reporting ........................................ 6

5. Improving data quality: quality control of sustainability indicators ............ 11

6. Conclusions......................................................................................................... 13

Next steps............................................................................................................... 14

References............................................................................................................. 15

Annex 1 – Selection of indicators ........................................................................ 15

Annex 2 – Dictionary of words associated with each indicator ......................... 16
**1. Introduction**

In recent years, awareness of social and environmental issues has been increasing, and consequently the demand for sustainability data has been growing exponentially. Although regulation and ESG reporting have been evolving and today there are numerous ESG requirements for companies, there is still a long way to go in terms of data availability, homogeneity, robustness and reliability. Despite its richness, the granular information available is still insufficient since:

- **It does not cover the entire population** of companies, mainly because (i) current regulations exclude small and medium-sized enterprises from obligatory reporting requirements and (ii) companies belonging to a group are not obliged to report ESG indicators if the parent company of the group reports this information.

- **It is not homogeneous** as there is no single definition of the indicators and metrics to be reported. Each company bases its reporting on a series of different standards with numerous ways of calculating and measuring the ESG indicators which makes comparison for the purposes of analysis difficult. Moreover, deriving from the first limitation, we have a mix of consolidated data (from companies which present their emissions for the entire group, i.e.) with individual data (from companies that directly report their emissions, i.e.). However, with the current regulatory advances, namely the Corporate Sustainability Reporting Directive (CSRD) and the European Sustainability Reporting Standards (ESRS) being developed in the European context by the European Financial Reporting Advisory Group (EFRAG) and the IFRS Sustainability Disclosure Standard being developed by the International Sustainability Standards Board (ISSB), as well as the Sustainability Accounting Standards Board (SASB) from the Value Reporting Foundation, the homogeneity limitation in relation to the indicators and metrics being reported is mostly being solved.

- **It is not digitalised**, which makes it difficult to download data for processing. The ESG information is still not reported in a digital reporting language such as XBRL (Extensible Business Reporting Language). In the near term it is expected to be digitalised like the financial statements.

Taking into account the limitations mentioned above and bearing in mind the need to assess the impact of economic policy measures on climate change, as well as the role of the financial system in channelling investments towards environmentally sustainable activities, the Banco de España is interested in using the information available in order to generate ESG statistics. Consequently, and in line with the institution’s objectives, the Statistics Department is working on the development of a microdata base for sustainability indicators for non-financial companies, the technical part of which has already been presented at the Irving Fisher Committee (IFC)¹ and, moreover, participating actively in three working groups relating to climate change².

---

¹ For further information regarding this web application prototype, see Koblents, Eugenia and Morales Alejandro, “Creation of a structured sustainability database from company reports: A web application prototype for information retrieval and storage”, presented at the IFC.

² Working groups: (1) Workstream on exposure of financial institutions to climate-related physical risks from the STC Expert Group on Climate Change and Statistics / (2) Workstream on climate-related data gaps from the FSC Task Force on Climate Impact Assessment Analytics / (3) Workstream on ESG non-financial information reporting from the XBRL Spanish association.
The objective of this paper is to present the process developed in the Central Balance Sheet Data Office for obtaining, processing and analysing the ESG indicators of non-financial corporations.

2. Researching and establishing the relevant indicators

Before determining what indicators to look for, a rigorous analysis of the different aspects that needed to be taken into consideration was conducted, as follows:

**Analysing the current regulatory obligations.** In this part it was detected that, under the Non-Financial Reporting Directive (Directive 2014/95/EU that was transposed into Spanish law by Ley 11/2018), companies were obliged to present a report containing non-financial information. However, the precise indicators and the format for reporting them were not specified. Additionally, they were not reported digitally and they did not follow a structured form.

**Researching the national and international ESG standards.** In order to better understand the current indicators reported by companies and the technical aspects of their preparation different ESG standards were analysed. We focused on the GRI (Global Reporting Initiative) technical papers since most Spanish companies report in line with this standard.

**Establishing a preliminary list of 120 indicators.** First, lists of indicators were consulted, such as the one compiled by the Spanish Association of Accounting and Business Administration (AECO) and the one compiled by the private company INFORMA. In conjunction with the indicators reflected in the GRI standards, these consultations enabled a preliminary list of 120 indicators to be drawn up from which to select the most feasible and interesting ones.

**Conducting a practical research exercise involving six listed companies.** In order to select the indicators with the highest probability of being found and of greatest interest a practical exercise was performed by searching for the 120 indicators on the preliminary list in the non-financial reports of six listed companies published in the year 2019. Additionally, a survey was conducted among three experts (one from the rating and risk sustainability field, another from the sustainability finance field and the last one from the accounting and ESG reporting field) who categorised each indicator into one of three categories (low, medium and high), according to its potential interest. The result of this exercise, taking into consideration various aspects of each indicator (difficulty, interest, feasibility, etc.), was a list of 39 ESG indicators (Annex 1) to search for. Figure 1 and Chart 1 below depict the distribution of the 39 selected indicators by type (environmental, social or governance) and subtype (energy, water, etc.):

![Figure 1. Distribution of the 39 selected indicators by type (ESG)](image-url)
Having established the indicators to look for it was necessary to address the process of retrieving them from non-financial reports. The main issue was the lack of homogeneity and digitalisation in the indicators reported. To solve this issue, and in close collaboration with data scientists of the Statistics Department, a semi-automatic search application was developed in order to obtain this data and transform it from an unstructured format to a structured database.\(^3\)

In order to facilitate the gathering of this information the search process was based on a set of pre-defined terms for each sustainability indicator. The dictionary of ESG terms associated with each indicator was prepared on the basis of the practical exercise mentioned before involving six listed companies and the information supplied by the indicator descriptions obtained from the various standards consulted. The result was a preliminary list of words (see Annex 2) which would help the web application locate this information.

The first list of words was compiled without any experience of this process of extracting ESG indicators from non-financial reports. Now, having gathered approximately 15,000 indicators, and having saved the search terms for each indicator, there is a record of the degree of success of each list of words that enables the current ontology to be improved in order to increase the success rate. Natural Language Processing (NLP) and Artificial Intelligence (AI) tools enable the application to learn from current searches and automatically improve the ontology in order to increase the degree of success.

\(^3\) For further information regarding this web application prototype, see Koblents, Eugenia and Alejandro Morales, “Creation of a structured sustainability database from company reports: A web application prototype for information retrieval and storage”, presented at the IFC.
4. Data gaps and limitations in current ESG reporting

From the experience of this project there is enough evidence to confirm that most companies in the sample analysed are aware of ESG risks and reporting. However, a wide variety of limitations were found during the process.

**Limitation 1 - Different metrics.** During the process of extracting ESG information we found a wide variety of metrics for some indicators which were, at first, a clear barrier to direct comparison. However, in most cases it was possible to perform a simple transformation of the indicator into the homogeneous metric defined. Some examples for specific indicators are shown below:

**Examples of “energy consumption within the organisation”**

<table>
<thead>
<tr>
<th>ENERGIA</th>
<th>2015</th>
<th>2016</th>
<th>2017</th>
<th>2018</th>
<th>2019</th>
</tr>
</thead>
<tbody>
<tr>
<td>Consumo total de Energía (MWh)</td>
<td>7,031,436</td>
<td>6,865,919</td>
<td>6,901,216</td>
<td>6,991,253</td>
<td>6,956,516</td>
</tr>
<tr>
<td>Electricidad (MWh)</td>
<td>6,612,778</td>
<td>6,391,248</td>
<td>6,461,695</td>
<td>6,543,895</td>
<td>6,574,002</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>ENERGY</th>
<th>2015</th>
<th>2016</th>
<th>2017</th>
<th>2018</th>
<th>2019</th>
</tr>
</thead>
<tbody>
<tr>
<td>Total energy consumption (MWh)</td>
<td>7,031,436</td>
<td>6,865,919</td>
<td>6,901,216</td>
<td>6,991,253</td>
<td>6,956,516</td>
</tr>
<tr>
<td>Electricity (MWh)</td>
<td>6,612,778</td>
<td>6,391,248</td>
<td>6,461,695</td>
<td>6,543,895</td>
<td>6,574,002</td>
</tr>
</tbody>
</table>

Table 1. This table shows “energy consumption within the organisation” indicator for telefonica (Spanish telecommunication company) in MWh

<table>
<thead>
<tr>
<th>CONSUMO ENERGÉTICO INTERNO POR FUENTE PRIMARIA (TJ)*</th>
<th>2017</th>
<th>2018</th>
<th>2019</th>
</tr>
</thead>
<tbody>
<tr>
<td>Carbón</td>
<td>244,784</td>
<td>221,079</td>
<td>81,927</td>
</tr>
<tr>
<td>Fuel-oleo</td>
<td>58,205</td>
<td>53,313</td>
<td>47,755</td>
</tr>
<tr>
<td>Gas natural</td>
<td>33,357</td>
<td>34,590</td>
<td>34,457</td>
</tr>
<tr>
<td>Total ENDESA consumption</td>
<td>684,142</td>
<td>615,336</td>
<td>507,614</td>
</tr>
</tbody>
</table>

Table 2. This table shows “energy consumption within the organisation” indicator for Endesa (Spanish energy company) in TJ
Table 3. This table shows “energy consumption within the organisation” indicator for Ferrovial (Spanish construction and transport company) in GJ

The variety of metrics used by companies to present this information is clear. However, in these specific cases, where there was a possibility of converting all the indicators to the same metric, a preliminary analysis was carried out to see which metric was the most common one (in the case of “energy consumption within the organisation”, MWH) and then all the data stored in the database were converted into that metric in order to facilitate comparison and economic analysis. Hence, the variety of metrics does not prevent analysis of this type of information, but it does make the process longer and more difficult. Hopefully, future regulatory standards will specify a common metric in order to minimise this limitation.

Examples of “GHG emissions intensity”

Table 4. This table shows the “GHG emissions intensity” indicator for Euskaltel (Spanish telecommunication company) in Kg CO2e / output

Chart 2. This table shows the “GHG emissions intensity” indicator for Inditex (Spanish retail company) in Kg CO2e / m2
In this second list of examples of regarding the “GHG emissions intensity” (Table 4 and 5 and chart 2) indicator there was a wide variety of ways in which companies presented this ratio as the current definition enables companies to use the business parameters they find most accurate (i.e. companies use in the denominator of the ratio total output, total sales, total employees, square metres of production plants, total km travelled, audio-visual production hours, etc.). Although some specific cases could be transformed (i.e. to use total employees or total sales, etc.), the numerator posed further difficulties as there was a variety of options to use (i.e. Scope 1, Scope 2 or Scope 1+2, among others). Consequently, in cases like this the data should only be taken into consideration for quality analysis, without attempting to make direct comparisons. In the case of this specific indicator it could be more beneficial to devise internally a ratio that can be calculated for all companies using the same parameters.

**Limitation 2 - Changes in data over time.** During the process it was not uncommon to see data change from one year to the next. There were a number of companies in the sample that changed data compilation criteria and, consequently, the number they had given for a specific indicator one year did not coincide with the number for the same indicator in the next year’s non-financial report. When an explanation was available in the non-financial report, they usually cited a change in the method of calculation. These changes are probably due to the novelty of this information and the scarcity of years of experience of collecting and preparing it. However, in these cases, the most up-to-date data were taken to be the relevant ones. Our intention is to closely follow the evolution of this information in order to see if it stabilises over time and if the number of changes falls significantly.
In the example reflected above, the company presents in its 2019 non-financial report 2018 data that have been recalculated using a different method. This leads to a substantial difference between years for the same indicator. During this initial period of adaptation to ESG reporting, which companies are currently in, such changes in criteria and data will be common. However, after a reasonable time of compiling and processing non-financial information, these changes in criteria and data can be expected to be fewer and of less impact.

**Limitation 3 - Lack of information.** Some indicators were not found for some companies in the search process. Moreover, it is important to highlight that searches for some indicators from the preliminary selection list had a higher success rate than for others.
The following conclusions drawn from the analysis presented above may be highlighted:

- Social indicators stand out as the type of indicator with the highest degree of success (80%) in their location, followed by governance (75%) and environmental (66%). This result was to be expected since social information (employees, diversity data, disability…) was being reported long before the current boom in environmental information.

- The overall success rate stands at 73%

- A substantial difference was observed between the degree of success in listed companies (81%) vs unlisted companies (50%). This is because listed companies have greater public exposure and stakeholder demand than unlisted companies.

**Limitation 4 – Comparability difficulties.** Although the main comparability issue is related to diverse metrics and can be relatively easily solved there are still some issues that generate difficulties in comparisons of this type of data such as:

- **Individual vs consolidated data.** The current regulations exempt individual companies from presenting the non-financial report when the company reports the information at a consolidated level. Consequently, we have the inconvenience of gathering data which in some cases are to be found at a consolidated level and in others (less often) at an individual level.

- **Global data from international companies.** Currently, many companies, especially the biggest, have an international presence all over the world. This affects the ESG information reported in their non-financial reports as it refers to the performance of the company or group worldwide, hindering the task of measuring national impacts (i.e. Scope 1 emissions in Spain). However, for a very small
percentage of companies and indicators, a breakdown is presented by country or economic region.

**Different calculation methodologies.** For some indicators and companies, the method of calculation varies so, although theoretically data could be comparable there is not enough evidence that the way they have obtained that data is through the same method or with similar criteria. A case in point are the Scope 2 emissions, which may be reported using a market approach or a location based approach. Consequently, if the company reports the method used there is no problem, but in the majority of the sample analysed the calculation method was not specified in the non-financial report.

**Limitation 5 - Lack of digitalisation of the information presented.** The information presented in the non-financial reports is still not digitalised and is reported in a wide variety of formats (i.e. charts, tables, etc.) which makes the process of locating it more difficult. However, the CSRD (Corporate Sustainability Reporting Directive) in progress includes the digitalisation and standardisation of reporting standards.

**Limitation 6 - Lack of official ESG verification.** Although there have been some advances in terms of ESG verification there is still no rigorous technical supervision of the information presented. Additionally, although the information presented could be checked, it is necessary to increase the amount of verification through the introduction of objective and technical measures that guarantee the quality and veracity of the data presented (i.e. that verify that the emissions a company reports are correctly calculated and real).

5. **Improving data quality: quality control of sustainability indicators**

The heterogeneity of the information and the novelty of having to analyse new types of information (e.g. greenhouse gas emissions and energy consumption) made it extremely important to conduct a rigorous and robust quality control assessment of the data. Consequently, each individual datum was reviewed, compared and contextualised, in terms of activity, in order to guarantee consistency in the database.

Moreover, during the process of analysis of each individual indicator the most common metric was identified in order to be able to define the standard metric in which to convert all the data to facilitate comparison. Additionally, and as a consequence of learning from this new information an interval analysis was conducted in order to establish the possible maximum and minimum values for each indicator. This exercise allowed the possible values of an indicator to be narrowed down and improved the process of searching for this information in future.

To increase the accuracy and quality of future uploadings of ESG data, individual guides were prepared for each of the indicators with the following fields:
Analyzing climate change data gaps

Figure 2. Example of the aspects covered in the first part (1-Information of interest on the indicator) of the guides prepared for each indicator

This first part (1-Information of interest on the indicator) of the guides relates to general information on the indicator in order to improve the process of locating the information. The specific aspects section contains information related to alternative ways of calculating the value or concrete aspects relevant to its interpretation. Additionally, the metrics sections help to establish if the data being uploaded is consistent in terms of value.

Figure 3. Example of the aspects covered in the second part (2-Analysis of the indicator by sector) of the guides prepared for each indicator

This second part (2-Analysis of the indicator by sector) of the guides relates to sample data for the indicator being analysed in order to provide a reference to establish if the new data being introduced into the database are consistent and realistic. This tables allows the possible intervals of the indicator to be narrowed down and gives the average values by sector and year in order to minimise the potential number of future errors that can be introduced into the database.
6. Conclusions

Sustainability reporting is becoming increasingly relevant to measure companies’ impact on climate change. The numerous demands from the economy’s stakeholders and the new green investment and financing decisions based on this information are key to the exponential growth in ESG reporting. However, despite current progress there is still a long way to go as there are still data gaps, heterogeneity in reporting, lack of digitalisation and verification, and regulatory limitations.

The experience obtained in gathering ESG indicators from Spanish non-financial corporations’ reports shows that listed companies present more and richer information than unlisted companies. Moreover, the ability to measure national ESG impacts is still a huge issue as multinational companies present information for the whole group and local companies either do not present ESG information, as they are exempted from doing so by the current regulations, or they belong to a group which presents consolidated data.

All in all, this project, in combination with the regulatory advances currently being made (CSRD) and the increasing awareness of reporting by companies, means that statisticians can be optimistic about ESG data availability in the near future.
Next steps

The achievements made and the experience obtained to date have provided the necessary catalyst to continue working on the provision of a robust ESG microdata base that can meet the various requirements of stakeholders and help minimise current ESG data gaps. Consequently, and bearing in mind the long-term goal of this project, the following lines of work are in the pipeline:

- Increasing the sample of companies from which information is currently extracted.
- Developing robust and solid quality control tests to minimise inconsistencies and errors in the data (automatic quality checks for future information uploads).
- Adapting the upcoming new ESG regulation to the microdata base.
- Analysing the possibility of increasing the number of indicators searched for.
- Refining the search process to increase the degree of success in terms of location and automation.
References


Annex 1 – Selection of indicators

Selection of indicators on which sustainability information has been collected (39)

<table>
<thead>
<tr>
<th>Type (E,S,G)</th>
<th>Subtype</th>
<th>Indicator</th>
</tr>
</thead>
<tbody>
<tr>
<td>E</td>
<td>Energy</td>
<td>Energy consumption within the organisation</td>
</tr>
<tr>
<td>E</td>
<td>Energy</td>
<td>Reduction of energy consumption</td>
</tr>
<tr>
<td>E</td>
<td>Energy</td>
<td>% of renewable energy among the total</td>
</tr>
<tr>
<td>E</td>
<td>Water</td>
<td>Water consumption</td>
</tr>
<tr>
<td>E</td>
<td>Water</td>
<td>Company located in a stress area regarding water</td>
</tr>
<tr>
<td>E</td>
<td>Greenhouse gases (GHG)</td>
<td>Direct GHG emissions (scope 1)</td>
</tr>
<tr>
<td>E</td>
<td>Greenhouse gases (GHG)</td>
<td>Indirect GHG emissions when generating energy (scope 2)</td>
</tr>
<tr>
<td>E</td>
<td>Greenhouse gases (GHG)</td>
<td>Other indirect GHG emissions (scope 3)</td>
</tr>
<tr>
<td>E</td>
<td>Greenhouse gases (GHG)</td>
<td>GHG emissions intensity</td>
</tr>
<tr>
<td>E</td>
<td>Greenhouse gases (GHG)</td>
<td>Reduction of GHG emissions</td>
</tr>
<tr>
<td>E</td>
<td>Waste</td>
<td>Non-hazardous waste</td>
</tr>
<tr>
<td>E</td>
<td>Waste</td>
<td>Waste generated</td>
</tr>
<tr>
<td>E</td>
<td>Waste</td>
<td>Waste not destined for disposal</td>
</tr>
<tr>
<td>E</td>
<td>Waste</td>
<td>Hazardous waste</td>
</tr>
<tr>
<td>E</td>
<td>Environmental policies</td>
<td>Environmental policy</td>
</tr>
<tr>
<td>E</td>
<td>Environmental policies</td>
<td>Circular economy</td>
</tr>
<tr>
<td>E</td>
<td>Environmental policies</td>
<td>ISO14001 regulatory compliance</td>
</tr>
<tr>
<td>S</td>
<td>Diversity, equality and well-being of staff</td>
<td>Employees</td>
</tr>
<tr>
<td>S</td>
<td>Diversity, equality and well-being of staff</td>
<td>Average age of employees</td>
</tr>
<tr>
<td>S</td>
<td>Diversity, equality and well-being of staff</td>
<td>Gender diversity</td>
</tr>
<tr>
<td>S</td>
<td>Diversity, equality and well-being of staff</td>
<td>Gender diversity on the board</td>
</tr>
<tr>
<td>S</td>
<td>Diversity, equality and well-being of staff</td>
<td>Average wage gap</td>
</tr>
<tr>
<td>S</td>
<td>Diversity, equality and well-being of staff</td>
<td>Work stability</td>
</tr>
<tr>
<td>S</td>
<td>Diversity, equality and well-being of staff</td>
<td>Disability</td>
</tr>
<tr>
<td>S</td>
<td>Diversity, equality and well-being of staff</td>
<td>Absenteeism</td>
</tr>
</tbody>
</table>
### Annex 2 – Dictionary of words associated with each indicator

<table>
<thead>
<tr>
<th>Type (E,S,G)</th>
<th>Subtype</th>
<th>Indicator</th>
<th>Dictionary of words (original queries in Spanish)</th>
<th>Dictionary of words (English)</th>
</tr>
</thead>
<tbody>
<tr>
<td>E</td>
<td>Energy consumption within the organisation</td>
<td>Consumo de energía eléctrico de la organización kWh, MWh, GWh, TWh, kJ, MJ, GJ, TJ, PJ “302-1” dato total vativos julios</td>
<td>Consumption of energy electrical energy electricity within the organization kWh, MWh, GWh, TWh, kJ, MJ, GJ, TJ, PJ “302-1” total data watts joules</td>
<td></td>
</tr>
<tr>
<td>E</td>
<td>Reduction of energy consumption</td>
<td>Reducción del consumo energético kWh, MWh, GWh, TWh, kJ, MJ, GJ, TJ, PJ “302-4” dato total evitar descender julios vativos</td>
<td>Reduction of energy consumption Energy kWh, MWh, GWh, TWh, kJ, MJ, GJ, TJ, PJ “302-4” total data avoid falling joules watts</td>
<td></td>
</tr>
<tr>
<td>E</td>
<td>% of renewable energy among the total</td>
<td>Porcentaje energía renovable total % ratio fuentes TJ consumo energético combustible fósil verde kWh GWh origen “menores emisiones” alternativa eólica solar 302 “no convencionales” kWh, MWh, GWh, TWh, kJ, MJ, GJ, TJ, PJ</td>
<td>Percentage of total renewable energy % ratio sources TJ energy consumption green fossil fuel kWh GWh origin “lower emissions” solar wind alternative 302 “non-conventional” kWh, MWh, GWh, TWh, kJ, MJ, GJ, TJ, PJ</td>
<td></td>
</tr>
<tr>
<td>E</td>
<td>Water consumption</td>
<td>“Consumo de agua” “uso de agua” “gasto de agua” “303-5” volumen caudal total dato m3 Hm3 megalitros litros millones toneladas masa captación dulce</td>
<td>“Water consumption” “water use” “water expenditure” “303-5” volume total flow data m3 Hm3 megalitres litres million tonnes fresh catchment mass</td>
<td></td>
</tr>
<tr>
<td>Type (E,S,G)</td>
<td>Indicator</td>
<td>Dictionary of words (original queries in Spanish)</td>
<td>Dictionary of words (English)</td>
<td></td>
</tr>
<tr>
<td>-------------</td>
<td>-----------</td>
<td>---------------------------------------------------</td>
<td>------------------------------</td>
<td></td>
</tr>
<tr>
<td>E</td>
<td>Company located in a water stress area</td>
<td>Estrés hídrico localización zona área agua escasez recursos hídricos H20 303</td>
<td>Water stress location area water scarcity water resources H20 303</td>
<td></td>
</tr>
<tr>
<td>E</td>
<td>Company located in a water stress area</td>
<td>Estrés hídrico localización zona área agua escasez recursos hídricos H20 303</td>
<td>Water stress location area water scarcity water resources H20 303</td>
<td></td>
</tr>
<tr>
<td>E</td>
<td>Direct GHG emissions (scope 1)</td>
<td>Emisiones directas de GEI &quot;alcance 1&quot; &quot;scope 1&quot; &quot;ambito 1&quot; &quot;Gases de Efecto Invernadero&quot; Mt Tn Tm CO2* &quot;305-1&quot; miles millones toneladas dato total kt calcul*</td>
<td>Direct GHG emissions &quot;scope 1&quot; &quot;scope 1&quot; &quot;Greenhouse Gases&quot; Mt Tn Tm CO2* &quot;305-1&quot; billion tonnes total data kt calcul*</td>
<td></td>
</tr>
<tr>
<td>E</td>
<td>Indirect GHG emissions when generating energy (scope 2)</td>
<td>Emisiones indirectas de GEI gener* energía energético &quot;alcance 2&quot; &quot;scope 2&quot; &quot;ambito 2&quot; &quot;Gases de Efecto Invernadero&quot; Mt Tn Tm CO2* &quot;305-2&quot; miles millones</td>
<td>Indirect GHG emissions generated* energy &quot;scope 2&quot; &quot;scope 2&quot; &quot;scope 2&quot; &quot;Greenhouse Gases&quot; Mt Tn Tm CO2* &quot;305-2&quot; billions</td>
<td></td>
</tr>
<tr>
<td>E</td>
<td>Other indirect GHG emissions (scope 3)</td>
<td>Otras emisiones indirectas de GEI gener* energía energético &quot;alcance 3&quot; &quot;scope 3&quot; &quot;ambito 3&quot; &quot;Gases de Efecto Invernadero&quot; Mt Tn Tm CO2* &quot;305-3&quot; miles millones toneladas eléctrico electricidad dato total kt calcul* TCO eq equivalente tCO* dióxido de carbono GHG</td>
<td>Other indirect GHG emissions generated* energy energy &quot;scope 3&quot; &quot;scope 3&quot; &quot;Greenhouse Gases&quot; Mt Tn Tm CO2* &quot;305-3&quot; billion tonnes electric electricity total data kt calcul* TCO eq equivalent tCO* carbon dioxide GHG</td>
<td></td>
</tr>
<tr>
<td>E</td>
<td>GHG emissions intensity</td>
<td>Intensidad de las emisiones de GEI &quot;Gases de Efecto Invernadero&quot; Ratio &quot;305-4&quot; dato tCO* CO2* kg kilogramo carbono TJ intensidad energética MtCO2e GHG</td>
<td>Intensity of GHG emissions &quot;Greenhouse Gases&quot; Ratio &quot;305-4&quot; data tCO* CO2* kg kilogram carbon TJ energy intensity MtCO2e GHG</td>
<td></td>
</tr>
<tr>
<td>E</td>
<td>Reduction of GHG emissions</td>
<td>Reducción de las emisiones de GEI objetivo evitar &quot;Gases de Efecto Invernadero&quot; Mt Tn CO2* &quot;305-5&quot; miles millones toneladas dato TCO* eq equivalente GHG</td>
<td>Reduction of GHG emissions objective to avoid &quot;Greenhouse Gases&quot; Mt Tn CO2* &quot;305-5&quot; billions of tonnes TCO data* GHG equivalent eq</td>
<td></td>
</tr>
<tr>
<td>E</td>
<td>Non-hazardous waste</td>
<td>Residuos generados &quot;no peligrosos&quot; reutilización reciclaje recuperación incineración compostaje toneladas Tn &quot;306-2&quot;</td>
<td>Waste generated &quot;non-hazardous&quot; reuse recycling recovery incineration composting tonnes Tn &quot;306-2&quot;</td>
<td></td>
</tr>
<tr>
<td>E</td>
<td>Waste generated</td>
<td>Residuos generados peso total toneladas Tn &quot;306-3&quot;</td>
<td>Waste generated total weight tonnes Tn &quot;306-3&quot;</td>
<td></td>
</tr>
<tr>
<td>E</td>
<td>Waste not destined for disposal</td>
<td>Residuos generados no destinados a eliminación Tn Toneladas composición &quot;306-4&quot;</td>
<td>Waste generated not intended for disposal Tn Tonnes composition &quot;306-4&quot;</td>
<td></td>
</tr>
<tr>
<td>E</td>
<td>Hazardous waste</td>
<td>Residuos generados peligrosos Tn Toneladas 306</td>
<td>Hazardous waste generated Tn Tonnes 306</td>
<td></td>
</tr>
<tr>
<td>E</td>
<td>Environmental Policy</td>
<td>Medio Ambiente Politica medioambiental Futuro limpio de emisiones Estrategia climática menor impacto productos bajos 30X</td>
<td>Environment Environmental policy Clean future of emissions Climate strategy less impact low products 30X</td>
<td></td>
</tr>
<tr>
<td>E</td>
<td>Circular economy</td>
<td>Proyectos de &quot;economia circular&quot; Reducción del uso de &quot;materias primas&quot; Ecología Circularidad Ecodiseño</td>
<td>&quot;Circular economy&quot; projects Reduction in the use of &quot;raw materials&quot; Ecology Circularity Ecodesign alternatives Efficiency in processes Innovation</td>
<td></td>
</tr>
<tr>
<td>Type (E,S,G)</td>
<td>Indicator</td>
<td>Dictionary of words (original queries in Spanish)</td>
<td>Dictionary of words (English)</td>
<td></td>
</tr>
<tr>
<td>-------------</td>
<td>----------</td>
<td>-------------------------------------------------</td>
<td>------------------------------</td>
<td></td>
</tr>
<tr>
<td>E</td>
<td>ISO14001 regulatory compliance</td>
<td>ISO 14001 cumplimiento normativa Estándar en gestión ambiental Responsabilidades medioambientales reducir impacto corporativa Requerimientos regulatorios</td>
<td>ISO 14001 regulatory compliance Environmental management standard Environmental responsibilities reduce corporate impact Regulatory requirements</td>
<td></td>
</tr>
<tr>
<td>S</td>
<td>Employees</td>
<td>Número de empleados Mujeres Hombres Distribución plantilla Contrato laboral Fijos Temporales &quot;102-8&quot;</td>
<td>Number of employees Women Men Workforce distribution Employment contract Permanent Temporary &quot;102-8&quot;</td>
<td></td>
</tr>
<tr>
<td>S</td>
<td>Average age of employees</td>
<td>edad media de la plantilla años empleo empleados</td>
<td>average age of the workforce years of employment employees</td>
<td></td>
</tr>
<tr>
<td>S</td>
<td>Gender diversity</td>
<td>Nº de hombre* mujer* total empleadas Empleo Distribución por sexo género igualdad laboral integración &quot;102-8&quot;</td>
<td>No. of men* women* total employees Employment Distribution by sex gender labour equality integration &quot;102-8&quot;</td>
<td></td>
</tr>
<tr>
<td>S</td>
<td>Gender diversity on the board</td>
<td>distribución por sexo &quot;102-8&quot; Consejo de administración mujer* Composición Compromiso igualdad Consejeros</td>
<td>distribution by sex &quot;102-8&quot; Board of Directors female* Composition Commitment to equality Directors</td>
<td></td>
</tr>
<tr>
<td>S</td>
<td>Average wage gap</td>
<td>Brecha salarial Sueldos por sexo igualdad de oportunidades Retribución Paridad Políticas de género a favor de hombre* mujer* equidad remuneración equitativa</td>
<td>Salary gap Salaries by sex Equal opportunities Remuneration Parity Gender policies in favour of men* women* equity equal pay</td>
<td></td>
</tr>
<tr>
<td>S</td>
<td>Work stability</td>
<td>Empleados por contrato laboral &quot;102-8&quot; Permanente Temporal Indefinidos Personal fijo tipo Tipología laborales</td>
<td>Employees by employment contract &quot;102-8&quot; Permanent Temporary Indefinite Permanent staff type Type of work</td>
<td></td>
</tr>
<tr>
<td>S</td>
<td>Disability</td>
<td>Discapacidad &quot;405-1&quot; Empleados Personal discapacitado grado Grupos vulnerables integración social inserción laboral igualdad de oportunidades accesibilidad exclusión diversidad discriminación respetuoso</td>
<td>Disability &quot;405-1&quot; Employees Disabled staff degree Vulnerable groups social integration job placement equal opportunities accessibility exclusion diversity discrimination respectful</td>
<td></td>
</tr>
<tr>
<td>S</td>
<td>Absenteeism</td>
<td>absentismo días puesto de trabajo ausencia ausentismo perdidos laborables totales horas jornadas</td>
<td>absenteeism days job absence absenteeism lost working days total hours working hours</td>
<td></td>
</tr>
<tr>
<td>S</td>
<td>Employee rotation</td>
<td>rotación de empleados abandono de puesto de trabajo fin relación laboral número personas abandonan &quot;404-1&quot; plantilla cese voluntario</td>
<td>employee turnover job abandonment termination of employment relationship number of people leaving &quot;404-1&quot; template voluntary resignation</td>
<td></td>
</tr>
<tr>
<td>S</td>
<td>Employee training</td>
<td>total horas de formación cursos empleado trabajador % año</td>
<td>total hours of training courses employee worker % year</td>
<td></td>
</tr>
<tr>
<td>S</td>
<td>Number of layoffs</td>
<td>Número de despidos empleados despedidos extinción rescisiones fin relación laboral cese indemnización terminación rescindir contrato trabajo</td>
<td>Number of layoffs dismissed employees termination terminations end of employment relationship cessation compensation termination terminate employment contract</td>
<td></td>
</tr>
<tr>
<td>S</td>
<td>Reconciliation measures</td>
<td>Conciliación bienestar empleados medidas de laboral políticas sociales reducir estrés emocional mental vida privada familiar equilibrio hijos mayores discapacitados compaginar conciliar</td>
<td>Reconciliation well-being employees labour measures social policies reduce mental emotional stress private family life balance older children disabled reconcile</td>
<td></td>
</tr>
<tr>
<td>Type (E,S,G)</td>
<td>Indicator</td>
<td>Dictionary of words (original queries in Spanish)</td>
<td>Dictionary of words (English)</td>
<td></td>
</tr>
<tr>
<td>-------------</td>
<td>-----------</td>
<td>---------------------------------------------------</td>
<td>-------------------------------</td>
<td></td>
</tr>
<tr>
<td>S</td>
<td>Human rights policy</td>
<td>Compromiso social político de derechos humanos</td>
<td>Social commitment human rights policy</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>412 vulneración abusos vulnerables</td>
<td>412 violation abuses vulnerable responsibility</td>
<td></td>
</tr>
<tr>
<td>S</td>
<td>Equality plan</td>
<td>Compromiso social plan político de igualdad oportunidades</td>
<td>Social commitment equal opportunities policy plan</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>&quot;102-8&quot; &quot;406-1&quot; &quot;404-2&quot;</td>
<td>&quot;102-8&quot; &quot;406-1&quot; &quot;404-2&quot;</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>inclusión discriminación raza sexo color</td>
<td>inclusion discrimination race sex colour</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>religión diversidad género mujer*</td>
<td>religion diversity gender woman*</td>
<td></td>
</tr>
<tr>
<td>S</td>
<td>Diversity plan</td>
<td>Políticas de inclusión</td>
<td>Inclusion policies</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>Igualdad de derechos LGBTI</td>
<td>Equal rights LGBTI</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>Plan de diversidad plantilla 405</td>
<td>LGBTI Diversity plan 405</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>origen étnico orientación identidad sexual sin prejuicios</td>
<td>template ethnic origin orientation sexual identity without prejudice</td>
<td></td>
</tr>
<tr>
<td>S</td>
<td>Health and safety policy</td>
<td>Salud y seguridad política de asistencia sanitaria</td>
<td>Health and safety health care policy</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>seguros médicos beneficios sociales planes de ayuda beneficios sociales 403 bienestar de los empleados</td>
<td>medical insurance social benefits assistance plans social benefits 403 employee welfare</td>
<td></td>
</tr>
<tr>
<td>S</td>
<td>Payment to suppliers</td>
<td>Período medio pago a proveedores días PMP cadena de suministro grupos de interés</td>
<td>Average payment period to suppliers days PMP supply chain stakeholders</td>
<td></td>
</tr>
<tr>
<td>S</td>
<td>Customer satisfaction level</td>
<td>Nivel de satisfacción cliente servicio postventa</td>
<td>Level of customer satisfaction after-sales service</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>trato experiencia compra usuario grado &quot;102-43&quot; GRI &quot;102-44&quot; insatisfacción grupos interés</td>
<td>experience grade &quot;102-43&quot; GRI &quot;102-44&quot; dissatisfaction with interest groups</td>
<td></td>
</tr>
<tr>
<td>G</td>
<td>Board average remuneration</td>
<td>Remuneración media percibida salario administración 405 miles anual euros variable especie monetaria fija a largo plazo</td>
<td>Average remuneration received salary of the board remuneration of the board of directors 405 thousand euros per year variable long-term fixed monetary species</td>
<td></td>
</tr>
<tr>
<td>G</td>
<td>Crime prevention policy</td>
<td>Modelo de Prevención de Delitos Certificado AENOR Sistema de gestión de compliance penal</td>
<td>Crime Prevention Model AENOR Certificate Anti-bribery criminal compliance management system</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>anticorrupción UNE-ISO 37001 UNE 19601 sooborno</td>
<td>Anti-corruption UNE-ISO 37001 UNE 19601 bribery</td>
<td></td>
</tr>
<tr>
<td>G</td>
<td>Number of corruption and bribery complaints</td>
<td>Casos de corrupción soborno políticas antisoborno nº &quot;205-3&quot; fraude prevenir luchar sanción indebido denuncia</td>
<td>Corruption cases bribery anti-bribery policies number &quot;205-3&quot; fraud prevent fight sanction improper complaint</td>
<td></td>
</tr>
<tr>
<td>G</td>
<td>Complaint channel</td>
<td>Canal de denuncias Ética y Cumplimiento sistemas de comunicación comunicar fraud* empleados transparencia ilegal irregular* conducta Anonimato Confidencialidad 205 de forma confidencial y anónima</td>
<td>Ethics and Compliance Whistleblowing Channel Communication Systems Communicate Fraud* Employees Transparency Illegal Irregular* Conduct Anonymity Confidentiality 205 Confidentially and Anonymously</td>
<td></td>
</tr>
</tbody>
</table>
MICRODATA BASE FOR SUSTAINABILITY INDICATORS (ESG) DEVELOPED AT THE BANCO DE ESPAÑA - ANALYSING CLIMATE CHANGE DATA GAPS

Borja Fernández-Rosillo San Isidro

Central Balance Sheet Data Office
Statistics Department

11\textsuperscript{th} IFC CONFERENCE ON “POST-PANDEMIC LANDSCAPE FOR CENTRAL BANKS STATISTICS”

BASEL
25/08/2022
INDEX

1. Introduction and context
2. Researching and establishing the relevant indicators
3. Extracting ESG information: from unstructured to structured form
4. Data gaps and limitations
5. Improving data quality
6. Conclusions
1. INTRODUCTION AND CONTEXT

Why does this need arise? What are the main challenges?

In recent years, awareness of social and environmental issues has been increasing, and consequently the demand for sustainability data has been growing exponentially.

1. Measure the exposure of the Spanish Economy to climate change at a disaggregated level
2. Analyse the implications of climate change and transition to a more sustainable economy
3. Evaluate the impact of economic policy measures on climate change
4. Facilitate the channelling of investment towards environmentally friendly activities

Despite its richness, the granular information available is still insufficient since:

1. It does not cover the entire population
2. It is not homogeneous
3. It is not digitalised

It is essential to increase the quantity, quality and harmonization of Environmental information.

NEED FOR A GRANULAR DATABASE ON ESG INFORMATION
1. INTRODUCTION AND CONTEXT
Most limiting aspects in current and future regulations

<table>
<thead>
<tr>
<th>MAIN CURRENT LIMITATIONS</th>
<th>Spanish Law 11/2018 on non-financial information and diversity</th>
<th>CSRD (Corporate Sustainability Reporting Directive) – in progress</th>
</tr>
</thead>
<tbody>
<tr>
<td>1 ONLY GROUPS? EXEMPTION FOR SUBSIDIARIES</td>
<td></td>
<td>EFRAG - ESRS (European Sustainability Reporting Standards)</td>
</tr>
<tr>
<td>2 HETEROGENEITY OF REPORTED INDICATORS</td>
<td></td>
<td>Information available in XBRL format</td>
</tr>
<tr>
<td>3 LACK OF DIGITALIZATION</td>
<td></td>
<td></td>
</tr>
<tr>
<td>4 LIMITED POPULATION (SMALL AND MEDIUM OUTSIDE)</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

TWO IMPORTANT LIMITATIONS REMAIN WITH THE NEW CSRD
Main aspects taken into consideration

1. Analysing the current regulatory obligations
   - Non-Financial Reporting Directive (Directive 2014/95/EU that was transposed into Spanish law by Ley 11/2018)
   - In the future CSRD will be transposed and will update Ley 11/2018

2. Researching the national and international ESG standards
   - Researching the national and international ESG standards
   - Establishing a preliminary list of 124 indicators

3. Establishing a preliminary list of 124 indicators
   - List of preliminary indicators
     TOTAL 124
     - E-Environmental: 62
     - S-Social: 44
     - G-Governance: 18

4. Conducting a practical research exercise involving six listed companies

5. List of 39 ESG indicators to search for
   - E 62
   - S 44
   - G 18
   - TOTAL 39

Distribution of indicators by subtype

Distribution of indicators by type (ESG)
2. RESEARCHING AND ESTABLISHING THE RELEVANT INDICATORS

List of indicators for searching

**TYPE E**

- Energy consumption within the org.
- Reduction of energy consumption
- % renewable energy among the total
- Water consumption
- Company located in a stress area regarding water
- Direct GHG emissions (scope 1)
- Indirect GHG emissions when generating energy (scope 2)
- Other indirect GHG emissions (scope 3)
- GHG emissions intensity
- Reduction of GHG emissions
- Environmental Policy
- Circular Economy
- ISO14001 regulatory compliance
- Non-hazardous waste
- Waste generated
- Waste not destined for disposal
- Hazardous waste

**TYPE S**

- Employees
- Average age of employees
- Gender diversity
- Gender diversity on the board
- Average wage gap
- Work stability
- Disability
- Absenteeism
- Employee rotation
- Employee training
- Number of layoffs
- Reconciliation measures
- Human Rights Policy
- Equality plan
- Diversity plan
- Health and safety policy
- Payment to suppliers
- Customer satisfaction level

**TYPE G**

- Board average remuneration
- Crime Prevention Policy
- Nº corruption and bribery complaints
- Complaint channel

**TOTAL:** 39
3. EXTRACTING ESG INFORMATION: FROM UNSTRUCTURED TO STRUCTURED FORM

Prototype developed with AI (Artificial Intelligence) support*

- Preliminary dictionary of ESG search terms: helps the semi-automatic search application locate this information.
- Context information (exact location and paragraph) is saved for each indicator.
- This labelled data enables to automatically optimize search terms using Machine Learning (ML) and Natural Language Processing (NLP). An strategy has been designed to propose, evaluate, compare and optimize queries.

(*) For further information regarding this web application prototype, see Koblents, Eugenia and Alejandro Morales, “Creation of a structured sustainability database from company reports: A web application prototype for information retrieval and storage”, presented at the IFC.
4. DATA GAPS AND LIMITATIONS

ESG data (1/2)

1. DIFFERENT METRICS
   - **Wide variety** of metrics for some indicators
   - Barrier to direct comparison
   - However, in most cases it was possible to perform a simple **transformation** to a homogeneous metric defined
   - Future regulatory standards will **specify a common metric** in order to minimise this limitation

2. CHANGES IN DATA OVER TIME
   - During the process it was not uncommon to see **data change** from one year to the next
   - These changes are probably due to the **novelty** of this information and the **scarcity of years of experience**
   - We will closely follow the evolution of this information in order to see if it stabilises over time

3. LACK OF INFORMATION
   - A substantial difference was observed between the degree of success in **listed companies** (81%) vs **unlisted companies** (50%).
   - The overall success rate stands at 73%
   - **Social indicators** stand out as the type of indicator with the **highest degree of success** (80%) in their location.

---

**Energy Consumption**

<table>
<thead>
<tr>
<th>Year</th>
<th>2015</th>
<th>2016</th>
<th>2017</th>
<th>2018</th>
<th>2019</th>
</tr>
</thead>
<tbody>
<tr>
<td>Total energy consumption (MWh)</td>
<td>7,081,458</td>
<td>6,865,919</td>
<td>6,901,216</td>
<td>6,991,281</td>
<td>6,958,516</td>
</tr>
<tr>
<td>Electricity (MWh)</td>
<td>6,012,778</td>
<td>6,391,248</td>
<td>6,501,295</td>
<td>6,540,895</td>
<td>6,574,002</td>
</tr>
</tbody>
</table>

**Internal Energy Consumption by Primary Source (TJ)**

<table>
<thead>
<tr>
<th>Fuel type</th>
<th>2017</th>
<th>2018</th>
<th>2019</th>
</tr>
</thead>
<tbody>
<tr>
<td>Total ENDES consumption</td>
<td>684,142</td>
<td>613,736</td>
<td>567,614</td>
</tr>
</tbody>
</table>

**Distribution of emissions according to scope**

<table>
<thead>
<tr>
<th>Scope</th>
<th>2018 (TJ)</th>
<th>2018 (comparable with 2019)</th>
<th>2019</th>
</tr>
</thead>
<tbody>
<tr>
<td>Scope 1</td>
<td>7,477</td>
<td>32,223</td>
<td>32,761</td>
</tr>
<tr>
<td>Scope 2</td>
<td>949</td>
<td>949</td>
<td>1,159</td>
</tr>
<tr>
<td>Scope 3</td>
<td>20,471</td>
<td>726</td>
<td>632</td>
</tr>
<tr>
<td>TOTAL</td>
<td>37,897</td>
<td>34,453</td>
<td></td>
</tr>
</tbody>
</table>

* Reported breakdown in 2019 non-financial statement

**Degree of success by type**

- **Environmental**
  - Found: 50%
  - Not found: 50%
- **Social**
  - Found: 20%
  - Not found: 80%
- **Governance**
  - Found: 80%
  - Not found: 20%

**Total degree of success**

- **Found**: 50%
- **Not found**: 50%
4. DATA GAPS AND LIMITATIONS

ESG data (2/2)

- Due to diverse metrics (easily solution in most cases) and different calculation methodologies
- Individual vs consolidated data. Current regulatory exemptions provokes some inconvenience
- Global data from international companies makes it difficult to measure national impacts

5. LACK OF DIGITALISATION OF THE INFORMATION PRESENTED

- Information presented in the non-financial reports is still not digitalised
- However, the CSRD (Corporate Sustainability Reporting Directive) includes the digitalisation and standardisation of reporting standards

6. LACK OF ESG VERIFICATION

- No rigorous technical supervision of the information presented
- Introduction of objective and technical measures that guarantee the quality and veracity of the data presented
The heterogeneity of the information and the novelty of having to analyse new types of information (e.g. greenhouse gas emissions and energy consumption) made it extremely important to conduct a rigorous and robust quality control assessment of the data → BANK OF SPAIN ESG GUIDES FOR INTERNAL USE*

This first part of the guides relates to general information on the indicator in order to improve the process of locating the information.

This second part of the guides relates to sample data for the indicator being analysed in order to provide a reference to establish if the new data being introduced into the database are consistent and realistic.

This third part of the guides relates to real examples of how this information can be presented in the non-financial reports in order to facilitate the location of this information.

(* To increase the accuracy and quality of future uploadings of ESG data, individual guides were prepared for each of the 39 indicators. These guides are only available for Bank of Spain technical employees involved in the process of gathering ESG data.)
6. CONCLUSIONS
Conclusions and challenges

✓ Real data shows the limitation of current ESG regulation (data gaps, heterogeneity and variability in reporting, comparability difficulties...)

✓ Information of groups is not enough. There is a need for information from individual companies.

✓ The relevance of having a homogeneous list of harmonised indicators with a clear criteria.

✓ The importance of information in electronic formats for its massive and automated treatment.
THANKS FOR YOUR ATTENTION
APPENDIX
## Selection of ESG indicators for the database – some examples

### DIRECTORATE GENERAL ECONOMICS, STATISTICS AND RESEARCH

<table>
<thead>
<tr>
<th>Indicator</th>
<th>Type</th>
<th>Subtype</th>
<th>Metric</th>
<th>Standards applied</th>
<th>Informa</th>
<th>Regulation</th>
<th>Search on non-financial statements</th>
<th>Statistic Department</th>
<th>Level of interest</th>
</tr>
</thead>
<tbody>
<tr>
<td>Energy consumption within the org.</td>
<td>E</td>
<td>Energy</td>
<td>KWh, MWh, GWh, TWh, kj, MJ, GJ, PJ</td>
<td>GRI 302-1 AECA</td>
<td>NO</td>
<td>YES</td>
<td>n.d.</td>
<td>1 - Currently collecting (*)</td>
<td>1 - High</td>
</tr>
<tr>
<td>Water consumption</td>
<td>E</td>
<td>Water</td>
<td>m3, Hm3, megalitros</td>
<td>GRI 303-5 AECA</td>
<td>YES</td>
<td>YES</td>
<td>n.d.</td>
<td>1 - Currently collecting (*)</td>
<td>1 - High</td>
</tr>
<tr>
<td>Direct GHG emissions (scope 1)</td>
<td>E</td>
<td>Green House Gases (GHG)</td>
<td>MtCO2e, TnCO2e</td>
<td>GRI 305-1 AECA</td>
<td>YES</td>
<td>YES</td>
<td>n.d.</td>
<td>1 - Currently collecting (*)</td>
<td>1 - High</td>
</tr>
<tr>
<td>Hazardous waste</td>
<td>E</td>
<td>Waste</td>
<td>Tn</td>
<td>Old GRI 306-2</td>
<td>YES</td>
<td>YES</td>
<td>n.d.</td>
<td>2 - Would like to collect</td>
<td>1 - High</td>
</tr>
<tr>
<td>% renewable energy among the total</td>
<td>E</td>
<td>Energy</td>
<td>%</td>
<td>GRI 302-1</td>
<td>NO</td>
<td>NFRD</td>
<td>YES</td>
<td>2 - Would like to collect</td>
<td>1 - High</td>
</tr>
<tr>
<td>Number of layoffs</td>
<td>S</td>
<td>Diversity, equality and well-being of staff</td>
<td>Nº</td>
<td>NO</td>
<td>YES</td>
<td>YES</td>
<td>n.d.</td>
<td>1 - Easy</td>
<td>2 - Would like to collect</td>
</tr>
<tr>
<td>Complaint channel</td>
<td>G</td>
<td>Corruption and bribery</td>
<td>Si/No/ND</td>
<td>NO</td>
<td>YES</td>
<td>YES</td>
<td>n.d.</td>
<td>1 - Easy</td>
<td>2 - Would like to collect</td>
</tr>
</tbody>
</table>

*(*) Information available since 2019 for 30 Spanish Groups listed*
A web application with (semi-automatic) full-text search and storage capabilities has been developed.

1. The user selects a year, company and indicator.

2. The tool searches the ontology in company documents and presents an ordered list of paragraphs in order of relevance.

3. The user validates the search results and stores the indicator value into the database.

4. Context information (user, data, page and paragraph, terms and search results…) is saved, which will serve to improve the automation in the search process.
Example of comparison and optimization of search queries:

- Original query: 'consumo de agua dulce salada subterránea m3' -> score = 50%
- New optimized query: 'consumo agua m3', -> score = 60%

Scores of original queries (blue), optimized queries (red) and achieved improvements (green).

Indicators with lower dispersion (more separable) achieve higher scores, as expected.
0. QUALITY CONTROL OF THE SUSTAINABILITY DATABASE

Relevance of the review process

1. Understanding the behaviour of ESG indicators:
   - Logical values (limit and comprehend the possible values of the indicators)
   - Analysing the metrics in which indicators are reported (KWH, MWH, GJ...)
   - Searching for alternative methods to obtain the data
   - Value changes between exercises

2. Improving in the automatic search process:
   - Suggestions of new words
   - Removal of old words that could distort the search
   - Sharing improvement in queries with Data Scientists

3. Increase quality of the database:
   - Error corrections
   - Delete inconsistent values
   - Review correct metric consignation
20 employees were working part-time during 2 months to make the first data ingestion:

- 10,000 records
- 39 indicators
- 164 groups
- 470 documents

Visualization with PowerBI
Constructing forward-looking climate-related physical risk indicators - The use of private owned data for forward-looking aggregates at sector-country level for official statistics

Maurice Fehr and Elena Triebskorn, Deutsche Bundesbank, and Jens Mehrhoff, IMF

---

1 This presentation was prepared for the conference. The views expressed are those of the authors and do not necessarily reflect the views of the BIS, the IFC or the central banks and other institutions represented at the event.
Constructing forward-looking climate-related physical risk indicators

The use of private owned data for forward-looking aggregates at sector-country level for official statistics

Maurice Fehr and Elena Triebskorn, Deutsche Bundesbank, DG Statistics*

Jens Mehrhoff, International Monetary Fund, Statistics Department*

Abstract

Timely and consistent climate-related data are essential for analyses of trends and shifting patterns in financial markets and the economy as a whole. The NGFS Bridging Data Gaps report states that “given the importance of forward-looking assessments of both physical and transition risks, the current reliance on mostly backward-looking data is unsatisfactory”.

With this paper, we aim at answering the question: How can we use existing climate data from private data providers to extract relevant forward-looking aggregates at a sector and country level? To do so, we use forward-looking data from a third-party data provider such as physical risk scores and the underlying financial risks.

One main issue we encountered is the limited coverage of company-level data. In addition, the variation between the different data providers is high, similar to other areas of sustainability data. Further, the hazards covered as well as their definitions are not consistent across data providers and therefore need to be considered when analyzing results. Finally, yet importantly, physical risk metrics should be comparable across years and scenarios and reflect financial damages.

Since results are likely to depend greatly on source data used, we focus on developing methods and concepts rather than on the actual results. These findings can contribute to planned work under the new G20 Data Gaps Initiative, specifically the recommendation on physical and transition risk indicators.

* The authors would like to thank their colleagues at the Bundesbank’s Research Data and Service Centre for valuable suggestions and feedback. All views expressed in this report are the personal views of the authors and do not necessarily reflect the views of the Deutsche Bundesbank or the Eurosystem, nor the views of the IMF, its Executive Board, or IMF management.
1. Motivation

Why climate-related data?

Timely and consistent climate-related data are essential for analyses of trends and shifting patterns in financial markets and the economy as a whole. The availability of rich and up-to-date data is key to good policymaking. Many central banks, including the Eurosystem, are addressing climate-related risks on their balance sheets and in their supervisory roles.

High quality data sets are essential for any decision maker. Governments need them to decide on the future path of carbon taxes or to benchmark carbon emission reductions. Investors need detailed information on the climate footprint of companies to weigh up their investment decisions. Scattered data sources and cumbersome access further complicate the situation. It is essential to improve the coverage and consistency of data collections. Looking at various data pools helps to unmask measurement uncertainties and differences in the coverage from third-party data providers.

Why forward-looking indicators?

Reports such as from the Financial Stability Board (FSB)\(^1\) and the Network on Greening the Financial System (NGFS)\(^2\) have all highlighted the need to accelerate progress in making climate-related data available. In these reports, stakeholders report the need for more forward-looking data (for example targets or emissions pathways) and granular data (for example geo-spatial data at entity and asset levels). As an outcome of its monetary policy strategy review, the European Central Bank (ECB)\(^3\) is developing new experimental indicators, covering relevant green financial instruments and the carbon footprint of financial institutions as well as their exposures to climate-related physical risks.

A survey of the Irving Fisher Committee on Central Banks Statistics (IFC)\(^4\) in 2021 finds that indicators needed to properly support progress assessment, on sustainable financial instruments as well as environmental indicators related to physical risk, emission trading and energy use pricing, are of key importance. However, as many indicators are backward-looking, it is useful to complement them with forward-looking data to track commitments towards a greener economy. Still, forward-looking metrics seem to be a newer area of analysis for many central banks, with their actual use remaining limited so far.

The NGFS Bridging Data Gaps Progress Report states that “given the importance of forward-looking assessments of both physical and transition risks, the current reliance on mostly backward-looking data is unsatisfactory. Stakeholders reported that they need to understand the point-in-time performance of an exposure against a transition pathway – hence the need for firms to disclose their transition plans – as

\(^{1}\) FSB, July 2021.

\(^{2}\) NGFS, May 2021.

\(^{3}\) ECB, July 2021(a); and ECB, July 2021(b).

\(^{4}\) IFC, December 2021.
well as the impact of adaptation and mitigation measures on the evolution of the risks.

Making use of available data

Several initiatives have sought to facilitate the availability of adequate data to support sustainable finance. While global initiatives on improving climate data are progressing, we need to leverage on readily available data sources and approaches in the short to medium-term. This is essential to fulfil the various pressing data needs in monetary policy, financial stability, and banking supervision. To this end, the IFC recommends for central banks to intensify the identification of sustainable finance data needs, cooperate with traditional and new stakeholders to close data gaps, and lead by example by improving the usage of the new data being collected. 5

With this paper, we aim at answering the question: How can we use existing climate data from private data providers to extract relevant forward-looking aggregates at a sector and country level? To do so, we use data from a third-party data provider on forward-looking physical risk metrics, including the underlying financial risks.

The forward-looking metrics refer to physical risk in 2050, depending on various Representative Concentration Pathways (RCPs) that represent different levels of global warming and are defined by the Intergovernmental Panel on Climate Change (IPCC). The RCPs can be loosely translated into “low” (RCP 2.6), “medium” (RCP 4.5), and “high” (RCP 8.5) levels of global warming. While the impact of global warming is visible in financial damages, it is less visible in overall risk scores. We will base aggregation and analysis on financial damages data from ISS ESG (or, short, ISS). 6

One main issue we encountered is the limited coverage of company-level data. In addition, the variation between the different data providers is high, similar to other areas of sustainability data. Further, the hazards covered as well as their definitions are not consistent across data providers and therefore need to be considered when analyzing results. Finally, yet importantly, physical risk metrics should be comparable across years and scenarios and reflect financial damages.

To account for these challenges, we will focus on sectors and regions with comparatively good coverage. Most promising in our sample are manufacturing sectors in the US, China, Japan, India, and the EU. Since results are likely to depend greatly on source data used, we focus on developing methods and concepts, rather than on the actual results. We hope these findings can contribute to planned work under the new G20 Data Gaps Initiative, specifically the recommendation on physical and transition risk indicators. 7

5 IFC, op. cit.
6 https://www.issgovernance.com/
7 FSB and IMF, October 2021.
2. Data, definitions, and methodology

Data and coverage

We performed our aggregation and analyses based on data from ISS, a commercial data provider, to which the Deutsche Bundesbank procured access. The data used here is part of the Climate Core Impact dataset and consists of multiple physical climate risk metrics at a company level. ISS also provides metadata on NACE\(^8\) sector classification and country of incorporation for the companies in the dataset.

Data coverage is a key issue in aggregating company-level data. The physical risk dataset covers around 23 000, mostly large public, companies, many of which in large economies. Almost half of those are manufacturing companies (NACE C) while other large sectors in the dataset are logistics (NACE G–I) as well as financial and insurance activities (NACE K). The best country coverage can be found for the US and Japan with around 2 200 companies each, and for China with almost 2 500 companies. For the combination of country and sector, the dataset contains data on more than 1 000 companies each for the US, Chinese, and Japanese manufacturing sectors. There is also rather good coverage for the euro area, the United Kingdom, Canada, and India.

The number of companies in many sector-country combinations would be too low to conduct meaningful aggregation. Therefore, we look at the largest economies with the best ISS data coverage as well as the euro area separately while building regional aggregates for the other countries (see table 1).

<table>
<thead>
<tr>
<th>Table 1. Number of companies by economy and sector</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>United States</strong></td>
</tr>
<tr>
<td>8</td>
</tr>
<tr>
<td><strong>Euro Area</strong></td>
</tr>
<tr>
<td>20</td>
</tr>
<tr>
<td><strong>Japan</strong></td>
</tr>
<tr>
<td>10</td>
</tr>
<tr>
<td><strong>United Kingdom</strong></td>
</tr>
<tr>
<td>4</td>
</tr>
<tr>
<td><strong>Canada</strong></td>
</tr>
<tr>
<td>11</td>
</tr>
<tr>
<td><strong>Other AEs</strong></td>
</tr>
<tr>
<td>35</td>
</tr>
<tr>
<td><strong>China</strong></td>
</tr>
<tr>
<td>40</td>
</tr>
<tr>
<td><strong>India</strong></td>
</tr>
<tr>
<td>18</td>
</tr>
<tr>
<td><strong>Other EMDEs Asia</strong></td>
</tr>
<tr>
<td>63</td>
</tr>
<tr>
<td><strong>EMDEs Europe</strong></td>
</tr>
<tr>
<td>4</td>
</tr>
<tr>
<td><strong>EMDEs Latin America and the Caribbean</strong></td>
</tr>
<tr>
<td>20</td>
</tr>
<tr>
<td><strong>EMDEs Middle East and Central Asia</strong></td>
</tr>
<tr>
<td>11</td>
</tr>
<tr>
<td><strong>EMDEs Sub-Saharan Africa</strong></td>
</tr>
<tr>
<td>11</td>
</tr>
<tr>
<td><strong>Rest of the World</strong></td>
</tr>
<tr>
<td>13</td>
</tr>
</tbody>
</table>

Source: ISS ESG.
Note: AE = advanced economy; EMDE = emerging market and developing economy.

---

\(^8\) European classification of economic activities, from the French title: *Nomenclature générale des Activités économiques dans les Communautés Européennes*. NACE is derived from ISIC, the International Standard Industrial Classification of All Economic Activities.
Methodology

The ISS physical risk metrics show climate-related physical risk both for the base year 2020 and forward-looking for the year 2050. The forward-looking assessment is based on two scenarios: The “worst-case” scenario corresponds to the RCP 8.5 scenario of the IPCC Fifth Assessment Report, while the more moderate “likely” scenario corresponds to the RCP 4.5 scenario of said report.

For each of the scenarios and for the base year, absolute financial damages in 2020 US dollars are estimated for each of the hazards. The relative damages, i.e., damages as a percentage of the base year’s revenues, are also made available by ISS for each company. The difference between the future relative damage and the base year relative damage forms the basis for the risk score. This score is bound to a range between 1 and 100 and measures the change in relative financial risk of each company compared to the base year and to their respective GICS sector. It is normalized in a way that a score of 50 corresponds to the median risk of the issuer’s sector in the respective scenario. It is therefore neither comparable across sectors nor across scenarios. This is shown in figure 1 for the full sample of companies.

While the expected financial damage is generally higher for the RCP 8.5 scenario than for the RCP 4.5 scenario, this does not translate into higher physical risk scores in RCP 8.5, due to the normalization process. This limits the usefulness of the risk score for our purposes, and we thus focus on the actual financial risks from here on.

Figure 1. Distribution of overall physical risk by climate scenario

Source: ISS ESG.
ISS conducts the analysis of physical climate risk by distinguishing operational risks and market risks.

- Operational risks materialize at the level of the individual physical asset, either in the form of direct damage through business interruptions and repair cost or through reduced labor productivity. They are estimated by linking asset locations (if available) to multiple climate hazard maps depending on the type of hazard. Where no individual asset locations are available (as most companies report them only at the country level), country level information on the facilities is combined with exposure maps derived from satellite imagery and population data.

- Market risks materialize at the level of national GDP. It therefore depends on in which countries the company generate its revenue and the physical climate risks of those countries.

Six types of climate-related hazards are considered: coastal floods, droughts, heat stress, river floods, tropical cyclones, and wildfires. The last three are considered to cause only operational risks through direct damage, while coastal floods impact operational risks through the same channel and at the same time impact market risks through GDP. Heat stress decreases labor productivity which affects both operational risks and market risks. Droughts exclusively affect market risk through its impact on agriculture and therefore GDP.

Comparison ISS – Carbon4 Finance

Besides ISS, the Bundesbank has access to physical risk scores from another data provider: Carbon4 Finance (C4F)\(^\text{(12)}\). This data provider covers physical risk scores but not financial metrics. Furthermore, the number of companies that are covered by the dataset amounts to only about a quarter of that of ISS. Given how critical an issue the coverage of the dataset is for aggregation, we continue to use ISS data for the analysis. However, an important issue becomes visible when we compare the physical risk metrics of ISS of C4F for all the companies in the common sample.

To this end, we compare forward-looking physical risk metrics for the same year (2050) and for the same two scenarios (RCP 4.5 and RCP 8.5). The C4F physical risk score is not normalized for each scenario, so it is comparable to the financial physical risk metric from ISS, expected annual losses from climate-related damage in percent of revenues. The comparison shows some correlation but first and foremost high variation (see figure 2). If there was agreement on how exposed companies are to physical risk, the points would all be arranged on a straight line. This variation between data from different providers is a common theme in sustainability data.\(^\text{(13)}\) In the specific case of physical risk data, this can be explained, among other things, by the aforementioned fact that different hazards are covered and their definitions are not consistent across data providers.

\text{\textsuperscript{12}} https://www.carbon4finance.com/

\text{\textsuperscript{13}} Berg et al., 2019.
3. Preliminary indicators and main findings

Aggregation

We construct aggregate indicators by economic sector and by country or region using the metadata provided by ISS. In sectoral terms, we use the NACE framework. For the country and region breakdown we use the IMF classification of “advanced economies” (AE) and “emerging market and developing economies” (EMDE) as well as the latter’s regional breakdown.14

A metric used by ISS to show the physical climate risk of a company is the expected annual loss from climate-related damages in percent of the company’s baseline revenue. To derive sectoral and regional aggregate indicators, we calculate the weighted average of this metric in the respective country or region and sector, using baseline revenue as the weight. For each sector-region combination, we derive two aggregate indicators, one for the “likely” (RCP 4.5) and one for the “worst-case” (RCP 8.5) scenario.

Results

1. Results by country

In terms of how damages are distributed between companies located in different countries, we arrive at similar results as many other publications on the topic such as the ND-GAIN indicator on climate change vulnerability and readiness.15 We find that companies in emerging market and developing economies generally face higher physical risk than those located in advanced economies (see table 2). Within the EMDE group, India and other Asian countries stand out as being subject to particularly high

---

14 IMF, October 2022.
15 University of Notre Dame, 2022; and NGFS, October 2022.
Constructing forward-looking climate-related physical risk indicators

physical risk. Although this is not too unexpected, the gap to e.g., Middle Eastern and Sub-Saharan countries seem a bit large given that those regions are expected to suffer immensely from drought and heat stress respectively.

A possible explanation is that some climate hazards are better researched than others, e.g., floods vs. droughts. It is also important to keep in mind that the physical risk of companies is not necessarily representative of the physical risk of countries in which they are based for, at least, two reasons. First, (large) companies usually have production facilities outside their country of incorporation. Second, physical risk varies strongly within a country and a company’s facilities might be located in particularly favorable or unfavorable areas. Not surprisingly, the physical risk is consistently higher for the RCP 8.5 scenario than for the RCP 4.5 scenario.

Table 2. Total annual physical risk as a percentage of revenues by scenario

<table>
<thead>
<tr>
<th>Country/Region</th>
<th>RCP 4.5</th>
<th>RCP 8.5</th>
</tr>
</thead>
<tbody>
<tr>
<td>United States</td>
<td>0.14</td>
<td>0.18</td>
</tr>
<tr>
<td>Euro Area</td>
<td>0.09</td>
<td>0.12</td>
</tr>
<tr>
<td>Japan</td>
<td>0.17</td>
<td>0.25</td>
</tr>
<tr>
<td>United Kingdom</td>
<td>0.14</td>
<td>0.19</td>
</tr>
<tr>
<td>Canada</td>
<td>0.14</td>
<td>0.17</td>
</tr>
<tr>
<td>Other AEs</td>
<td>0.28</td>
<td>0.39</td>
</tr>
<tr>
<td>China</td>
<td>0.00</td>
<td>0.00</td>
</tr>
<tr>
<td>India</td>
<td>0.43</td>
<td>0.59</td>
</tr>
<tr>
<td>Other EMDEs Asia</td>
<td>0.90</td>
<td>1.23</td>
</tr>
<tr>
<td>EMDEs Europe</td>
<td>1.23</td>
<td>1.73</td>
</tr>
<tr>
<td>EMDEs Latin America and the Caribbean</td>
<td>0.08</td>
<td>0.10</td>
</tr>
<tr>
<td>EMDEs Middle East and Central Asia</td>
<td>0.31</td>
<td>0.48</td>
</tr>
<tr>
<td>EMDEs Sub-Saharan Africa</td>
<td>0.53</td>
<td>0.75</td>
</tr>
<tr>
<td>Rest of the World</td>
<td>0.51</td>
<td>0.72</td>
</tr>
</tbody>
</table>

Source: ISS ESG.
Note: AE = advanced economy; EMDE = emerging market and developing economy.

2. Results by sector

The sectors most affected by physical climate risk according to the ISS data are NACE A (Agriculture, forestry and fishing), B, D and E (Mining and quarrying and other industry) and L (Real estate activities). While the NACE sectors A, B and E directly deal with natural resources, the sectors D and L have a large dependency on physical assets such as electricity grids, pipelines, and buildings. On the other end, many services sectors face relatively low physical risk, while sector C (manufacturing), which is the largest sector in terms of ISS data coverage, is roughly in the middle. Table 3 shows the aggregation of ISS data by the NACE sector for the two scenarios.

Table 3. Global annual physical risk as a percentage of revenues by sector

<table>
<thead>
<tr>
<th>Sector</th>
<th>RCP 4.5</th>
<th>RCP 8.5</th>
</tr>
</thead>
<tbody>
<tr>
<td>A – Agriculture, forestry and fishing</td>
<td>0.49</td>
<td>0.70</td>
</tr>
<tr>
<td>B, D, E – Mining and quarrying and other industry</td>
<td>0.47</td>
<td>0.62</td>
</tr>
<tr>
<td>C – Manufacturing</td>
<td>0.24</td>
<td>0.33</td>
</tr>
</tbody>
</table>
3. Results by country group and sector

When looking at physical risk by sector for each of the country groups separately, we see that the overarching theme of emerging market and developing economies being more affected by physical climate risk than advanced economies is reflected in each of the sectors (see table 4). However, the large difference we see in physical risk between sectors with more physical assets (A, B, D, E, L) and e.g., the services sectors, can be attributed mostly to the advanced economies. In developing economies, the different sectors are more similarly affected by climate change damages and the ranking order of sectors changes. For example, in EMDEs, the R–S sectors are most affected by physical risk, while showing relatively low risk in AEs. However, the data coverage for these sectors is generally low (see table 1) so the results might not be reliable. On the other hand, the real estate sector (L) does not stand out in EMDEs as in AEs. The latter results are somewhat surprising and warrant further research. It should also be kept in mind, that the data coverage for EMDEs is generally lower than for AEs and the amount of information the data provider has on the company’s facilities might vary between the country groups.

Table 4. Total annual physical risk as a percentage of revenues by country group

<table>
<thead>
<tr>
<th>Advanced Economies</th>
<th>Emerging market and developing economies</th>
</tr>
</thead>
<tbody>
<tr>
<td>RCP 4.5</td>
<td>RCP 8.5</td>
</tr>
<tr>
<td>A</td>
<td>0.37</td>
</tr>
<tr>
<td>B, D, E</td>
<td>0.31</td>
</tr>
<tr>
<td>C</td>
<td>0.17</td>
</tr>
<tr>
<td>F</td>
<td>0.08</td>
</tr>
<tr>
<td>G–I</td>
<td>0.10</td>
</tr>
<tr>
<td>J</td>
<td>0.12</td>
</tr>
<tr>
<td>K</td>
<td>0.12</td>
</tr>
<tr>
<td>L</td>
<td>0.35</td>
</tr>
<tr>
<td>M–N</td>
<td>0.11</td>
</tr>
<tr>
<td>O–Q</td>
<td>0.11</td>
</tr>
<tr>
<td>R–S</td>
<td>0.15</td>
</tr>
</tbody>
</table>

Source: ISS ESG.
Challenges

Our findings point to four main challenges when constructing forward-looking physical risk indicators (see figure 3):

1. Forward-looking indicators on physical climate risk are not standardized. They can differ in terms of the year they focus on, the scenarios they consider, the hazards they cover, the way they define the hazards, the methodology for risk calculation, and the way the results are presented (scores or expected annual losses, for example).

2. The most common form of forward-looking physical risk indicators is some sort of score. Besides not having a direct financial interpretation, scores might not be comparable across years, sectors, or climate scenarios. This may not be a big problem when comparing individual companies’ physical risk exposure but should be considered carefully when creating aggregate indicators.

3. While our aggregate indicators show a distribution of physical climate risks between regions and sectors that is roughly as expected, the overall magnitude of physical risk is significantly different from what for example the NGFS finds for comparable scenarios.16 It requires further research to see where these differences come from.

4. The coverage for company-level physical risk data is relatively low. This is a major challenge when calculating sector and country aggregates. Furthermore, the proprietary methodology for calculating the risk metrics is not completely transparent.

Figure 3. Challenges in constructing forward-looking indicators

16 NGFS, September 2022.
4. Way forward

This paper aims at calculating forward-looking physical risk indicators using third-party data. Results show that source data plays an important role when comparing results. Going forward, we want to contrast the findings using third-party data with other forward-looking indicators.

One important source are the climate scenarios of the NGFS, which are now in Phase III of their development pipeline. As such, the latest release reflects the new country-level commitments made at COP26 in November 2021. The scenarios also include latest trends in renewable energy and key mitigation technologies. Notably, the modelling of physical risks has been improved; including a first estimate of acute physical risks. Moreover, transition risks are represented with increased granularity in certain sectors. Robustness checks with those results can give additional insights on the importance of source data and methodologies applied.

In addition, aspects of the work carried out here could also be applied to forward-looking transition risk indicators. As the Bundesbank has this data available from other third-party data providers, we want to look at how to calculate sector and country aggregates from the data available to us.

While this paper highlights some of the challenges of working with forward-looking physical risk data, it is important to gain experience with this data and learn how to work with imperfect climate data for the time being. In addition to the forward-looking nature, the data used here were highly granular, at company level with economic sector information. The natural next step to support policy analysis would be to link forward-looking geo-spatial data to the granular and sectoral dimensions.

---

17 However, the current data do not yet account for Russia’s invasion of Ukraine.
References

http://dx.doi.org/10.2139/ssrn.3438533

European Central Bank, July 2021(a), An overview of the ECB’s monetary policy strategy

European Central Bank, July 2021(b), Detailed roadmap of climate change-related actions

Financial Stability Board, July 2021, The availability of data with which to monitor and assess climate-related risks to financial stability

Financial Stability Board and International Monetary Fund, October 2021, G20 Data Gaps Initiative (DGI-2): The Sixth Progress Report—Countdown to December 2021

Irving Fisher Committee on Central Bank Statistics, December 2021, IFC Report No 14—Sustainable finance data for central banks

Intergovernmental Panel on Climate Change, 2014, Climate Change 2014 Synthesis Report

International Monetary Fund, October 2022, World Economic Outlook—Countering the Cost-of-Living Crisis

ISS ESG, April 2021, Climate Physical Risk Methodology
(The document is not publicly available.)

Network for Greening the Financial System, May 2021, Progress report on bridging data gaps

Network for Greening the Financial System, September 2022, NGFS Scenarios for central banks and supervisors

12 Constructing forward-looking climate-related physical risk indicators
Network for Greening the Financial System, October 2022, *Dashboard on scaling up green finance*

University of Notre Dame, 2022, *Notre Dame Global Adaptation Index (ND-GAIN)*
https://gain.nd.edu/our-work/country-index/
Constructing forward-looking climate-related data

Maurice Fehr and Elena Triebskorn, Sustainable Finance Data Hub, Directorate General Statistics

The views expressed are those of the authors and do not necessarily reflect those of the Bundesbank.
Introduction
Constructing forward-looking climate data

1. Why forward-looking data? – Excerpts and results from relevant publications

2. Results from work on the constructing of forward-looking climate-related physical risk indicators

3. Challenges and way forward
1. Why forward-looking data?  
Excerpts and results from relevant publications

The Network on Greening the Financial System (NGFS) published a progress report on its Bridging Data Gaps Workstream in May 2021:

- Persistent gaps in climate-related data hinder the achievement of these objectives. Stakeholders report the need for more forward-looking data (for example targets or emissions pathways) and granular data (for example geographical data at entity and asset-levels).
- Given the importance of forward-looking assessments of both physical and transition risks, the current reliance on mostly backward-looking data is unsatisfactory. Stakeholders reported that they need to understand the point-in-time performance of an exposure against a transition pathway – hence the need for firms to disclose their transition plans – as well as the impact of adaptation and mitigation measures on the evolution of the risks.

The Irving Fisher Committee on Central Bank Statistics (IFC) released a report on Sustainable finance data for Central Banks in December 2021:

- Many indicators are backward-looking, it is useful to complement them with forward-looking data to track commitments towards a greener economy.
- In general, forward-looking metrics seem to be a newer area of analysis for many central banks, with their actual use remaining limited so far.
2. Constructing forward-looking climate-related physical risk indicators
   Underlying data and research question

− Bundesbank acquired a variety of climate-related indicators. From the same two data providers, we have physical risk data at the company level available for internal analysis.

− In one case, the data set consists only of physical risk scores; in the other case, it also includes the underlying financial risks.

− Both data sets are forward-looking and use the IPCC’s Representative Concentration Pathways (RCPs) that represent different levels of global warming. They can be loosely translated into: low (RCP 2.6), medium (RCP 4.5) and high (RCP 8.5) levels of global warming.

− Bundesbank Sustainable Finance Data Hub (Maurice Fehr and Elena Triebskorn) together with Jens Mehrhoff (IMF) are exploring the question:

   How can we use existing climate data from private data providers to extract relevant forward-looking aggregates at a sector and/or country level?
## 2. Constructing forward-looking climate-related physical risk indicators

### Coverage Physical Risk from ISS ESG

| Number of companies in the ISS physical risk dataset by country* and NACE sector | A | B-E | including: | F | G-I | J | K | L | M-N | O-Q | R-S |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| United States | | | | | | | | | | | | |
| Euro Area | | | | | | | | | | | | |
| Japan | | | | | | | | | | | | |
| United Kingdom | | | | | | | | | | | | |
| Canada | | | | | | | | | | | | |
| Other AE | | | | | | | | | | | | |
| China | | | | | | | | | | | | |
| India | | | | | | | | | | | | |
| Other EMDE Asia | | | | | | | | | | | | |
| EMDE Europe | | | | | | | | | | | | |
| EMDE Latin America Caribbean | | | | | | | | | | | | |
| EMDE Middle East and Central Asia | | | | | | | | | | | | |
| EMDE Sub-Saharan Africa | | | | | | | | | | | | |
| Rest of the World | | | | | | | | | | | | |

*Regional aggregates based on the classification from the IMF's World Economic Outlook April 2022

= not enough issuers for aggregation

= not enough issuers for statistical analysis

= barely enough issuers for statistical analysis

= enough issuers for statistical analysis
2. Constructing forward-looking climate-related physical risk indicators

Preliminary results

- The **coverage for many sector-country-combinations is too low** to construct aggregate indicators.

- Therefore, we **further aggregate the countries and regions** into an AE (Advanced Economies) group and an EMDE (Emerging Markets and Developing Economies) group.

- The next slide shows **physical risk indicators for sector-group combinations** and for **two climate scenarios** from the IPCC 5th Assessment Report, RCP4.5 being the more optimistic scenario, RCP8.5 the more pessimistic one.
## 2. Constructing forward-looking climate-related physical risk indicators

### Preliminary results

<table>
<thead>
<tr>
<th>Total annual physical risk in % of revenue by sector and country group</th>
<th>Advanced Economies</th>
<th>Emerging Markets &amp; Developing Economies</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>RCP 4.5</td>
<td>RCP 8.5</td>
</tr>
<tr>
<td>A</td>
<td>0.37%</td>
<td>0.53%</td>
</tr>
<tr>
<td>B, D, E</td>
<td>0.31%</td>
<td>0.40%</td>
</tr>
<tr>
<td>C</td>
<td>0.17%</td>
<td>0.24%</td>
</tr>
<tr>
<td>F</td>
<td>0.08%</td>
<td>0.12%</td>
</tr>
<tr>
<td>G-I</td>
<td>0.10%</td>
<td>0.15%</td>
</tr>
<tr>
<td>J</td>
<td>0.12%</td>
<td>0.16%</td>
</tr>
<tr>
<td>K</td>
<td>0.12%</td>
<td>0.17%</td>
</tr>
<tr>
<td>L</td>
<td>0.35%</td>
<td>0.47%</td>
</tr>
<tr>
<td>M-N</td>
<td>0.11%</td>
<td>0.16%</td>
</tr>
<tr>
<td>O-Q</td>
<td>0.11%</td>
<td>0.15%</td>
</tr>
<tr>
<td>R-S</td>
<td>0.15%</td>
<td>0.20%</td>
</tr>
</tbody>
</table>
3. Constructing forward-looking climate-related physical risk indicators

Challenges

− One main issue is **limited coverage** in company level data.

− The **variation** between the data of different providers is high, similar to other composite indicators.

− The hazard types covered and their **definitions are not consistent** across data providers and therefore need to be taken into account when analysing results.

− Physical risk metrics should be comparable across years and scenarios and **reflect financial damages**.
3. Constructing forward-looking climate-related physical risk indicators

Way forward

- Central banks could be a good place to construct physical risk indicators by combining climate-related with financial data, which is available in central banks (in the Eurosystem: Analytical Credit Datasets, Security Holding Statistics, Centralised Securities Database).

- In the short to medium-term: Making use of existing enterprise-level data from private sources. Combining them with public data can bridge data gaps.

- In the longer-term: To be more precise on detailed physical risks and their financial implications we would need:
  - **New skills**: GIS (Geographic Information System) knowledge required to work with climate-related data on a granular level and close cooperation with climate experts.
  - **Granular financial data**: Exact locations of collateral, counterparties and their branches / facilities as well as their valuation. In addition, data on households is limited.
Thank you for your attention!
References


- The availability of data with which to monitor and assess climate-related risks to financial stability, Financial Stability Board, July 2021, Link: The availability of data with which to monitor and assess climate-related risks to financial stability - Financial Stability Board (fsb.org)


- ECB presents action plan to include climate change considerations in its monetary policy strategy, July 2021, Link: ECB presents action plan to include climate change considerations in its monetary policy strategy (europa.eu)

- International Conference on Statistics for Sustainable Finance, Programme: International Conference on Statistics for Sustainable Finance | Banque de France (banque-france.fr); Presentations: Events (bis.org)

- IFC Report - Sustainable finance data for central banks, December 2021, Link: IFC Report no 14 on Sustainable Finance (bis.org)
Materiality of ESG factors in financial markets and financial statistics\(^1\)

Patrick Slovik\(^2\) and Farah Azman

---

\(^1\) This presentation was prepared for the conference. The views expressed are those of the authors and do not necessarily reflect the views of the BIS, the IFC or the central banks and other institutions represented at the event.

\(^2\) Lead author
Materiality of ESG Factors in Financial Markets and Financial Statistics

Patrick Slovik and Farah Azman

Abstract

The study assesses the materiality of Environmental, Social and Governance (ESG) factors in financial markets and financial statistics. The stocks of and flows in ESG financial assets have reached a systemically-relevant share in the overall financial system. The study explores the implications of materiality for ESG financial statistics while acknowledging that data gaps need to be addressed amid considerable uncertainty. It outlines the necessity to differentiate between single-materiality and double-materiality approaches and defines the concept of financial uncertainty in contrast to financial risk and its implications for materiality.

Keywords: Climate-related financial risks; Double materiality; Environmental, social and governance factors; ESG; Materiality; Sustainability; Sustainable finance; Uncertainty

JEL Classification: G11, G14, G21, M14, Q56
ESG Factors in Financial Markets

Sustainable finance evolved rapidly, both in terms of asset size and diversity of financial products. Sustainable finance became increasingly linked with Environmental, Social and Governance (ESG) criteria. The stocks of and flows in ESG financial assets reached a systemically-relevant share across key markets and asset classes. Financial markets and institutions have a critical financial intermediation role in sustainable finance.

ESG Factors in Debt Securities Markets

The market for sustainable-finance debt securities expanded at a fast pace during the last five years. ESG debt securities issued during 2021 alone amounted to more than USD 1 trillion. The market share of sustainable-finance debt securities in the overall global bond market increased to 11%. In Europe, ESG bonds comprised 20% of all debt securities issued in the region in 2021, representing a fourfold increase from a 5% share in 2017.

<table>
<thead>
<tr>
<th>Global ESG Bonds</th>
<th>Graph 1</th>
</tr>
</thead>
<tbody>
<tr>
<td><img src="image.png" alt="Graph" /></td>
<td></td>
</tr>
</tbody>
</table>

Sources: S&P CIQ; Authors’ calculations.

Notwithstanding Europe’s dominant position in sustainable finance, ESG debt securities issuance recorded strong growth across all other key global regions. Sustainable-finance debt securities issued in 2021 by European counterparties represented 55% of global ESG debt securities issued, followed by the Asia-Pacific region with a 22% share, the North American region with a 15% share, while the remaining regions collectively accounted for 8%.

Types of ESG Financial Instruments

ESG financial instruments evolved into several key types, reflecting the growing diversity of sustainable-finance products. While the categories originally derive from the principles developed for sustainable-finance bond markets (ICMA, 2020, 2021a, 2021b, 2021c), such categorisation can
also be applied to a broader set of sustainable finance products, with environmentally beneficial or socially beneficial uses of proceeds.

### Types of ESG Financial Instruments

<table>
<thead>
<tr>
<th>Types of ESG Financial Instruments</th>
<th>Table 1</th>
</tr>
</thead>
<tbody>
<tr>
<td>Green Financial Instruments</td>
<td>Financial instruments where the proceeds are used to fund environmentally beneficial projects. ¹</td>
</tr>
<tr>
<td>Social Financial Instruments</td>
<td>Financial instruments where the proceeds are used to fund socially beneficial projects. ²</td>
</tr>
<tr>
<td>Sustainability Financial Instruments</td>
<td>Financial instruments to fund both environmentally and socially beneficial projects.</td>
</tr>
<tr>
<td>Sustainability-Linked Financial Instruments</td>
<td>Financial instruments linked to the issuer achieving predefined sustainability targets.</td>
</tr>
</tbody>
</table>

*Sources: Authors’ review based on ICMA (2020, 2021a, 2021b, 2021c).*

Among ESG debt securities, green bonds had the largest market share of 56% in 2021. Social bonds surged to 23% during the pandemic due to an increase in projects with socially beneficial use of proceeds (Moody’s, 2021). Sustainability bonds, which combine characteristics of green and social bonds, also gained market share, representing 15%. Sustainability-linked bonds, with returns linked to achievements of predefined ESG targets, had a 6% market share.

### Global ESG Bonds Breakdown in 2021

*Sources: S&P CIQ; Authors’ calculations.*

1 Environmentally beneficial uses of proceeds cover areas such as renewable energy, energy efficiency, green buildings, pollution prevention, climate-change mitigation, biodiversity, clean transportation, or water management.

2 Socially beneficial uses of proceeds cover areas such as basic infrastructure, essential services, employment generation, affordable housing, food security and sustainability, socioeconomic development, or inequality reduction.
ESG Factors in Mutual Funds and Exchange-Traded Funds

Sustainable-finance funds represent mutual funds and exchange-traded funds that integrated ESG criteria into their investment strategies and portfolio selection processes (although the ESG investment methodologies may vary among funds). Assets under management of global sustainable-finance funds increased to over USD 2.7 trillion in 2021, representing about 7% of the total global mutual fund and exchange-traded fund industry (Morningstar, 2022).

<table>
<thead>
<tr>
<th>ESG Mutual Funds and Exchange-Traded Funds (ETFs)</th>
<th>Graph 3</th>
</tr>
</thead>
<tbody>
<tr>
<td><img src="image" alt="Bar Graph" /></td>
<td><img src="image" alt="Pie Chart" /></td>
</tr>
</tbody>
</table>

* Rest of the World category comprised funds domiciled in Asia, Japan, Australia, New Zealand, and Canada.
Sources: Morningstar; Authors' calculations.

Europe has also been at the forefront of the ESG mutual fund and exchange-traded fund market. Based on assets under management, more than 82% of such funds were domiciled in Europe, followed by the United States at 13%, and the rest of the world at 5%. The number of ESG mutual funds and exchange-traded funds expanded to close to 6,000 in 2021, a fourfold increase in the total number of sustainable-finance funds during the last five years.

ESG Factors in Equity Markets and Investment Principles

Concerning ESG factors in equity markets, publicly listed companies have increasingly considered ESG criteria, as reflected in their sustainability reports and disclosures. Out of the 500 largest listed companies in the United States, represented in the S&P 500, more than 90% issued sustainability reports or disclosures (G&A Institute, 2021). Thus, the considerations for ESG factors among listed companies could also be viewed as systemically relevant.

The Principles for Responsible Investment, supported by the United Nations, with close to 4000 signatories, are a further illustration of the integration of ESG criteria into the investment process and financial markets. The signatories of the principles committed to incorporating ESG issues into investment analysis and decision-making. The assets under management of the signatories of the principles amounted to over USD 100 trillion (PRI, 2021).
Materiality of ESG Factors in Financial Institutions

The materiality of ESG factors arises from two distinct approaches. The first approach to materiality reflects the impact of ESG factors on the entities' financial performance and risk profile. The second approach to materiality reflects the impact of the entities' business activities on the environment and stakeholders. These two approaches were utilised in various ESG-related frameworks with varying terminologies. Double materiality refers to the blend of both approaches.

<table>
<thead>
<tr>
<th>Different Terminologies of Materiality of ESG Factors</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Impact of ESG Factors on Entity</strong></td>
</tr>
<tr>
<td>Financial Materiality</td>
</tr>
<tr>
<td>Business Materiality</td>
</tr>
<tr>
<td>Outside-in Materiality</td>
</tr>
<tr>
<td>Impact on Business</td>
</tr>
<tr>
<td>Impact on Entity</td>
</tr>
</tbody>
</table>

Double Materiality

Sources: Authors’ review based on ESG frameworks and sustainability reports.

Impact of ESG Factors on Financial Institutions

Outside-in materiality refers to the impact of external environmental and social factors on financial institutions. The materiality approach is also referred to as financial materiality as it relates to the impact on entities' financial performance and business value. Financially material ESG factors could have implications for investors, regulators, and shareholders concerned with financial institutions’ performance, soundness, and enterprise valuation.

Impact of Financial Institutions on ESG Factors

Inside-out materiality refers to entities' impact through their business activities on sustainable development goals. It is also referred to as environmental and social materiality, as financial institutions' business affects environmental and socioeconomic factors. As business activities impact wider stakeholders, such as the environment, employees, customers, and communities, this approach is also known as stakeholder materiality.
Materiality of ESG Factors in Financial Statistics

Developing sustainable finance statistics requires distinguishing between the two approaches to materiality. The duality of the materiality of ESG factors necessitates separate sets of methodologies also for statistical purposes. The first set focuses on the impact of sustainability factors on financial institutions (outside-in materiality). The second set focuses on the impact of financial institutions on sustainability factors (inside-out materiality).

Measuring Sustainability Factors of Financial Exposures

The first set of statistical methods ascribes to measuring, modelling, or pricing sustainability risks of the financial institutions’ exposures. It broadly relates to financial materiality and requires measuring and disclosing material information to assess the risk profile, financial performance, and valuation of financial institutions. Such approaches require granular and forward-looking statistics and assessments on the sustainability of financial institutions’ exposures.³

Measuring Sustainability Factors of Business Transition

The second set of statistical approaches relates to the transition of business models of financial institutions towards sustainability. It broadly relates to stakeholder materiality and requires measuring and disclosing information relevant to a wider group of stakeholders on the external impact of business activities of financial institutions. Such approaches require measuring the transition of financial institutions and financial systems towards sustainability.

Materiality of ESG Factors in Financial Ratings

The dichotomy of materiality was also evident in the approaches of rating agencies. Credit ratings integrate material implications of ESG factors on the likelihood of default of borrowers on their financial obligations in given time horizons.⁴ Creditworthiness materiality is thus similar to financial materiality. In contrast, ESG ratings reflect wider spectrums of ESG factors addressing the impacts on broader groups of stakeholders, i.e. the environment and society.⁵

³ Along these lines, the International Conference on Statistics for Sustainable Finance identified pressing data gaps in the lack of granular firm-level data and the absence of forward-looking data on future sustainability paths of firm-level counterparties (IFC, 2022).

⁴ For instance, S&P Global Ratings reported 550 credit rating actions driven primarily by ESG factors during 2021 (S&P, 2022a). In addition to integrating ESG factors in credit ratings, S&P Global Ratings also conducts separate ESG evaluations. Most credit and ESG rating agencies practice the duality of approaches to materiality.

⁵ The dichotomy of materiality, when not properly recognised, may result in identifications of improper causal hypotheses in research studies. Such as testing causal relationships between inside-out ESG factors (particularly present in ESG ratings) and entities’ financial performance (impacted particularly by outside-in ESG factors).
Uncertainty of ESG Factors

The manner in which environmental risks translate into financial risks over time remains an area of significant uncertainty (EBA, 2022). Subject to considerable uncertainty, climate-related financial risks cannot be accurately measured (Chenet et al., 2019). While there is high confidence in the severity of climate-related hazards, how they will interact with future socioeconomic developments that determine the scale of exposures and vulnerabilities remains uncertain (OECD, 2021).

Financial Risk and Financial Uncertainty

Financial risk relates to stochastic positive or negative outcomes with determinable likelihoods (Knight, 1921). In contrast, financial uncertainty corresponds to outcomes with indeterminate probabilities (Knight, 1921; Keynes, 1937; Lawson, 1985). While financial risks are determinable based on available information, knowledge, and experience, financial uncertainty occurs due to a lack of adequate information, knowledge, and experience at the time (Slovik, 2010).

<table>
<thead>
<tr>
<th>Financial Risk</th>
<th>Financial Uncertainty</th>
</tr>
</thead>
<tbody>
<tr>
<td>Potential for adverse outcomes with determinate likelihoods of occurrence.</td>
<td>Potential for adverse outcomes with indeterminate likelihoods of occurrence.</td>
</tr>
<tr>
<td>Determinate due to available information, knowledge, and experience at the time.</td>
<td>Indeterminate due to unavailable information, knowledge, and experience at the time.</td>
</tr>
<tr>
<td>Predictable through conventional risk management approaches.</td>
<td>Not predictable through conventional risk management approaches.</td>
</tr>
</tbody>
</table>

Sources: Author’s review based on Knight (1921), Keynes (1937), Lawson (1985), and Slovik (2010).

Financial Uncertainty and Materiality

Financial risk and financial uncertainty represent two distinct concepts with different implications for the decision-making of financial institutions and financial authorities. In decision-making under risk, potential stochastic positive or negative outcomes and their probabilities are determinable. In decision-making under uncertainty, these probabilities are indeterminable with adequate precision, which often occurs due to structural shifts and a lack of historical precedence.

In view of climate-related financial uncertainty, conventional risk management approaches and estimation techniques might not be sufficient to determine the financial implications of climate-related hazards with satisfactory precision. As a result, reliance on financial materiality (outside-in materiality) alone might not be sufficient. In view of financial uncertainty, a double-materiality approach offers preferred policy options in lieu of a single-materiality approach.
Reference:


Materiality of ESG Factors in Financial Markets and Financial Statistics

Patrick Slovik and Farah Azman

Abstract

The study assesses the materiality of Environmental, Social and Governance (ESG) factors in financial markets and financial statistics. The stocks of and flows in ESG financial assets have reached a systemically-relevant share in the overall financial system. The study explores the implications of materiality for ESG financial statistics while acknowledging that data gaps need to be addressed amid considerable uncertainty. It outlines the necessity to differentiate between single-materiality and double-materiality approaches and defines the concept of financial uncertainty in contrast to financial risk and its implications for materiality.
ESG Factors in Financial Markets: Debt Securities Market

- Sustainable finance evolved rapidly in terms of asset size and diversity of financial products. ESG financial assets reached a systemically-relevant share across most key markets and asset classes.

- ESG debt securities issued during 2021 alone amounted to more than USD 1 trillion. The market share of sustainable-finance debt securities in the overall global bond market increased to 11%.

---

**Global ESG Bonds**

<table>
<thead>
<tr>
<th>Annual Global ESG Bond Issuance</th>
<th>Market Share of ESG Bonds</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>USD billion</strong></td>
<td><strong>Per cent</strong></td>
</tr>
<tr>
<td>2017: 100</td>
<td>0%</td>
</tr>
<tr>
<td>2018: 200</td>
<td>2%</td>
</tr>
<tr>
<td>2019: 300</td>
<td>4%</td>
</tr>
<tr>
<td>2020: 700</td>
<td>6%</td>
</tr>
<tr>
<td>2021: 1,200</td>
<td>8%</td>
</tr>
</tbody>
</table>

**Sources:** S&P CIQ; Authors’ calculations.
ESG Factors in Financial Markets: Types of Financial Instruments

- ESG financial instruments evolved into several key categories, reflecting the growing diversity of sustainable-finance instruments.
- Green bonds had a market share of 56%, followed by social bonds at 23%, sustainability bonds at 15%, and sustainability-linked bonds at 6%.

Global ESG Bonds Breakdown in 2021

<table>
<thead>
<tr>
<th>ESG Bond Types</th>
<th>Per cent</th>
</tr>
</thead>
<tbody>
<tr>
<td>Green Bonds</td>
<td>56%</td>
</tr>
<tr>
<td>Sustainability Bonds</td>
<td>23%</td>
</tr>
<tr>
<td>Social Bonds</td>
<td>15%</td>
</tr>
<tr>
<td>Sustainability-Linked Bonds</td>
<td>6%</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>ESG Bond Issuers</th>
<th>Per cent</th>
</tr>
</thead>
<tbody>
<tr>
<td>Non-Financial Corporates</td>
<td>16%</td>
</tr>
<tr>
<td>International Organisations</td>
<td>15%</td>
</tr>
<tr>
<td>Governments</td>
<td>37%</td>
</tr>
<tr>
<td>Government Agencies</td>
<td>11%</td>
</tr>
<tr>
<td>Financial Institutions</td>
<td>21%</td>
</tr>
</tbody>
</table>

Sources: S&P CIQ; Authors’ calculations.
ESG Factors in Financial Markets: Mutual Funds and Exchange-Traded Funds

- Sustainable-finance funds represent mutual funds and exchange-traded funds that integrated ESG criteria into their investment strategies and portfolio selection processes.

- Assets under management of global sustainable-finance funds increased to over USD 2.7 trillion, representing about 7% of the total global mutual funds and ETFs.

### ESG Mutual Funds and Exchange-Traded Funds (ETFs)

<table>
<thead>
<tr>
<th>Global ESG Mutual Funds and ETFs</th>
<th>Regional Distribution of ESG Mutual Funds and ETFs</th>
</tr>
</thead>
<tbody>
<tr>
<td>USD billion</td>
<td>Per cent</td>
</tr>
<tr>
<td>2017</td>
<td>Europe (82%)</td>
</tr>
<tr>
<td>2018</td>
<td>United States (13%)</td>
</tr>
<tr>
<td>2019</td>
<td>Rest of the World* (5%)</td>
</tr>
<tr>
<td>2020</td>
<td></td>
</tr>
<tr>
<td>2021</td>
<td></td>
</tr>
</tbody>
</table>

*Rest of the World category comprised funds domiciled in Asia, Japan, Australia, New Zealand, and Canada

Sources: Morningstar; Authors’ calculations.
ESG Factors in Financial Markets: Equity Markets and Investment Principles

ESG Factors in Equity Markets

- Publicly listed companies have increasingly considered ESG criteria, as reflected in their sustainability reports and disclosures.

- Out of the 500 largest listed companies in the United States, represented in the S&P 500, more than 90% issued sustainability reports.

ESG Factors in Investment Principles

- The signatories of the Principles for Responsible Investment committed to incorporating ESG issues into investment analyses and decision-making.

- Assets under management of the signatories amounted to over USD 100 trillion, a further illustration of the ESG integration into financial markets.
Materiality of ESG Factors

- Materiality of ESG factors arises from two divergent approaches. The first approach to materiality reflects the impact of ESG factors on the entities' financial performance and risk profile.

- The second approach to materiality reflects the external impact of the entities' business activities on the ESG factors. The blend of both approaches is referred to as double materiality.

<table>
<thead>
<tr>
<th>Different Terminologies of Materiality of ESG Factors</th>
</tr>
</thead>
<tbody>
<tr>
<td>Impact of ESG Factors on Entity</td>
</tr>
<tr>
<td>---------------------------------</td>
</tr>
<tr>
<td>Financial Materiality</td>
</tr>
<tr>
<td>Business Materiality</td>
</tr>
<tr>
<td>Outside-in Materiality</td>
</tr>
<tr>
<td>Impact on Business</td>
</tr>
<tr>
<td>Impact on Entity</td>
</tr>
</tbody>
</table>

Double Materiality

Sources: Authors’ review based on ESG frameworks and sustainability reports.
Materiality of ESG Factors in Financial Statistics

- The duality of materiality of ESG factors necessitates separate sets of methodologies for developing sustainable-finance statistics, differentiating between single materiality and double materiality.

Measuring Sustainability Factors of Financial Exposures

- The first approach focuses on the impact of sustainability factors on financial institutions (outside-in approach), requiring data to evaluate financial risks and performance of financial institutions.

Measuring Sustainability Factors of Business Transition

- The second approach focuses on the impact of financial institutions on sustainability factors (inside-out approach), requiring data on sustainability transformation and impact of financial institutions.

Materiality of ESG Factors in Financial Ratings

- The dichotomy of materiality was also evident in the approaches of rating agencies. Credit ratings reflect material ESG factors on the default likelihood. ESG ratings reflect wider materiality spectrums.
Uncertainty of ESG Factors

- Estimating financial implications of sustainability vulnerabilities remains subject to considerable uncertainty. As a result, reliance on financial materiality alone might not be sufficient.

- In view of considerable financial uncertainty of sustainability vulnerabilities, a double-materiality approach might offer preferred policy options in lieu of a single-materiality approach.

<table>
<thead>
<tr>
<th>Financial Risk</th>
<th>Financial Uncertainty</th>
</tr>
</thead>
<tbody>
<tr>
<td>Potential for adverse outcomes with a determinate likelihoods of occurrence.</td>
<td>Potential for adverse outcomes with indeterminate likelihoods of occurrence.</td>
</tr>
<tr>
<td>Determinate due to available information, knowledge, and experience at the time.</td>
<td>Indeterminate due to unavailable information, knowledge, and experience at the time.</td>
</tr>
<tr>
<td>Predictable through conventional risk management approaches.</td>
<td>Not predictable through conventional risk management approaches.</td>
</tr>
</tbody>
</table>

Sources: Authors’ review based on Knight (1921), Keynes (1937), Lawson (1985), and Slovik (2010).
An estimation of the carbon footprint in Spanish credit institutions’ business lending portfolio – Experimental statistics for credit institutions in Spain

Luis Ángel Maza,
Bank of Spain

---

1 This presentation was prepared for the conference. The views expressed are those of the author and do not necessarily reflect the views of the BIS, the IFC or the central banks and other institutions represented at the event.
An estimation of the carbon footprint in Spanish credit institutions’ business lending portfolio

Experimental statistics for credit institutions in Spain

Luis Ángel Maza (Banco de España).

Abstract

This study proposes a set of indicators to estimate the carbon footprint of the business lending of Spanish credit institutions. The growing interest of our societies in environmental issues has brought into the general debate the need to analyse how financial institutions perform in their role to facilitate the fight against climate change and the green transition. In this respect, it is essential to have quality environmental information and to establish robust methodologies to assess the climate exposures of the financial sector. This paper seeks to contribute to this debate, offering experimental statistics to measure the degree of exposure of the banking sector in Spain to the risks involved in the transition towards a more sustainable economic model in the financing granted to companies.

Keywords: climate change, carbon footprint, financial risks.

JEL classification: Q50, Q56, G10, G20.
1. Introduction

The growing social demand regarding environmental issues observed in recent decades and the response from public authorities (including the establishment of climate targets) have driven reflections on the role the financial system should play in this transformation towards more sustainable economic growth models. The financial sector needs to prioritise efficiency when channelling the resources needed to drive this transformation and pay attention to the potential transition and physical risks that will affect companies and households in the near term and will, consequently, have a bearing on financial institutions as well.

One of the most significant international environmental public policy commitments of recent years has been the 2015 Paris Agreement. This Agreement set targets for reducing greenhouse gas (GHG) emissions, which have been established as being the main cause of global warming. Since then, important initiatives have emerged in the field of financial sector regulation to incorporate environmental and green transition considerations (for example, in relation to financial stability and prudential supervision) that are compatible with the climate targets. These new regulations flag a need to assess the financial sector’s behaviour and its environmental responsibility and to pay greater attention to evaluating the exposures of financial institutions (in particular, credit institutions) through the use of environmental parameters. However, the initiatives for assessing financial institutions’ environmental risks are still in an initial phase of specification and international harmonisation, which prevents a clear understanding of the efforts and progress needed to meet the climate change targets.

This paper presents a proposal for an indicator to quantify the GHG carbon footprint of Spanish credit institutions’ portfolio of loans to resident firms, with the aim of evaluating the banking sector’s behaviour in: (i) channelling financial resources towards less polluting activities and (ii) measuring the transition risks to which financial institutions will be subject in the coming years, due to the economic and financial performance of certain economic sectors in the face of the decarbonisation of the economy.

This proposal will thus contribute to the methodological debate that currently surrounds the selection of potential indicators to measure these transition risks, in the absence of an internationally agreed definition. So far, there are only experiences and initiatives of an autonomous nature in the selection of such indicators.

In the spirit of ensuring that the proposed methodology can be reproduced by other interested countries or agents, information sources and calculation procedures that are broadly available in other jurisdictions have primarily been used when defining the methodology.

To this end, the second section briefly describes the data sources and the methodology used to calculate this carbon footprint indicator. The third section...
shows the main results obtained, in aggregate and individual terms. Lastly, in the fourth section, possible refinements in and extensions to the application of this procedure are reviewed. A final section of conclusions is also included.

2. Data sources and methodology

When describing the procedure and the methodology used to prepare the carbon footprint indicator, it may be of interest first to carry out a detailed review of the sources of information used. This description will allow a better understanding of the characteristics, benefits and limitations of the information content of the indicator and serve as a model for its replicability in other arenas and geographical areas.

Polluting emissions

In order to assess the carbon footprint of financial institutions’ loan portfolio, it is key that information on the GHG emissions generated by productive activity (mainly, firms) is available. An ideal approach to obtaining this information would be to have individual data (at company level) on (i) the direct emissions (called, in this context, “scope 1” emissions) generated by companies in their direct consumption of fossil fuels, and (ii) their indirect emissions (called “scope 2 and 3” emissions), which derive from suppliers’ fuel consumption in the inputs incorporated by companies into their production process. However, the current situation regarding information availability is a long way off this approach. The detailed information on polluting emissions is limited to a small number of companies (generally large corporations), and there are sometimes difficulties in assigning emissions to specific companies, as the data are disseminated in terms of business groups or installations, with no details on the company or their geographical allocation.

Such restrictions on access to information advise the use of aggregate statistics on the atmospheric polluting emissions of each sector of activity, which, while less accurate than the real data of each company, will allow a homogeneous comparison and provide complete information for the economy as a whole. These aggregate data are available in the environmental accounts that are usually drawn up by the national statistics institutes (in the case of Spain, by the Instituto Nacional de Estadística, or INE), following the methodology established by the United Nations for the System of Environmental-Economic Accounting.

The variable selected for measuring these polluting emissions in Spain is total GHG emissions, measured in thousands of tonnes of CO₂ equivalent in annual terms, which are available in the INE’s Air Emissions Accounts. The level of detail used corresponds to the breakdown of the 64 sectors of activity according to the National Classification of Economic Activities (NACE Rev. 2).

Output level

In order to calibrate the degree of pollution intensity of the sectors of activity, the emissions must be compared with a measure of the quantity of goods and
services produced by each economic sector. Of the various indicators available for measuring the activity generated, the level of output in a financial year, measured in terms of monetary units (millions of euro), has been selected. This variable is available in the annual National Accounting statistics that the INE also prepares for Spain.

Input-output table

In this exercise, to calculate the carbon footprint of the sectors of activity, the information on direct pollution needs to be completed with an assessment of indirect GHG emissions, which derive from the resources incorporated into the production process. These calculations can be made using the input-output table available in the National Accounts framework. This statistic contains a very detailed description of the characteristics of the productive sectors of the economies. In the case of Spain, the most recent information drawn up by the INE corresponds to 2016.

Bank loan portfolio

As regards access to the information on Spanish banks’ credit exposures to resident companies, the quarterly financial reporting statement that credit institutions send to the Banco de España is used. This reporting form contains details of the stock of loans according to the economic activity of the borrowers (NACE Rev. 2 sections).

This information has been completed with aggregate data from the Banco de España’s Central Credit Register (CCR) for those sectors of activity not available in the reporting statements, to standardise the information on loans with the breakdowns at the level of the 64 sectors of activity available from other sources.

Direct and indirect polluting emission coefficients

In the proposed procedure for calculating the carbon footprint indicator, a decisive factor is obtaining the carbon dioxide (CO₂) emission coefficients per unit of output for each sector of activity. These coefficients or ratios seek to assess the intensity of (direct and indirect) polluting emissions in the output of each economic sector.

The direct coefficients (q^direct) are calculated as the ratio between GHG emissions (expressed in thousands of tonnes of CO₂ equivalent) and the total output of each sector (in millions of euro), according to formula 1:

\[
q_{\text{direct}}^{i,t} = \frac{\text{Greenhouse Gas Emission}_{i,t}}{\text{Total output}_{i,t}}
\]

for each sector of activity (i) and year (t).

The total coefficients (q^totals) are the sum of the direct coefficients and the indirect effect of the polluting emissions produced in obtaining the intermediate
An estimation of the carbon footprint in Spanish credit institutions' business lending portfolio

inputs used by each sector of activity, as expressed in formula 2. This estimate draws on information from the National Accounts input-output matrix, which includes how the final output of each sector is incorporated as inputs by the other sectors.

\[ q_{it}^{\text{total}} = (I - A)^{-1} q_{it}^{\text{direct}} \]

A: the coefficient matrix of the input-output table in the Annual National Accounts of Spain for 2016

\((I - A)^{-1}\): the Leontief inverse matrix

The values of the direct coefficients will depend on the inherent characteristics of the productive structure of each sector of activity. This means that the changes in the coefficients over time are small and that the changes in the sectors’ relative positions in pollution intensity levels are also infrequent. Chart 1 shows the results obtained from the direct coefficients of the most polluting sectors in Spain. The results indicate that electricity and gas supply, transportation, manufacturing and agriculture are the productive sectors with the highest GHG emissions per unit of output.

Chart 1.

As regards the results of the total coefficients by sector of activity (i.e. direct emissions plus the polluting effects of the inputs), there are small changes in the ranking of the most emission-intensive sectors, although some new sectors are classified among the most polluting sectors (for example, mining and quarrying and storage activities). Turning to the indirect component, electricity and gas supply again stands out as the industry with the largest component stemming from emissions generated in obtaining the inputs used in its production process (see Chart 2).
Chart 2.

The time analysis of the total coefficients of the sectors of activity allows for an assessment of production efficiency over time and the degree of compliance with the GHG emission reduction targets. Chart 3 shows the results of the sectors’ coefficients in the period 2008-2019 (represented through the percentiles observed in each year). According to this information, emission intensity has decreased across the board in recent years, particularly in the most polluting sectors (decrease in the 90th and 75th percentiles).

Chart 3.

Indicator of the carbon footprint intensity of loans

Once the information on the total GHG emission coefficients at the sector of activity level is available, the indicator of the carbon footprint intensity of the business lending portfolio (IHCO2P) is calculated as the average of the coefficient totals of the sectors of activity, weighted by the stock of financing granted to each sector in the bank loan portfolio, as expressed in formula 3.

\[ IHCO2P_t = \frac{\sum_i P_{it} q_{it}^{total}}{\sum_i P_{it}} \]

\( P_{it} \): stock of loans from credit institutions at year-end by sector of activity (i) and year (t).
The IHCO2P, evaluated for all credit institutions, thus represents the average ratio of polluting emissions of productive activities that obtain bank financing to total bank loans granted in Spain.

This indicator is interpreted as follows: a drop in the level of the indicator signifies an improvement in the carbon footprint (relative reduction in polluting emissions), while an increase represents a worsening (increase in GHG pollution intensity).

3. Main results

The methodology proposed in the previous section for calculating the IHCO2P indicator for all credit institutions in Spain in the period 2008-2019 can be used to assess the behaviour of the polluting emission intensity of bank financing. Chart 4 shows the results of the IHCO2P and identifies how in recent years the indicator has shown a downward trend, in line with the general improvement in the polluting emission coefficients.

Chart 4.

Given the difficulties in interpreting the indicator in its original units (tonnes of CO₂ per million euro) and for comparison purposes, the indicator is re-expressed in terms of an index (with a base year of 2008, the start of the series). A comparison of the IHCO2P with changes in the emission intensity of the Spanish economy in 2008-2019 shows a cumulative reduction of a very similar magnitude (see Chart 5).
To identify the elements behind the changes in the IHCO2P, the factors underlying this behaviour have been analysed, identifying the influence of (i) the changes in the polluting emission intensity of the sectors and (ii) the changes due to shifts in the composition of the loan portfolio of all banks.

To identify these effects, a simulation exercise has been carried out, where the baseline scenario envisages a stable composition of the base-year loan portfolio (data from 2008). The results obtained indicate that the primary factor behind the improvement in the IHCO2P is the decrease in the emission coefficients, while, in cumulative terms, the contribution of the changes in the composition of credit exposures (the result of a combination of supply and demand factors in bank financing during the period analysed) appears to be quite marginal.

A more detailed analysis of the changes in the loan portfolio composition can be carried out by classifying the sectors of activity according to the values of their total emission coefficients. In the absence of a fully accepted taxonomy for categorising
productive sectors based on their pollution intensity, in this exercise the sectors of activity have been divided into two groups: more or less polluting, depending on their average polluting emission intensity in 2008-2019. Thus, those sectors that exceed the median emission coefficient of the 64 analysed sectors are classified in the "more polluting" category, while those whose emission coefficients are below the median of the distribution are classified as "less polluting".

Chart 7 shows the result of this analysis and evidences a slight shift in the composition of Spanish credit institutions’ loan portfolio towards less polluting sectors in the later years of the analysed period.

Chart 7.

**Detail of the composition of loans from highly polluting sectors of activity (*)**

(*) “Highly polluting” sectors of activity are those whose total emission coefficients exceed the 75th percentile of the distribution.
More specifically, the bulk of financing granted to the more polluting productive sectors was concentrated in (i) transportation and storage, (ii) electricity, gas, steam and air conditioning supply and (iii) the food industries (see Chart 8).

The results analysed above in the behaviour of the IHCO2P derive from the application of the methodology drawing on the aggregate information of the Spanish credit institutions sector. However, the approach proposed in the carbon footprint calculation can be applied to the individual data of the loan portfolio of each credit institution operating in the Spanish market, and the changes in this indicator can be evaluated individually. This approach is complementary to the aggregate analysis and can be used to analyse behaviours (for example, to identify influential observations) in the distribution of the IHCO2P for the entire sample of credit institutions in Spain.

To this end, the IHCO2P values have been calculated for each credit institution (160 banks) at two moments in time (in 2017 and 2020). The results obtained have been represented by kernel density functions, to show graphically the distribution of IHCO2P values within the population of institutions (see Chart 9).

Chart 9.

IHCO2P Analysis based on individual data (credit institution level)

This approach, which is complementary to the aggregate view and uses more granular information, would indicate a shift to the left of the density function in 2020 compared to 2017, suggesting that the improvement observed in the IHCO2P in aggregate terms is compatible with a general reduction in the carbon footprint of credit institutions’ loan portfolio at an individual level, meaning that the reduction in the carbon footprint is widespread in the overall population of credit institutions in Spain in the most recent period.
4. Possible methodological extensions in the calculation of the indicator

The methodology proposed in the foregoing sections for the calculation of the loan portfolio carbon footprint to obtain this experimental statistic enables both aggregate and individual data to be obtained that can be used to assess the composition of the financial flows channelled towards more polluting activities and to analyse the risks assumed by the financial sector in these exposures. However, this approach may be subject to refinement and extensions with potential complements, depending on the additional information available and the approaches that are to be incorporated into the indicator’s information content. This section details a set of reflections and improvements of a methodological nature that could be incorporated into the definition of the indicator.

Use of features related to the productive and financial structure of the sectors

When evaluating the carbon footprint in business financing, it may be interesting to incorporate elements that estimate the impact on the granting of funds to carry out productive activities and that, in turn, affect the level of polluting emissions. These factors could be related to the intrinsic characteristics of the economic activities and would refer to (i) the productive structure of the economic sectors and (ii) their financial structures.

As regards the productive structure, one highly conditioning element when gauging the influence of the financing granted is the level of investment needed to carry out productive activities (for example, the amount of capital goods needed). This is often specific to each sector of activity. For example, industrial sectors generally have higher capital goods investment needs than companies belonging to the services sectors.

Likewise, financing structures are not unique to each company and, on many occasions, they are also determined by the sector of activity in which the business project is developed. For example, companies engaged in retail trade can access financing sources with no financial cost for relatively larger volumes (financing granted by trade suppliers as a result of payment deferrals), while in other sectors of activity, financial liabilities are only available from external sources, at a cost (through bank financing).

Thus, both the productive and the financial structure may be relevant when assessing the carbon footprint of bank loans, given that the same volume of financing granted to two companies engaged in different activities may have a completely heterogeneous effect on the scale of the activity to be financed and its environmental impact.
Formula 4 incorporates this modification into the definition of the **polluting emission coefficients by including information on the economic and financial characteristics** of the sectors of activity.

Given that these variables are often only available for samples of companies, rather than for the population as a whole, in the adjustment of the calculations of the corrected emission coefficients \( q_{t}^{\text{modified}} \), **ratios from representative information bases for non-financial corporations**, such as those prepared by the Central Balance Sheet Data Office of the Banco de España, are used. In order to facilitate reproducing this exercise for other countries, Spain’s contribution to the BACH (Bank for the Accounts of Companies Harmonized) database of the European Committee of Central Balance Sheet Data Offices is used. **The ratios used to incorporate the productive and financial structures** of the sectors of activity into these calculations correspond to the BACH variables R41 (Asset-turnover ratio) and L (Liabilities), respectively. To establish a more structural nature for the data provided by these coefficients, the averages of these values in the period 2008-2019 for the sample of Spanish non-financial corporations have been used for the breakdown of the 64 sectors of activity.

\[
q_{t}^{\text{modified}} = q_{t}^{\text{totals}} X \frac{\text{Output (net turnover)}_t}{\text{Total balance sheet}_t} \frac{\text{Productive structure ratio}}{\text{Leverage ratio}}
\]

Applying this new definition of the polluting emission coefficients for the sectors of activity, **the results of the carbon footprint indicator have been obtained, making it possible to adjust the “true” influence of bank loans on the impact of the companies’ polluting activities**. As can be seen in Chart 10, once these adjustments are included, the results would show a less pronounced decrease, almost nearing stability, in the IHCO2P throughout the analysed period.
Limitations in the use of information on loans according to economic activity rather than purpose

One of the most significant limitations in analyses of the impact of the financial sector’s actions on polluting emissions derives from the restrictions on access to information on the specific environmental characteristics of the business projects that are being financed, beyond the general activity carried out by the company. Usually, information on granted financing only contains reference data on the risk holder, not on the purpose for which the funds are used. The exercises to quantify the footprint could be enriched if the purpose of the loans were included in the loan portfolio classification (for example, the acquisition of electric vehicles, or energy efficiency or environmental sustainability projects carried out by companies), which would allow for a better measurement of the carbon footprint.

The current classifications of economic activities do not contain adequate details to assess polluting emissions

The current international classifications of economic activities do not incorporate details of the sectors of activity that would allow a detailed and tailored analysis of the carbon footprint. This hampers the availability of information on changes in production models that need to be identified and driven through public policy action. One example of such information deficits is the lack of differentiation of electricity generation activities through renewable energies within electricity production.

The significant shortcomings of the current international codes of economic activities (ISIC Rev 4 and NACE) are well known in the ongoing discussions on the reform and adaptation of these classifications. Thus, issues related to improving access to the details of environmental information have been incorporated into the deliberations on future updates of these classifications. A successful adaptation of these details will enable higher quality statistical work to be prepared, which is crucial for economic agents and public authorities in their decision-making.

Access to individual data on emissions and loans (at company level) could refine this measurement

The methodology proposed in this paper for calculating the carbon footprint exclusively uses aggregate information (at the sector of activity level) both in the analysis of the behaviour of polluting emissions and as regards the composition of loan portfolios. One possible improvement to the estimation of the carbon footprint of bank loans could be directed towards the use of information from those segments of companies (mainly, large companies) for which detailed and comprehensive information on polluting emissions and on loans is available and can be incorporated into these calculations. However, for the bulk of the population of companies, such detailed information will not be available on an individual basis, so aggregate data will always need to be used or individual data will have to be estimated and imputed.
Treatment of loans to holding companies and head offices

In these exercises to estimate the carbon footprint, there is some weakness which has a bearing on the perfect calibration of the impact of the funds received by the holding companies and head offices of business groups, corresponding to the NACE Rev. 2 sectors of activity 6420 and 7010, respectively. The companies classified under these groups frequently channel the financing received towards the group companies that actually carry out the business activities and that, therefore, effectively generate the polluting emissions. However, as a result of this channelling, the relationship between the initial classification of the credit exposures by sector carried out by the credit institutions and the final recipients of these funds is lost.

In order to overcome these limitations, one appropriate way to assign the carbon footprint to the bank loans granted to holding companies and head offices would be to have detailed information on the economic activities carried out by the “productive” subsidiaries of their business groups (for example, industrial or service activities) and to assign their environmental characteristics to the estimate of the carbon footprint of such credit exposures. In the case of Spain, bank financing granted to holding companies and head offices represents around 10% of the total financing to non-financial corporations and, as a result, the effect of incorporating this adjustment into the carbon footprint calculations would not be of great significance, but it would help to better calibrate the results.

5. Conclusions

The greater social awareness of environmental deterioration and global concern about climate change have been reflected in recent years in various initiatives undertaken by public authorities to set targets for reducing GHG emissions. Similarly, the need to calibrate the role that the financial system will play in fostering the green transition towards more sustainable economic models has been added to the environmental agenda of international forums.

In this respect, establishing indicators that calibrate financial institutions’ carbon footprint will make it possible to analyse both developments in the funds channelled towards less polluting activities and the identification of the risks that financial agents will assume in this transformation of the consumption and production patterns of our economies. Within the current debate on the most appropriate indicators for measuring these phenomena, this paper seeks to make a methodological contribution and serve as an example for measuring the carbon footprint in credit institutions’ business lending.

In the method selected in this proposal for the experimental calculation of the carbon footprint of the Spanish credit institutions’ portfolio of business loans, elements that facilitate extrapolating this experience, both to other geographical areas (countries or jurisdictions) and to levels of aggregation (individual institutions or complete sectors), have been selected to make it easier to compare the results.
this end, easily accessible information sources and reproducible calculation methods have been used. However, this paper also flags the limitations of the methodology selected and suggests potential lines of improvement for refining the results obtained.

The carbon footprint index of Spanish credit institutions’ loans shows some improvement in the pollution intensity of the financing granted by banks in recent years. However, this is mainly a consequence of the reduction in the productive sectors’ polluting emissions, while the influence of the shift in the composition of the credit portfolio towards more sustainable activities has been small.

The experimental statistics obtained for measuring the carbon footprint of business loans by credit institutions in Spain and the methodological reflections presented in this paper seek to contribute to the current debate on the selection of indicators for measuring the financial sector’s impact on environmental targets and to be used for assessing climate change risks. They also have the potential to be extended to other instruments (for example, securities and household loans) and segments of the financial industry (such as investment funds and insurance companies).

Sources

AN ESTIMATION OF THE CARBON FOOTPRINT IN THE SPANISH CREDIT INSTITUTIONS’ PORTFOLIO OF LOANS TO FIRMS

Luis Ángel Maza

11TH BIENNIAL IFC CONFERENCE

25-26 August 2022

STATISTICS DEPARTMENT
OUTLINE

1. Project objectives and basic characteristics
2. Information sources and methodology
3. Main results
4. Limitations and possible improvements
Goals

- This work is an experimental statistics for the quantification of the carbon footprint of the portfolio of loans to resident firms held by Spanish credit institutions.
- There is no internationally agreed methodology: only autonomous experiences and initiatives
- This first contribution tries to be an input to the debate on the potential indicators to be used in the climate change strategy and for the measurement of the carbon footprint of the financial sector.
- It would make it possible to assess the transition risks linked to the change in the production model (decarbonization of the economy) in the credit exposures of financial institutions

Basic features

- Information sources used contain aggregated data (not at firm or loan level)
- Important assumptions and the incorporation of simplifying hypotheses
- Potential improvements (extensions to other instruments and sectors) and enrichments (methodological and access to new data)
Data sources

- Environmental accounts. Emissions into the atmosphere by branches of activity
  - Indicator: Total greenhouse gases (thousands of tons of CO2 equivalent)
  - Frequency: annual
  - Source: INE (last data: 2020)

- Annual national accounts of Spain: aggregates by branches of activity: Level of production (euro million)
  - Indicator: Production level (millions of euros)
  - Frequency: annual
  - Source: INE (last data: 2019)

- Annual National Accounts of Spain: Input-Output tables
  - Indicator: coefficients of the total inverse matrix
  - Source: INE (last data: 2016)

- Statement FINREP with breakdown according to economic activity (NACE sections)
  - For additional details (2-digit branches) in the manufacturing, mining and telecommunications sectors, the Central Credit Register (CCR) data are used.
  - Indicator: balance amount of loans (euro millions)
  - Frequency: quarterly
  - Source: Bank of Spain (last data: June 2021)

- Annual National Accounts of Spain: Input-Output tables
  - Indicator: coefficients of the total inverse matrix
  - Source: INE (last data: 2016)

- Environmental accounts. Emissions into the atmosphere by branches of activity
  - Indicator: Total greenhouse gases (thousands of tons of CO2 equivalent)
  - Frequency: annual
  - Source: INE (last data: 2020)

- Annual national accounts of Spain: aggregates by branches of activity: Level of production (euro million)
  - Indicator: Production level (millions of euros)
  - Frequency: annual
  - Source: INE (last data: 2019)

- Annual National Accounts of Spain: Input-Output tables
  - Indicator: coefficients of the total inverse matrix
  - Source: INE (last data: 2016)

- Statement FINREP with breakdown according to economic activity (NACE sections)
  - For additional details (2-digit branches) in the manufacturing, mining and telecommunications sectors, the Central Credit Register (CCR) data are used.
  - Indicator: balance amount of loans (euro millions)
  - Frequency: quarterly
  - Source: Bank of Spain (last data: June 2021)

- Environmental accounts. Emissions into the atmosphere by branches of activity
  - Indicator: Total greenhouse gases (thousands of tons of CO2 equivalent)
  - Frequency: annual
  - Source: INE (last data: 2020)

- Annual national accounts of Spain: aggregates by branches of activity: Level of production (euro million)
  - Indicator: Production level (millions of euros)
  - Frequency: annual
  - Source: INE (last data: 2019)

- Annual National Accounts of Spain: Input-Output tables
  - Indicator: coefficients of the total inverse matrix
  - Source: INE (last data: 2016)

- Statement FINREP with breakdown according to economic activity (NACE sections)
  - For additional details (2-digit branches) in the manufacturing, mining and telecommunications sectors, the Central Credit Register (CCR) data are used.
  - Indicator: balance amount of loans (euro millions)
  - Frequency: quarterly
  - Source: Bank of Spain (last data: June 2021)

The information-sets contain **data aggregated by branches of activity**

In general, we work with the detail of **64 branches** of activity according to NACE 2009 (when there is no perfect correspondence, we proceed to homogenization, merging or estimating details that allow us to work with greater granularity)

The information on **loans to companies** corresponds to the amounts of the principal drawn down (balances) that are classified according to the **NACE** of the companies
Methodology (I)

1. **CO2 emission coefficients** per unit of production: try to identify the intensity of emissions (direct and indirect) of carbon dioxide in the production of an economic branch

1. **Direct coefficients**: ratio between emissions and production

   \[
   q_{\text{direct}}^{it} = \frac{\text{Greenhouse Gas Emission}_it}{\text{Total production}_it}
   \]

   *For each branch of activity (i) and year (t)*

2. **Total coefficients**: add to the direct coefficients the indirect effect of the emissions produced in obtaining the intermediate inputs used by each industry

   \[
   q_{\text{total}}^{it} = (I - A)^{-1} q_{\text{direct}}^{it}
   \]

   *A: is the matrix of coefficients of the input-output table of the Annual National Accounts of Spain corresponding to 2016*

   *(I – A)\(^{-1}\): is the inverse Leontief matrix*
In the calculation of the **total coefficients** (incorporating the indirect effects), the values of the coefficients will depend on the **characteristics inherent to the productive structure** of each branch of activity. These coefficients **vary over time**, although the modifications are of small magnitude and the changes in the relative positions between branches are infrequent.

The branch of **electricity and gas supply** is the industry with the **largest indirect component**, derived from the effect of the emissions of the inputs used in its production process.

In recent years there has been a decrease in the intensity of emissions in all the branches of activity analysed, especially in the **most polluting ones** (fall in the 90th and 75th percentiles).
Methodology (II)

2. **Indicator of the intensity of the carbon footprint in the portfolio of loans to companies (IHCO2P):** it is calculated as the average of the total coefficients by NACE based on the weight of each branch in the portfolio of bank loans.

\[
IHCO2P_t = \frac{\sum_i P_{it} q_{it}^{total}}{\sum_i P_{it}}
\]

*Pit*: Stock of loans from credit institutions at the end of the year by branch of activity (i) and year (t)

**Indicator Interpretation:**
- **Falls** → Improvement of the carbon footprint
- **Rises** → Represents a worsening

The IHCO2P represents the average of the polluting emission ratios carried out by productive activities that obtain bank financing in relation to the total loans granted by credit institutions in Spain.

This is an approach similar to the one used in the works of:
- “The Higher Carbon Intensity of Loans, the Higher Non-Performing Loan Ratio: The Case of China (2017)”
- the **International Monetary Fund** (IMF) in its dashboard on climate change, in the block of financial indicators (link)
Comparison with the economy’s total emissions intensity

Given the difficulty of interpreting the indicator in its original units (tons of CO2) and to facilitate its comparability, it is restated in terms of an index (base year: beginning of the 2008 series).

A comparison of the IHCO2P against the evolution of the intensity of emissions in the Spanish economy between 2008-2019 would show an accumulated reduction of a very similar magnitude.
RESULTS OBTAINED (II)

Classification of the productive branches based on the emission coefficient (average 2008-2020), in the absence of a standard, this rule is used:

- More polluting ≥ P50
- Less polluting < P50

A slight recomposition of the loan portfolio of Spanish credit institutions towards less polluting branches is identified

Structure of the loan portfolio of Spanish credit institutions based on the CO2 emissions of the financed productive activities

Analysis of the IHCO2P in individual data (at the credit institution level)

The carbon footprint of the loan portfolio has been calculated for each credit institution in 2020 and 2017 (around 160 institutions)

Through the representation of Kernel functions, a shift to the left of the density function is observed in 2020 compared to 2017

This information would indicate a general decrease in the carbon footprint in the loan portfolio of credit institutions in Spain in recent years
1. Incorporation in the calculation of the carbon footprint of the factors related to the productive and financial structure of the branches

\[ q_{\text{modified totals}}^{it} = q_{\text{totals}}^{it} \times \frac{\text{Production (net turnover)}_{it}}{\text{Total balance sheet}_{it}} \times \frac{\text{Total balance sheet}_{it}}{\text{Financial liabilities}_{it}} \]

Central Balance Sheet data: sample information and representative ratios from the BACH database (R41 and L ratios, respectively). Averages 2008-2019

This modification would make it possible to adjust the "true" influence of bank loans on the development of business activities in the calculation of emission coefficients.

The results would show a less pronounced decline, almost stability.
2. The current NACE has significant limitations in capturing detailed information on economic activities and polluting emissions
   An example is the non-differentiation in the electricity production branch of generation with renewable energies. It is known that the current international classifications of economic activities (ISIC rev 4) have important deficiencies. The climate issue has been incorporated into the discussion for future updates.

3. Treatment of loans to holding companies and headquarters: assignment of the economic activity of the subsidiary companies.

4. Barriers to the use of loan information according to economic activity and non-purpose
   The exercise of quantifying the footprint could be enriched if the purpose of the loans were included in the classification of the credit portfolio (for example, the acquisition of electric vehicles, energy saving investments) would allow a better measurement of the footprint of carbon.

4. Access to individual data on emissions and loans (at company or company group level) could refine this measurement
   This improvement in the estimation of the footprint of bank loans should distinguish between segments of companies with detailed and complete information versus other groups of companies for which it would be necessary to make estimates and imputations (individual or aggregate).

5. Extension to other financial instruments and sectors: Loans to households, securities and investment fund portfolio.
THANKS FOR YOUR ATTENTION
Measuring the emissions profile of green exchange-traded funds – initial finding and lessons for official statistics¹

Hendrik Christian Doll, Maurice Fehr, Gabriela Alves Werb and Ece Yalcin-Roder
Deutsche Bundesbank

¹ This presentation was prepared for the conference. The views expressed are those of the authors and do not necessarily reflect the views of the BIS, the IFC or the central banks and other institutions represented at the event.
Measuring the Emission Profile of Self-Proclaimed Sustainable Exchange-Traded Funds

Gabriela Alves Werb*, Hendrik Christian Doll**, Maurice Fehr**, Ece Yalcin-Roder***

Abstract

The number of exchange-traded funds (ETFs) that proclaim themselves as sustainable has proliferated in the past years and attracted substantial investor interest. Until new regulations enforce standardized and comparable classification criteria, it remains difficult to understand how self-proclaimed sustainable ETFs differ in terms of sustainability strategy and metrics. We investigate the construction of self-proclaimed sustainable ETFs and assess their environmental footprint through scope 1 greenhouse gas (GHG) emission intensities. We combine public information about fund assets from fund issuers' websites and ETF databases, fund-level textual information extracted from fund disclosure documents, and proprietary firm-level emission data. Our analyses rely on 178 self-proclaimed sustainable ETFs, covering the largest global issuers, and 38 reference ETFs, i.e., the non-sustainable conventional ETFs that issuers state as the benchmark reference.

We find that self-proclaimed sustainable ETFs have lower average emission intensities than their reference ETFs. Part of this reduction is driven by divesting from emission-intensive sectors. We find little evidence of best-in-class selection effects, i.e., of funds selecting firms that spearhead emission reductions in their sector. Our results suggest that investors may reduce the carbon footprint of their investments by investing in self-proclaimed sustainable ETFs, when compared to reference ETFs. However, investors looking to cover a broad market while rewarding the lowest emitters within a sector cannot generally do so by investing in self-proclaimed sustainable ETFs.

Keywords: Green finance, exchange-traded funds, sustainable finance, carbon emissions, GHG emissions, ESG investing, sustainable investing

JEL classification: F18, G00, G10, Q56

* Deutsche Bundesbank, Data Service Centre and Frankfurt University of Applied Sciences
** Deutsche Bundesbank, Sustainable Finance Data Hub
*** Deutsche Bundesbank, Data Service Centre

The authors would like to thank colleagues in Bundesbank’s Data Service Centre and Sustainable Finance Data Hub, as well as attendants of the eleventh biennial IFC Conference for their valuable suggestions and feedback. All views expressed in this report are personal views of the authors and do not necessarily reflect the views of Deutsche Bundesbank or the Eurosystem.
## Contents

1. Motivation........................................................................................................................................... 3
2. Literature............................................................................................................................................. 4
3. Data ....................................................................................................................................................... 5
4. Empirical Analysis............................................................................................................................. 9
5. Discussion ........................................................................................................................................ 15
6. Conclusion ....................................................................................................................................... 16
7. References........................................................................................................................................ 17
1. Motivation

In the past years, the volume of investment into passive exchange-traded funds (ETFs) has increased considerably, especially among private investors. In 2021, global ETFs assets hit $9 trillion invested, with record net inflows (Wursthorn, 2021). Compared to individual stocks, ETFs are attractive in that they enable investors to cover a broad, diversified market with low transaction costs, high liquidity, intra-day trading, and high tax efficiency (Antoniou et al., 2022; Hasbrouck, 2003).

Similarly, the past years have witnessed a steady increase in investor interest in sustainable investment. As a result, the number of ETFs that are marketed as “sustainable” or focused on environmental, social, and governance (ESG) aspects is rapidly rising (Aramonte and Zabai, 2021). Sustainable investment is usually referred to as ESG investment, underlining the multi-faceted nature of the issue.

However, measuring these multiple dimensions of sustainability for investment products remains difficult. Aiming to close this gap, the European Union (EU) Regulation 2020/852 introduces a taxonomy for sustainable activities and provides a framework for classifying investments as “sustainable”. In addition, the Sustainable Finance Disclosure Regulation (SFDR) requires issuers to classify funds in three progressively stricter categories: Article 6, Article 8, and Article 9. Article 6 covers funds that do not consider sustainability aspects. Article 8 requires funds to “promote” either environmental or social aspects, or both. In a further step, Article 9 requires funds to have sustainable investment as their objective.

Issuers can generally construct self-proclaimed sustainable ETFs in three ways or a mixture of them: negative, positive, and best-in-class selection (BaFin, 2021). In a “negative selection”, issuers may start from an established market index and remove stocks based on thresholds (e.g., for greenhouse gas (GHG) emissions) or other exclusion criteria (e.g., no firearms, alcohol, or tobacco) globally applied to all sectors. Alternatively, issuers can also include stocks from the ground up, passively, or actively based on pre-defined criteria, constituting a “positive selection”. Finally, issuers may apply sector-specific criteria or thresholds to account for different sector characteristics and dynamics, which represents a “best-in-class” selection.

Consequently, it is difficult for investors to glimpse through the different metrics and strategies involved in the composition of self-proclaimed sustainable ETFs, which compromises their ability to make informed decisions. Furthermore, anecdotal evidence suggests that self-proclaimed sustainable ETFs may still be “stuffed full of polluters and sin stocks”, invest in firms “not aligned with the goals of the Paris agreement”, and do not substantially differ from “traditional” funds (Barclays Research, 2020; The Economist, 2021; Time, 2021).

While new regulations are welcoming steps towards increasing transparency in sustainable investments, there is still surprisingly little research on how self-proclaimed sustainable ETFs differentiate themselves in terms of investment strategy and resulting sustainability performance. Shedding light into this topic is important for several stakeholders.

For investors, this information can support informed decisions and enable them to measure the suitability footprint of their investments. For firms, it can inform their future decisions towards sustainable initiatives and business models, as the composition strategy of self-proclaimed sustainable ETFs affects their cost of capital and incentives. Finally, for policymakers, it can help them to assess the impact of recent regulations and inform future policy decisions.

We contribute to closing this gap by investigating the composition of self-proclaimed sustainable ETFs and comparing them with their respective reference ETFs, i.e., the conventional ETFs that issuers state as the

---

1 Regulation (EU) 2019/2088
benchmark reference. We assess their sustainability metrics and analyse whether differences are related to a particular composition strategy – i.e., positive, negative, or best-in-class selection. Because of the current data gaps in measuring societal and governance impact, we focus our analyses on their environmental impact and measure differences in scope 1 greenhouse gas (GHG) emission intensities. Scope 1 refers to the GHG emissions that result directly from the production process of the company instead of e.g., its electricity consumption (scope 2) or activities along the value chain (scope 3). We measure GHG emission intensity in tons of CO₂-equivalent per million € of revenue.

We combine public information about fund assets from fund issuers’ websites and ETF databases, fund-level textual information extracted from fund disclosure documents, and proprietary firm-level emission data. We match the equity assets and emissions data using the International Securities Identification Number (ISIN) as a common identifier and by performing a string-matching-based record linkage. Our analyses rely on 178 self-proclaimed sustainable ETFs and their respective 38 reference ETFs, covering the largest global issuers.

Issuers classify most of the self-proclaimed sustainable ETFs in our sample as Article 8 under SFDR. Our analyses point to a large variance in emission intensities within self-proclaimed sustainable ETFs and between them and their reference ETFs. Furthermore, out of 101 self-proclaimed sustainable ETFs that have a reference ETF, five have higher average emission intensity (AEI) than their reference ETFs. The differences are statistically significant at a confidence level of 99%.

Our results suggest that self-proclaimed sustainable ETFs shift away from the two most emission-intensive sectors (energy and mining and quarrying) towards the two least emission-intensive sectors (finance and information technology). Therefore, differences in sector composition explain part of the difference in emission intensities between self-proclaimed sustainable ETFs and their reference ETFs. Within emission-intensive sectors, we find little evidence for a “best-in-class” selection strategy.

Our paper provides important insights into the composition of self-proclaimed sustainable ETFs and uncovers avenues for future research. The rest of this paper is structured as follows. In section 2, we briefly discuss related literature. Following in section 3, we describe the data used for this paper before proceeding with the empirical analysis in section 4. We discuss findings in section 5 and conclude in section 6.

2. Literature

Several studies in the marketing-finance interface literature highlight the interplay between sustainability and marketing claims. For instance, Luo and Bhattacharya (2009) find that the effect of sustainable engagement on firms’ financial market performance is more pronounced under higher advertising levels. In addition, Martin and Moser (2016) find in an experimental setting that investor reactions to firms’ efforts towards reducing carbon emissions are not solely based on the expected effect of these efforts on firms’ future cash flows. Instead, their reactions also depend on how firms frame these activities in their disclosures.

Consequently, firms have an incentive to maintain a veneer of sustainable engagement (Torres et al., 2012) and may overly emphasize it for marketing purposes, a phenomenon known as “greenwashing” (Laufer, 2003). Previous studies find evidence of firms and financial products exhibiting greenwashing behaviour while publicly claiming adherence to sustainability standards (Liang et al., 2021; Tucker, 2021).

Responding to the record inflows to self-proclaimed sustainable financial products (Kerber and Jessop, 2021), an emerging literature stream in finance investigates whether investors can “do well by doing good”. Most recent studies focus on understanding whether self-proclaimed sustainable financial products differ in terms of return and risk, and on the possible drivers and moderators of this relationship, with somewhat
inconclusive evidence. As Gillan et al. (2021) summarize it, there are still many conflicting hypotheses and results.

When looking at ETFs, multiple studies find no statistically significant difference between the risk-weighted performance of self-proclaimed sustainable and conventional ETFs (Lobato et al., 2021; Weston and Nnadi, 2021). Other studies examine this potential performance difference during economic downturns and find conflicting results (Folger-Laronde et al., 2022; Omura et al., 2021; Pavlova and de Boyrie, 2022).

Clements (2021) underscores that given the rapid increase in the number of self-proclaiming sustainable ETFs and the currently high effort involved in acquiring and standardizing information, it is not realistic for investors or even professional advisors to compare and distinguish them. The U.S. Securities Exchange Commission (SEC) has also recognized this problem and issued a series of recommendations for investors, issuers, and policymakers (SEC, 2020).

Understanding the degree of product differentiation within self-proclaimed sustainable ETFs and between them and conventional ETFs is crucial for informed investment decisions. Furthermore, obtaining reliable information on the environmental, social, and governance footprint of different sustainability investment strategies is equally relevant for investors and for policymakers. Anecdotal evidence in the popular press suggests that self-proclaimed sustainable ETFs continue to invest in emission-intensive industries and firms, whilst maintaining a substantial overlap with conventional ETFs (e.g., Mohr, 2022; The Economist, 2021; Time, 2021).

However, most of the existing literature focuses on the “doing well” part, that is, on differences in risk and return. Despite the topic’s relevance, there is surprisingly little research on how self-proclaimed sustainable ETFs differentiate themselves in terms of investment strategy and resulting sustainability performance, in comparison to conventional ETFs. Our study aims to contribute towards closing this research gap.

Closest to our study is the work of Reiser and Tucker (2019), who analyse hand-collected data of 31 actively and passively managed self-proclaimed sustainable ETFs and 7 conventional ETFs. The authors note that “the underlying variation across funds is largely opaque to consumers, who rely on the ESG acronym at their peril”.

3. Data

The lack of unified metrics and standardised reporting framework presents the first challenge for our study. Fund- and fund asset-level information is generally public, but dispersed across many sources and partly in unstructured, textual documents. Emissions data on firm-level mostly originates from public sustainability reports, but in practice, these data are restricted in usability due to their unstructured nature, dispersion across multiple sources, as well as non-standardised definitions and reporting frameworks.

Therefore, our study relies on data from multiple sources. We start by collecting decentralised public information about ETF assets provided by fund issuers. We scrape this information automatically from the websites of five fund issuers’ and one public ETF database. Our data encompass the largest issuers worldwide, based on issuers’ total assets under management (AUM).

We collect a list of ETFs per issuer, including the assets for each fund, their weights, and respective ISINs (for equity assets). For the sample of self-proclaimed sustainable ETFs, we retain all ETFs that our data sources classify as “ESG”, “SRI”, or “sustainable”. In the cases in which the data source provided no such automatic filtering or labels, we search for a list of terms in the ETF titles that typically refer to sustainability: clean,
sustainable, green, wind, solar, renewable, climate, water, decarbonization, environment, cleantech, low carbon, planet, and forestry. Furthermore, we only consider equity ETFs for our analyses. Fund assets typically include equity (firms' stocks) and cash. Only few occurrences of forwards, futures, swaps, and contracts for difference in our sample of equity funds.

To compare the emission intensity between self-proclaimed sustainable ETFs and their respective non-sustainable reference ETFs, we create two groups of ETFs. The first is the self-proclaimed sustainable ETFs’ group, which contains ETFs screened for sustainability considerations. For each of these ETFs, we also retrieve similar data for their “non-sustainable” reference version, whenever one exists. We identify reference indices as they are provided in the self-proclaimed sustainable ETF’s fund documents. We retrieve ETF data in March 2022 with updates in June 2022. The combined data set contains 178 self-proclaimed sustainable ETFs and 38 reference ETFs. Out of the self-proclaimed sustainable ETFs, 101 ETFs have a matching (non-sustainable) reference ETF.

An example of a self-proclaimed sustainable ETF with a reference index would be one based on the “MSCI World”, e.g., a hypothetical “MSCI World SRI”. In such cases, issuers typically create the self-proclaimed sustainable ETFs funds starting from the conventional index (“MSCI World” in this example) and then excluding assets by applying global or sector-based thresholds based on various ESG criteria. Multiple self-proclaimed sustainable ETFs can have the same reference index – in this example, many self-proclaimed sustainable ETFs may be based on “MSCI World”. Self-proclaimed sustainable ETFs without a reference ETF are typically topic-based, e.g., a hypothetical “Global Clean Water Index”. In such a case, there is typically no conventional non-sustainable reference ETF to match it.

Then, we complement these data with information provided by fund issuers in the ETFs’ prospects in textual form. This source provides information on the investment goal and fund strategy stated by the issuers, as well as on the stated classification according to the SFDR, whenever applicable. We manually categorize this textual information into structured data for further analyses.

<table>
<thead>
<tr>
<th>Exemplary and simplified data structure</th>
<th>Table 1</th>
</tr>
</thead>
<tbody>
<tr>
<td>Information is available on fund-level and asset-level.</td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>ETF</th>
<th>Inception date</th>
<th>Self-proclaimed sustainable</th>
<th>Assets</th>
<th>Weight</th>
<th>Emission intensity (CO2-eq/revenue)</th>
</tr>
</thead>
<tbody>
<tr>
<td>A</td>
<td>2020-01-01</td>
<td>1</td>
<td>Firm 1</td>
<td>60%</td>
<td>0.5</td>
</tr>
<tr>
<td>A</td>
<td>2020-01-01</td>
<td>1</td>
<td>Firm 2</td>
<td>40%</td>
<td>0.4</td>
</tr>
<tr>
<td>B</td>
<td>2017-01-01</td>
<td>0</td>
<td>Firm 1</td>
<td>33%</td>
<td>0.5</td>
</tr>
<tr>
<td>B</td>
<td>2017-01-01</td>
<td>0</td>
<td>Firm 2</td>
<td>33%</td>
<td>0.4</td>
</tr>
<tr>
<td>B</td>
<td>2017-01-01</td>
<td>0</td>
<td>Firm 3</td>
<td>34%</td>
<td>0.8</td>
</tr>
</tbody>
</table>

Note: This table is simplified and made-up for illustrative purposes containing no real-world data.

Finally, we combine the asset-level information from equity assets with proprietary emissions data on firm-level greenhouse gas emissions. This information is based on firms’ non-financial reports and third-party estimations and are attributed to a single ISIN. We obtain scope 1 GHG emission intensity and the NACE sector of the firms from ISS ESG, a proprietary data provider.
The information available from issuers' websites contains different sets of variables and are presented in different formats. We standardize their formats and measuring units. Furthermore, we use a string-matching record linkage procedure for issuers that provide only non-unique asset names but no ISINs. We rely on observations from other issuers for which we have both the asset names and ISINs to fill the missing ISINs in all cases where we have an exact name match, after standardizing the asset names.

Ultimately, we obtain a cross-sectional dataset with 216 ETFs after joining the combined data sets with fund- and asset-level variables with the data set containing emission intensities and the NACE sector classification at the asset level. Fund-level variables include the fund's inception date, the number of assets, and assets under management. Asset-level variables include firms' names and identifiers, asset weights, firms' economic sectors, and scope 1 and 2 GHG emission intensities. Table 1 provides an example of our data structure. Within our sample of 216 ETFs, we compare 101 self-proclaimed sustainable ETFs with 38 reference ETFs. The 101 self-proclaimed sustainable ETFs contain 38,876 assets with 6,474 unique firms. The 38 reference ETFs contain 20,227 assets that represent 6,977 unique firms. Furthermore, in our sample, we have 77 self-proclaimed sustainable ETFs without a reference index, which cover a total of 20,959 assets and 8,185 unique firms. In our sample 97.5% of all assets are equity. We exclude all non-equity assets from our sample. Table 2 presents selected summary statistics.

<table>
<thead>
<tr>
<th>Indicator</th>
<th>Fund age in days</th>
<th>Number of assets</th>
<th>Average Scope 1 GHG* Emission intensity</th>
<th>Average Scope 2 GHG Emission intensity</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>all</td>
<td>sps**</td>
<td>ref**</td>
<td>all</td>
</tr>
<tr>
<td>Minimum</td>
<td>150</td>
<td>150</td>
<td>444</td>
<td>12</td>
</tr>
<tr>
<td>1st Quartile</td>
<td>443</td>
<td>398.2</td>
<td>2341</td>
<td>62.8</td>
</tr>
<tr>
<td>Median</td>
<td>905.5</td>
<td>774.5</td>
<td>3448</td>
<td>163.5</td>
</tr>
<tr>
<td>Mean</td>
<td>1728</td>
<td>1356.5</td>
<td>3471</td>
<td>364.1</td>
</tr>
<tr>
<td>3rd Quartile</td>
<td>2304.5</td>
<td>1468.2</td>
<td>4758</td>
<td>380</td>
</tr>
<tr>
<td>Maximum</td>
<td>7657</td>
<td>6398</td>
<td>7657</td>
<td>5205</td>
</tr>
<tr>
<td>Number of Missing Values</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
</tbody>
</table>

* Greenhouse gas.

** sps: self-proclaimed sustainable ETFs, ref: reference ETFs

Our data suggests that the number of assets held per fund is smaller in self-proclaimed sustainable ETFs than in reference ETFs – self-proclaimed sustainable ETFs have an average of 330 assets, compared to 525 assets in reference funds. We also observe that average weighted scope GHG emission intensities are smaller for self-proclaimed sustainable ETFs than for their reference ETFs (123.7 vs. 174.0 tCo2eq/rev). Considering average weighted scope 2 emission intensities, the difference almost disappears (41.2 vs. 44.7 tCo2eq/rev).

The average age of self-proclaimed sustainable ETFs is 3.7 years, whereas reference ETFs are on average 9.5 years old. Figure 1 shows the inception dates of self-proclaimed sustainable ETFs and their reference
ETFs. In line with previous studies, we also observe a steady increase in the inception of self-proclaimed sustainable ETFs in recent years.

Using the textual information from the fund documents, we analyse the stated funds’ strategy, ESG criteria considered, and reported data sources for ESG information. From 178 self-proclaimed sustainable ETFs in our sample, 163 are managed passively and 15 ETFs are managed actively. All those passively managed track a sustainable index. Hence, when talking about an ETF’s impact in designing a sustainable investment opportunity, in practice this choice is closely reflects the index provider’s asset choices.

By far, MSCI is the most often cited index provider in our sample (or provider of sustainability-related information). It appears in fund documents with 101 occurrences, followed by S&P (21), and Deutsche Boerse AG (10). We note that these mentions can add up to more than the number of self-proclaimed sustainable ETFs, because some funds report multiple data sources for ESG information. 122 ETFs in our sample consider investments in a broad range of sectors, whereas 56 funds target one or few specific industry sectors, such as the hypothetical “Global Clean Water Index” in our previous example.

![Figure 1: The inception of self-proclaimed sustainable ETFs and reference ETFs in our sample over time](image)

The self-proclaimed sustainable ETFs or the indices they are based on often report following a negative-selection approach with 117 mentions, followed by a positive (33), best-in-class (21), and active selection (14). Only one ETF in our sample states that it is based on an index that explicitly does not exclude firms, but reweights the assets based on ESG information. Similarly, we note that these mentions can add up to more than the number of self-proclaimed sustainable ETFs, because some funds report a combination of selection methods. In particular, 9 self-proclaimed sustainable ETFs report a combination of global thresholds (negative selection) and sector-specific thresholds (best-in-class).

As self-proclaimed sustainable ETFs can target a combination of ESG aspects or can incorporate only one or two aspects, we regard their stated sustainability criteria. We see that 169 mention the general “ESG” goal as a criterion in their composition. In addition, 34 mention low carbon emission as a criterion, and 5
mention “social” aspects. Mentions can add up to more than the number of self-proclaimed sustainable ETFs, because some report a combination of criteria. Specifically, 29 self-proclaimed sustainable ETFs report a general combination of ESG aspects and low-carbon emissions criteria.

Where this information is available, 81 self-proclaimed sustainable ETFs report themselves as adhering to SFDR Article 8 and 14 to Article 9. This information is missing for a considerable fraction of our sample, as it is typically only available for European targeted ETFs.

The summary statistics indicate that, on average, self-proclaimed sustainable ETFs have smaller emission intensities than their conventional counterparts. In our empirical analysis, we investigate how this reduction in emission-intensities may relate to different possible asset selection strategies.

4. Empirical Analysis

In a first step of analysis, our intention is to discover patterns in our dataset, whether investors could differentiate between self-proclaimed sustainable ETFs and reference ETFs, short of reading the fund documents. Therefore, we try to mimic investor decisions based on available information by classifying ETFs using k-Means clustering. This can be seen as a continuation of Reiser and Tucker (2019), who noted that investors have to rely on the ESG label of the ETF. We are interested in how this holds up following legislative changes to standardise definitions and disclosures in the meantime.

We prepare a dataset without any labels about self-proclaimed sustainable ETFs and reference ETFs and let the algorithm classify them into two different groups. We choose the Hartigan-Wong k-Means clustering algorithm (Hartigan and Wong, 1979), which is a popular unsupervised machine learning method. The algorithm requires the pre-defined number of clusters and assigns each data point to a cluster. We choose the number of clusters as two. As this is an unsupervised method, the clusters are not automatically assigned to a specific group. After finding the highest performance model, our aim is to identify and explore which variables are the most informative and which variables should we take a deeper look at our following analysis.

We base this analysis on available sector and emission information, as well as fund-level data. In order to classify funds, we consider the following available fund-level variables: fee, age of the fund, total assets under management and number of assets.

As features providing information on funds’ emissions, we consider the average GHG emission intensities (scope 1 and scope 2) for each fund. We compute these by aggregating the weighted average of the scope 1 (scope 2) GHG emission intensity at the asset-level for a given ETF. We do so by calculating the weighted sum over all N equity assets in an ETF for which we have emission intensity information available:

\[ AEI_{ETF} = \sum_{i=1}^{N} \frac{E_i}{R_i} w_{ETF}, \quad i \in 1, \ldots, N \]

Where \( AEI_{ETF} \) represents the weighted average emission intensity (AEI) of a given ETF, \( E_i \) asset i’s total scope 1 GHG emission, \( R_i \) asset i’s (firm) revenue, and \( w_{ETF} \) the asset i’s weight in the ETF composition. AEI’s unit is:

\[ \frac{Total \ Scope \ 1 \ emissions \ in \ Tons \ of \ CO_2 \ equivalents}{Revenue \ in \ million \ EURO} \]

Furthermore, we consider the maximum emission intensity of a fund’s asset:

\[ \max_i (E_i_{ETF}) = \frac{E_i}{R_i}. \]
As proxies for the sector composition of the fund we compute the sector weights (SW), i.e. the % of a funds holding for each sector \( j \), where \( j \) is the NACE2 sector on level 1 ("letter level"):

\[
SW_{\text{ETF}} = \sum_{j=1}^{S} \sum_{i=1}^{N} I_{w_i}(j)w_i , \quad i \in 1, \ldots, N, \quad j \in 1, \ldots, S
\]

Where \( I \) is the indicator function, 1 if asset \( I \) belongs to sector \( j \), 0 if not, \( S \) is the number of NACE2 sectors (on the highest level 1, 21 NACE2 sectors exist). Furthermore, we compute the number of a fund’s asset in each sector \( j \), which is similar to SW, without considering the weights \( w \).

As a proxy, for “best-in-classness” BC, we compute the average differences in emission intensity between all assets of a given ETF in a given sector as opposed to the sector average emission intensities:

\[
BC_{\text{ETF}} = \sum_{j=1}^{S} \sum_{i=1}^{N} I_{w_i}(j) \left( \frac{E_i}{R_i} - \text{avg}(EI)_{(j)} \right) , \quad i \in 1, \ldots, N, \quad j \in 1, \ldots, S
\]

As we obtain a relatively large numbers of features for the sector-based variables, we also compute principal components in order to reduce the dimensions.

In this way, we have the features required for our model. To find the model gives us the highest predictive power, we try these features in a variety of combinations.

![kMeans clustering results](image)

**Figure 2: Resulting clusters of self-proclaimed sustainable and reference ETFs using kMeans.**

*Note: In the lower left corner, the green dots and grey x’s form one cluster, the grey dots and green x’s form the other cluster.*

The results somewhat surprisingly show that the most informative variables are the maximum scope 1 GHG emission intensities (not weighted) and the ETFs’ age. This means that the best predictive power is obtained using the highest emitter’s emission intensity in any given ETF and not the average emission intensity. Contrary to expectations, sector composition does not seem to add much predictive value (nor do
its principal components). Neither do emission intensities compared to sector average emission intensities play a significant role.

Figure 2 displays the ETF’s distribution across the two most informative variables “maximum scope 1 GHG emission intensity in fund” and “fund age”. It shows how the 216 ETFs are assigned to two separate categories, which we later label as sustainable and non-sustainable, based on the majority of the true available labels. Based on these labels, we assess whether the algorithm classified each ETF in the correct category. We see two clusters roughly emerging. The first one consists of green dots and grey x’s in the lower left corner, and the second cluster consists of grey dots and green x’s. The green and grey dots (“x” marks) represent the correctly (incorrectly) predicted self-proclaimed sustainable ETF and reference ETF, respectively.

<table>
<thead>
<tr>
<th>Confusion Matrix</th>
<th>Table 3</th>
</tr>
</thead>
<tbody>
<tr>
<td>Predictions of the clustering algorithm vs. self-descriptions of ETFs</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Self-proclaimed sustainable ETF</td>
</tr>
<tr>
<td>Predicted sustainable ETF</td>
<td>153</td>
</tr>
<tr>
<td>Predicted reference ETF</td>
<td>25</td>
</tr>
</tbody>
</table>

Source: Own calculations based on ETF data and ISS ESG

To measure the performance of the algorithm, we use a confusion matrix. Table 3 shows that, out of 178 self-proclaimed sustainable ETF, 153 are predicted correctly and 25 are predicted incorrectly. From the confusion matrix, we can also calculate the model’s precision and recall. In our analysis, precision is the share of self-proclaimed sustainable ETF that are correctly classified out of all self-proclaimed sustainable ETFs. Our model classifies 86% of self-proclaimed ETFs correctly. On the other hand, the recall of 93% provides a measure how our model’s ability to identify self-proclaimed ETFs. As we have highly imbalanced classes, a combination of these two metrics, the F1-score, provides a better assessment of our model performance. A F1-score of 89% indicates the harmonic mean of the precision and recall.

The cluster classification performs relatively well; even though it has some limitations, such as the sensitivity of clusters on the initial seed, an overrepresented majority class issue, and the lack of additional sustainability variables on asset-level and firm characteristics that could improve future analyses. Following this clustering from an investor’s perspective, we proceed to analyze the environmental footprint.

We compare the AEI of the self-proclaimed sustainable ETFs with their non-sustainable reference ETFs and calculate the respective differences in emission intensity for each pair:

\[ AEI_{\text{difference}} = AEI_{\text{self-proclaimed}} - AEI_{\text{reference ETF}} \]

Out of the 101 self-proclaimed sustainable ETFs that have a reference ETF, we identify 5 ETFs that have higher average emission intensity than their respective reference ETFs. For the remaining 96 self-proclaimed sustainable ETFs, we find their AEI to be lower than their reference ETFs. We test our results with a paired t-test and the differences in AEI are statistically significant (p < .01). However, we should interpret these results with care, as groups in our setting are not independent and reference group observations repeat themselves, because many self-proclaimed sustainable ETFs have the same reference ETFs. Consequently, degrees of freedom are not obvious to compute, and we cannot rule out the presence of heteroscedasticity. Nevertheless, we note that other test settings lead to the same substantive conclusions. Figure 3 shows the distribution of AEI for self-proclaimed ETFs and the reference ETFs, as well as the paired differences.
To shed light into the drivers of these differences, we examine further the sector composition across the pairs. In particular, we are interested in understanding the reductions in AEI – do they originate from a best-in-class selection strategy or from simply delisting emission-intensive sectors?

If a best-in-class selection takes place, we would expect the sector composition to remain relatively stable in a self-proclaimed sustainable ETF when compared to its reference ETF. Under this scenario, self-proclaimed sustainable ETFs would feature assets of firms with lower greenhouse gas emission intensities within emission-intensive sectors in a comparison against peers in their own industry, even if their absolute emission intensities remain high compared to other sectors. Instead, if negative selection criteria are more predominant, we would expect funds to remove assets of firms in emission-intensive sectors, as they would tend to remove those beyond a certain ESG threshold in a comparison against everyone. Under this approach, funds would tend to end up with higher weights in less emission-intensive sectors.
For this analysis, we examine the pairs of 101 self-proclaimed sustainable ETFs and their respective 38 reference ETFs using the NACE classification system. The NACE Level 1 Codes includes 21 sectors identified by the letters A to U. We compute the sector shares in each ETF based on the sum of weights of the firms that belong to the respective sector. Figure 4 shows that the most dominant sector in our sample is manufacturing (NACE C), followed by financial and insurance activities (NACE K), and information and communication (NACE J). Even though the rank of most dominant sectors is the same for both self-proclaimed sustainable ETFs and reference ETFs, self-proclaimed sustainable ETFs contain substantially more assets in the sector information and communication (NACE J).

![Average difference in sector shares](image)

Source: Own calculations based on ETF data and ISS ESG

Figure 5: Average differences in sector shares between self-proclaimed sustainable ETFs

Then, we further compare the percentage differences in sector composition with their original starting point in the reference ETFs. The results in Figure 5 suggest that self-proclaimed sustainable ETFs shift away from the two most emission-intensive sectors, electricity (NACE D) and mining and quarrying (NACE B), towards investing more into financial and insurance activities (NACE K) and information and communication (NACE J), which are among the least scope 1 GHG emission-intensive sectors. Even though the reductions in percentage points seem small – e.g., 2 percentage points less in mining and quarrying – these amount to a substantial proportion of the original sector shares in the reference ETFs. Overall, we find that differences in sector composition partly explain the differences in scope 1 GHG emission intensity between self-proclaimed sustainable ETFs and their reference ETFs.

To further shed light into the sources of the observed differences in emission intensities, we proceed to assessing whether assets in the self-proclaimed sustainable ETFs have substantially lower emission intensities than other assets in their respective sectors. Our aim is to assess whether there is any evidence for a “best-in-class” asset selection between self-proclaimed sustainable ETFs and reference ETFs.

To do so, we first take the universe of assets in the ISS data, which are based on emissions of 29,264 firms and calculate the average emission intensity for each sector, weighted by the firm’s revenue. Figure 6 shows that the most emission intensive sectors are electricity (NACE D) and mining and quarrying (NACE B). In comparison to these sectors, manufacturing (NACE C) has a moderate average emission intensity, whereas finance and information technology (NACE K) and information and communication (NACE J) have a substantially lower average emission intensity.
Then, for each asset in an ETF, we calculate the difference between its emission intensity and the average emission intensity in its respective sector, as calculated from the ISS universe. Then, we aggregate these differences at the ETF level by weighting them using the asset weights in each ETF's composition. If self-proclaimed sustainable ETFs predominantly followed a best-in-class asset selection strategy, then we would expect their aggregated differences from sector averages to be predominantly negative, as the assets they select would systematically have lower scope 1 GHG emission intensities than their sector peers.

The results in Figure 7 indicate that even though the deviations from sector averages in self-proclaimed sustainable ETFs are predominantly negative, their distribution has a substantial overlap with the distribution for their respective reference ETFs. Their selected assets do not seem to be substantially better than their peers in the corresponding sectors, compared to the assets present in the reference ETFs. Furthermore, for a few self-proclaimed sustainable ETFs, we find positive deviations from sector averages, which suggests that they are investing in assets of firms that are, on average, more emission-intensive than their sector peers. Taken together, these results provide little evidence corroborating a “best-in-class” asset selection in the self-proclaimed sustainable ETFs in our sample. We replicate these analyses considering scope 2 emission intensities and combined scope 1 and scope 2 GHG emission intensities and find similar substantive results. In the next section, we discuss these results’ main insights and implications.
5. Discussion

When simulating the typical information set available to investors to differentiate between self-proclaimed sustainable ETFs and their respective non-sustainable reference ETFs using k-Means clustering, we find that the most informative variables are the maximum (non-weighted) scope 1 GHG emission intensities and the ETF’s age. Interestingly, neither the sector compositions, nor their differences or average emission intensities have much value in distinguishing self-proclaimed sustainable ETFs from non-sustainable reference ETFs. These results suggest that many self-proclaimed sustainable ETFs follow a global approach with a negative asset selection – i.e., removing the most emission-intensive assets. This finding substantiates claims in self-proclaimed sustainable ETFs’ prospects, which predominantly report following a negative-selection approach, with a total of 117 mentions.

Out of the 101 self-proclaimed sustainable ETFs that have a non-sustainable reference ETF, 96 have lower average emission intensities than their reference ETFs. These differences are statistically significant (p<.01). Out of the 5 identified ETFs that have a higher average emission intensity (AEI) than their reference ETF, none focuses specifically on carbon reductions based on their fund documents. Indeed, in the textual analyses of the ETFs’ prospects, we find that most of them mention general “ESG goals” as their composition...
criterion. Maybe noteworthy could be that three out of these five are focused on specific sectors (energy, communication, and consumer goods respectively).

Manufacturing (NACE C) is the most dominant sector in the ETFs represented in our paired sample of self-proclaimed sustainable ETFs and their reference ETFs, followed by finance (NACE K), and information technology (NACE J). The most scope 1 GHG emission-intensive sectors are energy (NACE D) and mining and quarrying (NACE B). Manufacturing has moderate scope 1 GHG emission intensity, while finance and information technology produce very little scope 1 GHG emissions. Our results suggest that self-proclaimed sustainable ETFs shift away from the two most emission-intensive sectors towards the two least emission-intensive sectors.

We further analyze the deviations between the emission intensity of each asset in an ETF and the average emission intensity in its respective sector, as calculated from the ISS universe, and aggregate them for each ETF using the assets’ weights. Our results provide little evidence that self-proclaimed sustainable ETFs follow a “best-in-class” asset selection. This finding only partly substantiates the information reported by fund issuers in their fund prospects, as the textual data indicates that 21 self-proclaimed sustainable ETFs claim to follow a “best-in-class” approach, whereas nine report a combination of global thresholds (negative selection) and sector-specific thresholds (best-in-class).

Our findings corroborate the initial results provided by Reiser and Tucker (2019) – they also find that investment strategies vary widely among self-proclaimed sustainable ETFs. Furthermore, they also find that self-proclaimed sustainable ETFs’ assets substantially overlap with those of “conventional” ETFs. Because the degree of a firm’s sustainable activities depends on the sector in which it operates (Banerjee, Iyer, and Kashyap, 2003), self-proclaimed sustainable ETFs should exhibit different sustainability performance, depending on their sector composition. Indeed, our results suggest that differences in sector composition explain a great proportion of the differences in emission intensities.

Interestingly, Folqué et al. (2021) find that funds that only apply negative selection obtain worse ESG risk scores and worse carbon risk on average. The prevalence in negative selection in our study, which is stated in the fund documents and substantiated by our empirical analyses, opens interesting further roads for future studies.

6. Conclusion

In the past years, investor interest in sustainable investments has increased substantially. As a result, the number of ETFs that are marketed as “sustainable” also rose to record levels. However, despite recent regulatory steps to increase the transparency in this market, there is no coherent set of international regulations as to what defines a sustainable financial product. It is therefore challenging for multiple stakeholders to discern from the different metrics and strategies involved in the composition of self-proclaimed sustainable ETFs, which compromises their ability to make informed decisions.

We contribute to closing the research gap in this area by investigating the composition of 101 self-proclaimed sustainable ETFs and outlining the differences in comparison to their respective 38 reference ETFs. We further investigate whether differences arise from a particular composition strategy – e.g., positive, negative, or best-in-class selection – and assess whether the extent to which they lead to reductions in scope 1 GHG emission intensities.

We find that self-proclaimed sustainable ETFs generally have lower average emission intensities than their reference ETFs, but there is a high variance in the measured differences. We find them to be statistically significant at the confidence level of 99%. Part of this reduction is driven by divesting from emission-
intensive sectors, in particular, from energy and mining and quarrying. Self-proclaimed sustainable ETFs shift their assets to least emission-intensive sectors, such as finance and information technology.

We find little evidence of a best-in-class approach within emission-intensive sectors. Only 21 self-proclaimed sustainable ETFs explicitly mention this strategy in their as stated in their fund disclosure documents. Nevertheless, our data shows that little signs of a best-in-class selection effect on average for a sample. For future work, it seems valuable to investigate the differences in sector composition specifically in those ETFs that mention best-in-class as their selection criteria.

Our results indicate that investors may, on average, reduce GHG exposure by investing in self-proclaimed sustainable ETFs. However, investors looking to cover a broad market while rewarding the lowest emitters within a sector cannot generally do so by investing in self-proclaimed sustainable ETFs. If investors wish to incentivize the least emission intensive firms per sector (best-in-class) they currently do not have an easy way to identify such funds.

Investors who wish to support the transition to a less carbon-intensive economy through a specific strategy could make a deep analysis of the fund disclosure documents. They could assess whether self-proclaimed sustainable funds aspire to be ecological-, social-, or governance-oriented (or any combination of these) and assess their adherence to the claimed strategy with hard data.

7. References


Measuring Emissions Profiles of Self-Proclaimed ESG ETFs
Initial findings and lessons for official statistics

Gabriela Alves Werb, Hendrik Christian Doll, Maurice Fehr, Ece Yalcin-Roder

The authors would like to thank colleagues in the RDSC for their valuable suggestions and feedback. All views expressed in this report are personal views of the authors and do not necessarily reflect the views of Deutsche Bundesbank or the Eurosystem.
Understanding the investment strategy of self-proclaimed ESG\(^1\) ETFs\(^2\) is important for informed investing

**IMPORTANCE**
- Increasing interest in ESG investing
- Self-proclaimed “ESG” investments reached a market capitalization of $1.7 trillion in 2020 and continue to grow\(^1\)

**ISSUE**
- Increasing number of self-proclaimed “ESG” ETFs
- Difficult to measure how they differ in terms of sustainability strategy and metrics
- Lack of transparency, limited and scattered information available

**IMPACT**
- Critical information for informed investment decisions and for policy-making

---

Measuring emission profiles of self-proclaimed ESG ETFs
August 2022

---

1. ESG: Environmental, Social, Governance
2. ETF: Exchange Traded Fund
3. Jessop and Howcroft (2021)
Current evidence suggests incentives for ETFs to self-classify as “ESG”

For marketing purposes, funds have an incentive to maintain a veneer of ESG engagement.

Anecdotal evidence suggests that self-proclaimed ESG ETFs may still:

- be “stuffed full of polluters and sin stocks”
- invest in companies “not aligned with the goals of the Paris agreement”, and
- not substantially differ from “traditional” funds.

Findings in literature suggest that hedge funds exhibit greenwashing behavior while publically endorsing adherence to ESG standards.

Investors may not be able to adequately estimate the expected impact of ESG investments.

How do self-proclaimed “ESG” ETFs carry out their investment strategy?

I.e. do they choose “best-in class” assets per sector or do they divest from emission-intensive sectors?

Does their strategy lead to a consistently lower emission-intensity in their holdings?

Literature

1. Torres et. al (2012)
2. The Economist (2021)
3. Time (2021)
4. Barclays Research (2020)
5. Liang, Sun, and Teo (2021)
7. Martin and Moser (2016)
Data for this study comes from ETF issuers and proprietary emission data

**MATCH EXAMPLE**

MSCI World Index

- **“iShares MSCI World SRI UCITS ETF”**
- **“iShares Core MSCI World UCITS ETF”**

Top 3 Holdings

- Microsoft (4.5%)
- Tesla (4.4%)
- NVIDIA (4%)

**Top 3 Holdings**

- Apple (4.4%)
- Microsoft (3.5%)
- Amazon (1.2%)

**FINAL SAMPLE**

216 ETFs

Self-proclaimed ESG ETFs

- 101 ETFs
- 38,876 holdings
- 6,474 unique companies

Reference ETFs

- 216 ETFs
- 20,227 holdings
- 6,977 unique companies

**CHALLENGES**

- No central public data source for self-proclaimed ESG ETFs (webscraping from each fund issuer)
- Available information heterogeneous and in different formats

**EXAMPLE**

“Global X Clean Water ETF”

- ETFs constructed “ground up” with no reference index

1 Institutional Shareholder Services (ISS) ESG climate core package, data as of March 2022
2 The 38 reference ETFs serve as a benchmark for 101 self-proclaimed ESG ETFs. Reference ETFs can be identified whenever ESG ETFs are based on a large reference index (e.g., “MSCI Europe ESG Screened” is based on “MSCI Europe”). For the remainder of 77 self-proclaimed ESG ETFs, we cannot map a reference index, usually because these ETFs are constructed “ground up” and do not have a regular reference index.
To mimic investor information, we cluster ETFs based on sectors and emissions

**RESULTS**

- We cluster using k-Means and a range of features based on emissions, sectors, best-in-class proxies, and fund-level variables.
- Most informative features are maximum scope 1 emission intensities (not weighted) and the ETF’s age.
- Sector composition doesn’t seem to add much predictive value (neither single nor aggregated using PCA**).
- Cluster classification is decent, however findings have some limitations***.

---

Initial results

Measuring emission profiles of self-proclaimed ESG ETFs
August 2022

Page 5

- Included variables in this graphic are “maximum holding’s scope 1 emission intensity in fund” and “fund age”. We try a wide range of predicted features and feature combinations including sector weights, average emission differences in fund holdings compared to sector averages, and others (details in Annex).
- PCA: Principal Component Analysis
- **Limitations in the analysis include the variance of clusters depending on initial seed, an overrepresented majority class issue and more potential ESG variables on company-level to be considered.**
We compare emission intensities to further explore the drivers of incorrect cluster classifications.

**RESULTS**

- Out of 101 self-proclaimed ESG ETFs that have a reference ETF, we identify 5 ETFs that have higher average emission intensity (AEI)* than their non-ESG reference ETFs.
- For the remainder of 96 self-proclaimed ESG ETFs, we find their AEI to be lower than their reference ETFs.
- Differences in AEI are significant between self-proclaimed ESG ETFs and reference ETFs (p-value<.01)**

* Average emission intensity is the weighted sum of emission intensities of all holdings in a fund. Emission intensities are scope 1 CO₂ emissions as a fraction of company revenue.
** Challenges for a t-test in our setting are that groups are not independent, reference group observations repeat themselves, degrees of freedom are thus not obvious to compute, and there is heteroscedasticity present. We consider a paired t-test, however other test settings similarly let us reject the null hypothesis that AEI between groups follows the same distribution.

---

Measure of emission profiles of self-proclaimed ESG ETFs

August 2022

Page 6
An analysis of sector composition suggests that funds reduce emissions by relocating capital across sectors.

**SECTOR COMPOSITION**

- The most **dominant sector** in our sample is **manufacturing** (NACE C), followed by finance (NACE K) and information technology (NACE J).
- The **most emission intensive sectors** are **energy** (NACE D) and mining and quarrying (NACE B).
- Manufacturing has moderate emission intensity, while finance and information technology produce very little scope 1 emissions.

**Average share of most important sectors**

- n = 139, out of which 101 self-proclaimed ESG ETFs and 38 reference ETFs

**Average emission intensity of most important sectors**

- n = 29,264 companies in the ISS universe

Source:
- Own calculations based on ETF data and ISS ESG
- Deutsche Bundesbank
ESG ETFs seem to reduce investments in emission-intensive sectors

**SECTOR COMPOSITION**

- Self-proclaimed ESG ETFs seem to **shift away from the two most emission-intensive sectors** towards the two least emission-intensive sectors.

- Therefore, part of the **emission intensity difference** between self-proclaimed ESG-ETFs and their reference ETFs can be explained by differences in sector composition.

- Within emission-intensive sectors, there is **little evidence for a “best-in-class” asset selection**.

---

*Measuring emission profiles of self-proclaimed ESG ETFs
August 2022*
Our results suggest that self-proclaimed “ESG” ETFs reduce emissions through a “sustainable sectors strategy”

- **CARBON FOOTPRINTS**
  - Self-proclaimed ESG ETFs seem to have **lower average emission intensities** than their reference ETFs
  - Part of this reduction is driven by **divesting from emission-intensive sectors**
  - We find **little evidence of a best-in-class** (positive selection) approach

- **TAKE AWAYS**
  - **Investors** on average reduce carbon exposure by investing in self-proclaimed ESG ETFs
  - **Investors** looking to cover a **broad market**, while rewarding **lowest emitters within a sector**, cannot generally do so by investing in self-proclaimed ESG ETFs
  - **Policymakers** need to ensure better data availability and transparency

- **ROAD FORWARD**
  - **Standardization** of sustainability criteria, enhanced transparency and data availability is **underway on company-level.**
    - Standardization on fund-level is yet to come
  - Further analyses may focus on how positive and negative selection in self-proclaimed ESG ETFs affects companies’ cost of capital and incentives

---

* e.g. “EU taxonomy” (Regulation (EU) 2019/2088), Corporate Sustainability Reporting Directive (CSRD), Procedure (EU) 2021/0104/COD


Sentiment analysis of user’s reviews on non-bank payment service apps,¹

Muhammad Hafiruddin, Mohammad Khoyrul Hidayat, Arinda Dwi Okfantia and Nursidik Heru Praptono, Bank Indonesia

¹ This presentation was prepared for the conference. The views expressed are those of the authors and do not necessarily reflect the views of the BIS, the IFC or the central banks and other institutions represented at the event.
Sentiment Analysis of User's Reviews on Non-Bank Payment Service Apps

Nursidik Heru Praptono1,2, Arinda Dwi Okfantia, Mohammad Khoyrul Hidayat1, Muhammad Hafiruddin1

Abstract

The digital payment activities on non-bank payment service environment have shown to be increasing, especially during the COVID-19 pandemic. Any systemic risks on monetary and payment system stability affected by mobile apps quality should be anticipated as early as possible. We present an inference model to analyse the user’s review sentiments on mobile apps quality on some aspects, address some related inference problems, and construct an index based on the inferred sentiments. Having some promising results, we suggest that our approach can be used by policy makers in order to timely monitor the performance of non-bank payment service providers.

Keywords: Probabilistic Inference, Machine Learning, Limited Training Data, Non-Formal Text, Text Mining, Sentiment Analysis, User Reviews, Mobile Apps, Non-Bank Payment Service Provider, Monetary and Payment System Stability.

JEL classification: C44, E42

-The views expressed in this work belong to the authors and do not necessarily reflect the institution-

1 Statistics Department, Bank Indonesia.
2 nursidik_hp@bi.go.id/pra.heru@gmail.com (corresponding email)
Contents

1. Background.................................................................................................................. 3
2. Literature Review.......................................................................................................... 4
3. Methodology ................................................................................................................ 5
   3.1. Data Collection and Annotation ............................................................................ 5
   3.2. Data Preprocessing ................................................................................................ 7
   3.3. Inference Models .................................................................................................... 8
   3.4. Index Formulation .................................................................................................. 9
4. Result and Discussion .................................................................................................. 10
   4.1. Result on the Inference Models ............................................................................ 10
   4.2. Result on the Index Series .................................................................................... 12
5. Conclusion and Future Works ...................................................................................... 14
   5.1. Conclusion .............................................................................................................. 14
   5.2. Future Works .......................................................................................................... 15
References........................................................................................................................ 15
1. Background

More banking and payment activities are currently being conducted into mobile application platform than the traditional ones. Such services can legally be provided by the authorised entities, either bank or non-bank institutions. A number of nonbank payment service providers are increasingly running these services. In Indonesia for example, by the end of 2021, there have been at least 41 non-bank payment service providers (PSP) who take a part in this area. Through mobile application platforms that they provide, an amount of services related to the payment activities are offered including for example e-wallet and e-money. The digital economy is thus accelerated through this landscape.

The growing amount of non-bank PSP customers using such services has also been found to be increasing recently, moreover since the pandemic of COVID-19. By the beginning of year 2022, at Google Play (for Android users), more than 100 million installs non-bank PSPs were reported. In addition to that, at AppStore (for Apple Iphone users) although the number of installs of this top non-bank PSPs is not explicitly reported, the number of reviews reached more than 800 thousand. Considering the whole 41 providers, the non-bank PSP’s have taken a part and infiltrated on nearly the whole of about 270 million people of the Indonesian population. This amount describes that the payment services provided by non-bank PSP are non-trivial thing and have a significant role in promoting and accelerating economic growth. They however require authorisation, supervision and monitoring from Bank Indonesia under the Payment System Policy.

This massive implementation of payment services provided by non-bank PSP is not without challenges. Many risks can occur if there is any problem on running such services. Of this phenomena we identify that there are some perspectives that need to concern about based on the entities involved (customers, non-bank PSPs, and policy makers):

1. Customers however need the reliable mobile application provided, it is intuitive that they would either oversee the review before install and/or give any reviews on the mobile apps’ performance.
2. Non-Bank PSPs have to secure their reputational issue so that their business strategy can run properly as planned. Any problem related to the reliability of their mobile apps would matter, since mobile apps take the closest proxy of their representation to the customers. Thus they need to be able to provide reliable payment services mobile apps for the customers.
3. Policy maker concerns about monetary stability and payment system security & efficiency. It is important for policy maker to oversee the performances of non-bank PSP’s since any problem on running payment services through mobile apps can potentially impact the stability in the macroeconomic perspective. This is to enable the policy maker to consider an early decision to the non-bank PSP’s in advance.

Measuring the quality of mobile apps of payment services therefore becomes important as it represents the quality of non-bank PSP. However, performing such measurement itself is not a straightforward way to do. A proxy for this measurement is by leveraging user review’s sentiment on the quality of mobile apps on some aspects. Analysing sentiment is not without any problem. Through this paper, we
propose a methodology in order to analyse the user’s sentiment of such mobile apps. Generally, the contribution of our paper can be listed as the following:

1. We conduct sentiment analysis approach as the proxy for mobile payment apps measurement of nonbank PSP, as thus a proxy to oversee the condition of the non-bank payment system in the macroeconomic perspective.
2. We perform the comparison of several inference models and address some issues related to their performance due to the possibility of limited training data access.
3. We construct series of index based on the inferred sentiments.

The organisation of this paper is as follows: Section 2 provides literature review related to the assessment of mobile apps that may be related to our domain problem, Section 3 describes our methodology to infer the sentiment and to construct the sentiment’s index. Section 4 presents our experimental results and analysis. Finally, we conclude our works and discuss future direction in Section 5.

2. Literature Review

The importance of quality assessment of internet based proxy for banking or financial related activities can initially be described in the work by Rod et al (2009). In their work SEM PLS methodology was leveraged in order to analyse the questionnaire they gathered. Their work showed that customer service quality, online information system quality, and banking service product have positive impact to the customer satisfaction.

Following their works, a similar research goal had also been conducted by Ganguli et al (2011). A generic service quality dimension assessment of the internet banking was conducted through an exploratory factor analysis. They found four important aspects: customer services, technology security and information quality that affect the customer satisfaction, while technology convenience, and easiness and realibility that affect the customer loyalty.

The banking services are however expanding into mobile banking due to the freedom of time and place. Therefore mobile banking then obviously becomes a promising measurement source for proxy. For the banking sustainability reason, it is thus important to retain bank’s customer’s loyalty. Thakur (2013) investigated some factors that could have positive impact on customer’s loyalty to the bank. The work gives analytical result that mobile interface usability and service had positive impact on customer satisfaction. Arcand et al (2017) later, utilised SEM and found that security and practicity impact on trust, while enjoyment and sociality impact the commitment and satisfaction. The related works were also performed by Sagib and Zapan (2014), Rahman et al (2017), and Khan et al (2021), on the measurement of m-banking service quality.

Another work on more general domain related to the mobile apps quality assessment method can be seen in the work by Vu et al (2015). They proposed a methodology assessing user’s opinion on mobile apps that can be the concern for app’s developer. An amount of keywords related to the was found, and the work gives the result that the technology is one of the important factors to concern in order to improve user’s experience and satisfaction.
Related to the bank’s reputation, Bach et al (2020) investigated some aspects that determine the relationship between mobile banking and bank reputation. Their work leverage 500 clients on a number of local banks in Croatia. The work suggest that safety, simplicity, and service variability of mobile banking application gives positive impact to the bank reputation.

It is thus important to analyse the user’s sentiment on mobile payment apps. In addition to the work of Vu et al (2015), there have also been another related works related to the customer’s sentiment. Fang and Zhang (2015), Singla et al (2017) conducted the on sentiment analysis on review data of Amazon product utilising some NLP and machine learning techniques. These works however share the same property in the perspective of measurements of a product quality: utilising the customer’s opinion.

Recently, and more related to our case, the work of Leem & Eum (2021) utilised text mining in order to analyse the sentiment of mobile banking service quality. They leveraged user’s review obtained from application store. There were 5 aspects/dimension they investigate which are security/privacy, practicity, design/aesthetics, sociality and enjoyment. The analysis reported that the customers mostly complained about four topics which are process, interaction, customer convenience, and technology and function.

Considering the above conducted researches, we then consider some aspects related to the mobile apps applied to non-bank PSP. The case of bank PSP and non-bank PSP however shares the same property: utilising mobile apps in financial and payment system area. In our work, we propose an alternative, quick and timely methods as the proxy for qualitative measurements of non-bank PSP using text mining. Of those aspects mentioned in the literature, we deterministically define 5 aspects related to the mobile apps in order to demonstrate the sentiment analysis:

- **Security Aspect**: Represents the security of mobile apps. including e.g. some issue related to the problem on security aspects including log in/log out, registration, security of e-money.
- **Feature Aspect**: Represents the variability and functionality of the features offered by non-bank PSP’s.
- **Design Aspect**: Represents how well the design related to the intuitiveness and user experiences.
- **Customer Service Aspect**: Represents how responsible, reliable, and the effectiveness of the customer services.
- **Technology Aspect**: Represents how advanced the technology the non-bank PSP’s has, how well the application is from the technological issues.

3. Methodology

3.1. Data Collection and Annotation

**Data Collection**

In order to conduct the experiment, first we collect the data from Goole Play Store and Apple Appstore by scrapping methods, with Python programming language. The structure of information available can be seen in Fig 1 below.
Available Information on the Store

In general, the information that is available at the Google Play Store as on Fig 1 above consists of application id (appId), application name, application logo, number of rating reviews, number of download, application’s description, it’s latest update, user’s name, rating, review and time stamp of his/her review. The information available on Apple Appstore is however relatively similar. Of those information, for each of the appId we then extract the user’s review based on the username, name of user, and time stamp.

From about 41 nonbank-PSP’s total population, we obtained 35 mobile apps of nonbank-PSP available on the store, therefore the number of nonbank-PSP investigated though our work is 35. The total of the review scraped on Google Play Review and Apple AppStore for the period of December 2011 up to December 2021 in this experiment is 4,637,980 reviews, with language in Bahasa Indonesia.

Data Annotation

In order to develop and to validate our inference model, we annotate a number of samples. We sample 1270 datapoints (i.e. reviews) across the players over the monthly period. To this data we flag both aspects and sentiment’s magnitude. Here, a review may contain more than one aspect. For example, a review may discuss a bad feature of mobile app, but good at the user interface design. The distribution of this annotated dataset can be seen as on Table 1 below.

The sentiment category formulation in our case is slightly different compared to other sentiment analysis case study in general. For each aspect, we put the positive or not related aspect to be in the same category, while negative ones are categorised as negative. This is due to the purpose, that the policy maker may prefer to see only the negative sentiment. The challenge in our case is then the limited amount of training data as we are dealing with new specific domain problem in the specific language: mobile apps’ user’s reviews in Bahasa Indonesia.
### Annotated Dataset

<table>
<thead>
<tr>
<th>Aspect</th>
<th>Negative</th>
<th>Positive or Not Related</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>Security Aspect</td>
<td>449</td>
<td>821</td>
<td>1270</td>
</tr>
<tr>
<td>Feature Aspect</td>
<td>420</td>
<td>850</td>
<td>1270</td>
</tr>
<tr>
<td>Design Aspect</td>
<td>306</td>
<td>964</td>
<td>1270</td>
</tr>
<tr>
<td>Customer Service Aspect</td>
<td>390</td>
<td>880</td>
<td>1270</td>
</tr>
<tr>
<td>Technology Aspect</td>
<td>427</td>
<td>843</td>
<td>1270</td>
</tr>
</tbody>
</table>

1 Annotation is performed by 3 annotators.

### 3.2. Data Preprocessing

#### Text Normalisation.

The review obtained are of less formal lexical form. It is because the users are free to write anything expression as the review. Lexical normalisation is then an important stage so that the further feature extraction will become more optimum. We construct a dictionary and leverage deterministic regular expression in order to normalise the unnormalised lexical form. Any digit, punctuation and emoticon are removed. In addition to that, any review with length less than 25 characters are eliminated as after our inspection, most of them are meaningless/noise. Some examples of unnormalised vs. normalised text can be seen as in Table 2.

#### Unnormalised vs. Normalised Text

<table>
<thead>
<tr>
<th>Unnormalised Text (raw text)</th>
<th>Normalised Text (cleaned text)</th>
</tr>
</thead>
<tbody>
<tr>
<td>KAK SAYA SUDAH MELAKUKAN ISI SALDO VIA KLIK XX. SALDO SAYA DI XX TERPOTONG TAPI KOK TIDAK MASUK DI SALDO yy. TERUS UANGNYA NYANGKUT DIMANA???? mohon bantuannya</td>
<td>kak saya sudah melakukan isi saldo via klik xx saldo saya xx terpotong tapi kok tidak masuk saldo yy terus uangnya nyangkut dimana mohon bantuannya (in English: I have bought the credits phone using xx and the balance is debited, but no yy points added I got, where is my money? please help)</td>
</tr>
<tr>
<td>kecewa banget keamanan kurang, akun XX saya bisa disadap orang jadi kena penipuan ah anjir keselllll 1juta hilang</td>
<td>kecewa banget keamanan kurang akun xx saya bisa disadap orang jadi kena penipuan anjir kesell juta hilang (in English: very disappointed for bad security my xx account was tapped so I got scammed damn I lost millions)</td>
</tr>
<tr>
<td>Aplikasi yg Bagus.. Tampilannya menarik..</td>
<td>aplikasi yang bagus tampilannya menarik (in English: nice application, yet interesting user interface)</td>
</tr>
</tbody>
</table>

#### Feature Extraction.

We further apply tfidf bag of words weighting in order to extract the textual feature. This can be expressed as in the Eq 1 following.
each feature word \( w \) is converted to numerical representation according to the number of occurrence of words \( w \) in document \( d \) that is \( tf_{w,d} \) weighted by the log inverse of document frequency i.e. the frequency of document containing words \( w \), that is \( df_w \). Here \( N \) is number of all reviews in the collection.

3.3 Inference Models

**Rule Based Models (Deterministic Approach)**

The first and relatively the most straightforward inference model is to leverage rule based model. We provide predefined keywords that both belongs to aspects, and also sentiments. Some sample of list of keywords can be seen in Table 3 following.

<table>
<thead>
<tr>
<th>Keywords Category</th>
<th>List of Keywords Example</th>
<th>Translation (English)</th>
<th>Notes</th>
</tr>
</thead>
<tbody>
<tr>
<td>Security Aspect</td>
<td>log in, log out, penipuan, otp, verifikasi, username, password...</td>
<td>log in/out, scamming, otp, verification, username, password</td>
<td></td>
</tr>
<tr>
<td>Feature Aspect</td>
<td>Fitur, bayar, simpan, akses, servis, menu, saldo, emoney, pembayaran...</td>
<td>Feature, pay, save, access, service, menu, balance, emoney, payment.</td>
<td></td>
</tr>
<tr>
<td>Design Aspect</td>
<td>tampilan, desain, smooth, scroll, klik...</td>
<td>(user) interface, design, scroll, click...</td>
<td>In implementation, regular expression (regex) is utilised to reassure that the lexical/writing variety is captured.</td>
</tr>
<tr>
<td>Customer Service Aspect</td>
<td>aduan, complain, keluhan, tanggapan...</td>
<td>complain, rant, (customer service’s) response.</td>
<td></td>
</tr>
<tr>
<td>Technology Aspect</td>
<td>memory, lemot, lelet, error, memory, slow, error, crash...</td>
<td>memory, slow, error, crash</td>
<td></td>
</tr>
<tr>
<td>General Negative Sentiment</td>
<td>menyebaikan, menyusahkan, sampah, konyol, lemot, error, crash, jelek...</td>
<td>annoying, bothering, rubbish, awful, slow, error, bad</td>
<td></td>
</tr>
<tr>
<td>General Positive Sentiment</td>
<td>bagus, baik, mantap...</td>
<td>good, nice, excellent...</td>
<td></td>
</tr>
<tr>
<td>Negation</td>
<td>Tidak, belum, kagak...</td>
<td>not, not yet.</td>
<td></td>
</tr>
</tbody>
</table>

In order to infer the aspect and sentiment, we leverage those keywords into the simple rule as shown in Eq 2. Basically, the rule \( r \) tries to find aspect and its sentiment based on the lexical occurrence given the keywords on a review.

\[
r(S) = \exists (kw_{aspect}) \in S \bigwedge \exists (kw_{negative\_sentiment}) \in S
\]

Here \( S \) is review text and \( kw \) is keyword. This rule will return 1 if there is any negative sentiment on the particular aspect. Any negative sentiment that is followed by negation keywords will turn positive, and thus will give 1 value for \( r(S) \). However, after our manual inspection, this kind combination is relatively rare, as usually user
states single token opinion sentiment. Although the knowledge representation can be expressed explicitly, this model costs on the pattern complexity: it has to be able to do a generalisation well to accurately extract the information.

**Models Built from Data**

In order to perform the sentiment using models that rely on the data, we then utilise machine learning. We demonstrate the models which are: Decision Tree with gini index as splitting criteria, SVM with radial basis function (RBF) kernel, Logistic Regression. In our case, each aspect's sentiment is analysed with different models each other. Therefore, we have in total 15 machine learning models to assess (3 models for each aspect's sentiment analysis, and we have 5 aspects). These models are trained and evaluated onto the annotated data at the Table 1. We use 5-fold cross validation scenario in order to evaluate those models. Furthermore, onto the model from data that achieve high result, we also perform experiment on elaborating prior knowledge as discussed on the Result and Discussion Section.

**Imbalanced Data Handling**

As we have seen on Table 1 that on per aspect’s perspective, the distribution between negative and non-negative classes (either positive or not related) are relatively imbalance. We then leverage the oversampling technique for the minority class (in this case negative class) with Synthetic Minority Oversampling Technique (SMOTE) (Chawla et al, 2002) to enable the data synthesis. We perform such approach in the feature space, before the parameter estimation of those machine learning models.

**Evaluation of the Inference Models**

Our problem can be seen as a binary problem, with negative sentiment is our interest magnitude. To evaluate the performance, we use F1 scores as formulated as follows:

\[
F1 Score = \frac{2 \times (\text{Precision} \times \text{Recall})}{\text{Precision} + \text{Recall}}
\]

where

\[
\text{Recall} = \frac{TP}{(TP + FN)}
\]
\[
\text{Precision} = \frac{TP}{(TP + FP)}
\]

**3.4. Index Formulation**

After the full data are completely inferred by applying the best model trained from the annotated data, the next stage is to construct an index for each aspect. A review on each aspect should have either 0 (positive/neutral) or 1 (negative). The index is formulated as shown in Eq 3 below:

\[
\text{id}_{x_{o,t}} = \sum_{i} I(f(x_{o,t}) = 1) \frac{x \log N_{o,t}}{N_{o,t}}
\]

Generally speaking, the index is about the proportion of negative sentiment of a NB-PSP o within time t over the total of its review at the time t that is, \( N_{o,t} \).

In our case study, we consider prefer negative sentiment to positive sentiment. Negative sentiment is rather more useful to the policy makers as they would oversee
any problem so that they could consider further decision, to prevent any subsequent systemic risk on monetary and payment system.

The log term in Eq 3 is to weight the magnitude of the proportion. In other words, it is considered to quantitatively represent how important a NB-PSP is within the share. Intuitively, a number of negative sentiment on the top player should be take care of compared to the player whose small portion of market share.

4. Result and Discussion

4.1. Result on the Inference Models

The experimental result can be seen in Table 4. To note that we also experimented the inference scenario without SMOTE beforehand (that is let the imbalanced data as it is when training the model). However, the result is not better than when the SMOTE is utilised. Therefore, we only show the model when the SMOTE is performed during the training phase. We performed the experiment with 5-fold cross validation.

<table>
<thead>
<tr>
<th>Performance matrix of models from data (F1 in %)</th>
<th>Table 4</th>
</tr>
</thead>
<tbody>
<tr>
<td>Security Aspect</td>
<td>Feature Aspect</td>
</tr>
<tr>
<td>Rule Based</td>
<td>56.49</td>
</tr>
<tr>
<td>SVM</td>
<td>76.14</td>
</tr>
<tr>
<td>Decision Tree</td>
<td>72.00</td>
</tr>
<tr>
<td>Logistic Regression</td>
<td>78.17</td>
</tr>
</tbody>
</table>

According to the performance evaluation, the Logistic Regression model gives highest F1 score in general, at least in 4 aspects. In the Technology aspect however SVM gives the higher result than the Logistic Regression although the score differences is relatively small. The rule based on the other hand gives the lowest performance (less than 60% in all aspects).

Having the performance result as shown in Table 4 above, it can be seen that in general Logistic Regression takes over the other inference models. We then performed further experiment in the next part on how we incorporate the prior knowledge into our Logistic Regression.

Incorporating Prior Knowledge

The models from data as implemented previously rely on the data. We in this case attempt to incorporate our prior knowledge into the model for two reasons. First usually it is hard to access the labelled data to train the model properly i.e. annotation is a costly process. Second, we have prior knowledge related to the information we want to find. In our case we have some keywords and rules as our prior knowledge.

We then demonstrate on how to incorporate the prior knowledge. In general, Logistic Regression can be expressed as Eq 4 follows:
\[ p(y = 1|x) = \sigma(f(x)) = \frac{1}{1 + e^{-f(x)}} \quad (4) \]

where \( f(x) = w\phi(x) \) represents our linear function between parameter \( w \) and basis function \( \phi(x) \). Here \( x \in \mathbb{R}^d \) is feature vector representation (our input), and \( y \in \{0,1\} \) is the target (our label). The objective function is to minimise negative log likelihood as formulated in Eq 5:

\[ J = \sum_i \ln(1 + e^{-(2y_i-1)f(x_i)}) \quad (5) \]

We recapture the idea of the work by Schapire et al, 2002 on how to incorporate prior knowledge into the model. First we introduce a distribution that quantifies our prior belief from rule based model, that is:

\[ \pi = p(y = 1|x) = \begin{cases} 0.9, & \text{if } r(S_x) = 1 \text{ (True)} \\ 0.1, & \text{otherwise} \end{cases} \quad (6) \]

The objective function as defined on Eq 5 is then modified, becomes:

\[ J = \sum_i \ln(1 + e^{-(2y_i-1)f(x_i)}) + \eta D_{KL}(\pi(x_i)||\sigma(f(x_i))) \quad (7) \]

The second part of the addition in Eq 7 above intuitively represents the control of our prior knowledge. It is quantified by KL Divergence between the prior information \( \pi \) and the information obtained from training data \( f(x) \). The variable \( \eta \) controls the importance of the prior information. In our experiment, we leave its value into 1 as in this work it is not our main focus. We then estimate the model’s parameter by minimising cost function in Eq 7 by applying BFGS approximation algorithm.

The obtained parameter is then used in the decision function. First, we transform our prior probability into \( h_0(x) \) as follows:

\[ h_0(x) = \sigma^{-1}(\pi(x)) = \ln\left(\frac{\pi(x)}{1 - \pi(x)}\right) \quad (8) \]

Subsequently, the \( h_0(x) \) is blended to the final function: \( f^*(x) = f(x) + h_0(x) \). The decision function as expressed in Eq 4 thus becomes \( p(y = 1|x) = \sigma(f^*(x)) \).

The experimental results of this scenario can be seen as in Table 5 below.

---

### Performance on Logistic Regression with Prior Knowledge (F1 in %)

<table>
<thead>
<tr>
<th>Security Aspect</th>
<th># of Training Data</th>
<th>Logistic Regression</th>
<th>Prior Knowledge (Rule Based Model)</th>
<th>Logistic Regression with Prior Knowledge</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>100</td>
<td>54,11</td>
<td>56,49</td>
<td>55,00</td>
</tr>
<tr>
<td></td>
<td>200</td>
<td>66,67</td>
<td>56,49</td>
<td>67,86</td>
</tr>
<tr>
<td></td>
<td>500</td>
<td>72,64</td>
<td>56,49</td>
<td>75,20</td>
</tr>
<tr>
<td></td>
<td>800</td>
<td>75,76</td>
<td>56,49</td>
<td>78,82</td>
</tr>
<tr>
<td></td>
<td>1016</td>
<td>78,17</td>
<td>56,49</td>
<td>81,36</td>
</tr>
<tr>
<td>Feature Aspect</td>
<td>100</td>
<td>50,39</td>
<td>56,14</td>
<td>57,65</td>
</tr>
<tr>
<td></td>
<td>200</td>
<td>51,81</td>
<td>56,14</td>
<td>74,90</td>
</tr>
<tr>
<td></td>
<td>500</td>
<td>62,50</td>
<td>56,14</td>
<td>77,67</td>
</tr>
<tr>
<td></td>
<td>800</td>
<td>64,15</td>
<td>56,14</td>
<td>77,87</td>
</tr>
</tbody>
</table>

---

Sentiment Analysis of User’s Reviews on Non-Bank Payment Service Apps
Table 5 shows that incorporating prior knowledge can help improve our previous best model (Logistic Regression). The model’s performance on varying number of training set is evaluated on a fixed testing set for fairness/consistency reason. Initially, when the number of training data is small, the Logistic Regression gives result the lowest F1 score. Adding prior knowledge at this stage helps increase the inference process, although the performance is still below the rule based model. As the size of training data grows, the performance of both two models (model Logistic Regression and Logistic Regression with Prior Knowledge) increases. The Logistic Regression with Prior Knowledge however gives the better performance at any training size over the Logistic Regression model. The rule based model shows constant score as this is a deterministic function. Once the training data is sufficient, the Logistic Regression with Prior Knowledge achieves highest score. We then leverage the trained Logistic Regression with prior knowledge to be applied into full data so that the index can further be constructed based on the inferred results.

4.2 Result on the Index Series

After all the reviews are inferred (by Logistic Regression with prior knowledge), we then construct the index based on the polarity i.e. negative sentiment over all reviews on each particular NB_PSP’s and specific time granularity. In this case, we discuss on two examples of the NB_PSP’s (NB_PSP1 and NB_PSP2) having slightly different trends, yet still describing the increase of the digitalisation during the pandemic of COVID-19.

We investigate the series, mainly since 2019, because at this year the license for all the NB-PSP’s was issued. In addition to that, the number of customers started to increase since 2019 as any financial activities were more becoming digitalised.

The series index for NB_PSP1 for 5 aspects can be seen in Fig 3 while NB_PSP2 for 5 aspects in Fig 4. The blue line (for NB_PSP1)/red line (for NB_PSP2) represents the index (left hand side), while the grey bar represents the total reviews (right hand side). For confidentiality reason, we denote the total reviews on the graphic as N_{NB_PSP1} (or N_{NB_PSP2}) for the number of review of NB_PSP1 (or NB_PSP2). The higher index indicating the more negative sentiments received. The period shown is in the monthly basis. The list of words on the annotation (on event analysis) represents the list of
topics being reviewed. They are based on the order of most frequent words extracted from the inferred negative sentiments.

Series of Index of NB_PSP1

![Graphs showing the index of different aspects over time.](image)

* The period is from Jan 2019 until Dec 2021

From Fig 3 we can see that during the period of 2019 until 2021, the number of reviews gradually increased. In March and April 2019, the NB_PSP1 showed the increasing index on Security aspect, Feature aspect, as well as Customer Service aspect. At this stage, a number of customers were identified to complain about problems on the e-money in the payment services provided. It was the period when the NB_PSP1 expansively promoted their service product after legally licensed by the authority.

In the period of July 2020 when the pandemic of COVID-19 had started to hit the country, the customer's complaint and negative sentiments were about the Security, Feature and Customer Service's aspect. At this period, the physical distancing and lock down policy by the government were conducted. The number of customers increased as most of them practiced the digital transactions, and thus the mobile payment apps were in very high demand. The Technology and Design aspects, however, showed only once dramatic increasing curve in July 2019. Afterwards, the number of customers complained about its technological and design aspect declined. It is identified that the NB_PSP1 made some improvements to enrich the customer experiences and apps reliability, especially during the pandemic of COVID-19.
On the other hand, the NB_PSP2 (Fig 4) depicted the relatively different trend. We can see that during the period of 2019 until 2021 the number of overall reviews showed relatively no significant increase, except in September 2020. The index, however, started to increase at the time the pandemic of COVID-19 hit the country. It is identified that the app’s demand was highly increasing, and some possible trouble or unsatisfactory things were reported by user more than before. However, almost 5 aspects of NB_PSP2 showed relatively less differences of the trend, except in Q2 2020 as well as Q1 2021. In addition to that, in these two periods the Technology aspect was the most complained aspect.

Series of Index of NB_PSP2

Figure 4

* The period is from Jan 2019 until Dec 2021

5. Conclusion and Future Works

5.1 Conclusion

Through this work, we have performed the sentiment analysis method of user’s reviews collected from the mobile application’s store in order to oversee the quality of mobile payment apps of a nonbank-PSP. We conducted some experiments on the inference models: deterministic model (rule based), model-from-data (machine learning), and model-from-data incorporating prior knowledge. In our case study, the
model-from-data incorporated with prior knowledge helps infer the sentiment when the access to the training data is limited. Having demonstrating it onto 5 aspects, we suggest that this approach can be used to automatically infer the user’s sentiment of mobile non-bank payment apps by some aspects.

We also propose an index based on the inferred sentiment and the number of overall reviews, that can represent the recent condition. By automatically analyse the sentiment and monitoring the series periodically we can suggest that this approach can be used in order to timely monitor the nonbank-PSP performance, e.g. as a leading indicator. Therefore, any risks on monetary and payment system affected by the issues in non-bank payment system environment hopefully can be anticipated in advance, as early as possible.

5.2. Future Works

We notice that there are still some improvements needed in our works. We thus highlight some future directions:

- **Advancement on Text Normalisation Methods.**
  In our model, we implement a deterministic approach by utilising regular expression. In order to improve the inference’s performance, more sophisticated text normalisation leveraging more advanced language modelling should be used.

- **Enrich more various user’s expression.**
  The method presented in this work relies on textual feature. In the future, the model should also be able to cover sarcastic text as well as emoji/emoticon.

- **Add more Human Languages.**
  Although still in a small portion, some users write review in non-Bahasa Indonesia, for example English. Thus in the future, the model should be able to tackle non-Bahasa Indonesia language, although its proportion is relatively small.

References


Sentiment Analysis of User’s Reviews on Non-Bank Payment Service Apps

N. Heru Praptono, Arinda D. Okfantia, M.Khoyrul Hidayat, M. Hafiruddin

Statistics Department

Basel, Switzerland
August 2022

1The expressed views belong to the authors and do not necessarily reflect the institution.
2Corresponding email: nursidik.hp@bi.go.id/pra.heru@gmail.com
Outline

Background

Methodology
   Overall Methodology
   Annotated Data and Model Evaluation
   Incorporating Prior Knowledge

Non-Bank PSP Apps Index

Conclusion and Future Directions
Background

- More banking and payment activities are currently being conducted into mobile application platform than the traditional ones, legally provided by the authorised entities either bank or non-bank institutions.

- In Indonesia for example,
  - By the end of 2021, there have been at least 41 non-bank payment service providers (PSPs) who take a part in such banking and payment services.
  - The growing amount of non-bank PSP customers has found to be increasing recently, moreover since the pandemic of COVID-19.
  - Have infiltrated on nearly the whole of about 270 million people of the Indonesian population → play significant role in promoting and accelerating the economic growth.
  - Need to be supervised and monitored by the policy maker (i.e. Bank Indonesia) to anticipate any systemic risk.

- Measuring and monitoring the quality of non-bank payment service apps is difficult – but possible by considering user’s review. Survey is costly, solution: apps review (i.e. in Google Play Store, Apps Store) as a proxy.

- Related works on mobile apps user’s review analysis, e.g. Vu et al 2015 (mobile apps), Leem & Eum 2021 (m-banking).

- Some remaining issues (including but not limited to): non-formal text, limited training data, imbalanced training data, further utilisation for monitoring.
Overall Methodology

Unnormalised Text (raw text) | Normalised Text (cleaned text)
---|---
Aplikasi yg Bagus.. Tampilannya menarik.. | aplikasi yang bagus tampilannya menarik
*(in English: nice application, yet interesting user interface)*

$$tfidf_{w,d} = tf_{w,d} \times \log \frac{N}{df_w}$$

1. **Security**: Represents the security of mobile apps, some issues related to the problem on security aspects including log in/log out, registration, security of e-money.
2. **Feature**: Represents the variability and functionality of the feature offered by non-bank PSP’s
3. **Design**: Represents how well the design related to the intuitiveness and user experiences.
4. **Cust. Services**: Represents how responsible, reliable, and effectiveness of the customer services.
5. **Technology**: Represents how advanced the technology the non-bank PSP’s is, and represent how well the application is from the technological issues.

*e.g.: Bach et al, 2020; Leem & Eum 2021*

Overall Methodology

Annotated Data and Model Evaluation

Incorporating Prior Knowledge
Annotated Data and Model Evaluation

Annotated data are relatively imbalanced → utilise SMOTE (Chawla et al, 2002) during learning process. Some experimented models → rule based model, SVM (with RBF Kernel), Decision Tree (Gini Splitting Criteria), and Logistic Regression. The experiment is performed with 5-fold cross validation.

<table>
<thead>
<tr>
<th>Annotated Data</th>
<th>Table 1</th>
</tr>
</thead>
<tbody>
<tr>
<td>Annotated Dataset</td>
<td></td>
</tr>
<tr>
<td>Security Aspect</td>
<td>449</td>
</tr>
<tr>
<td>Feature Aspect</td>
<td>420</td>
</tr>
<tr>
<td>Design Aspect</td>
<td>306</td>
</tr>
<tr>
<td>Customer Service Aspect</td>
<td>390</td>
</tr>
<tr>
<td>Technology Aspect</td>
<td>427</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Model Evaluation</th>
<th>Table 4</th>
</tr>
</thead>
<tbody>
<tr>
<td>Performance matrix of models from data (F1 in %)</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Security Aspect</td>
</tr>
<tr>
<td>Rule Based¹</td>
<td>56,49</td>
</tr>
<tr>
<td>SVM</td>
<td>76,14</td>
</tr>
<tr>
<td>Decision Tree</td>
<td>72,00</td>
</tr>
<tr>
<td>Logistic Regression</td>
<td>78,17</td>
</tr>
</tbody>
</table>

¹Rule based: \( r(S) = \exists (kw_{aspect}) \text{ in } S \land \exists (kw_{negative\_sentiment}) \text{ in } S \)

1 Annotation is performed by 3 annotators.
Incorporating Prior Knowledge

We adopt the methodology incorporating prior knowledge described by Schapire et al, 2002.

- **Prior Knowledge**: $r(S_x)$ (essentially the result of rule based model)
- **Learning**: Given the training dataset $D = \{(x_i, y_i)\}_{i=1}^N$, and $y \in \{0, 1\}$. The objective function is to minimise negative log likelihood, controlled by the prior information.

$$J = \sum_i \left[ \ln(1 + e^{-(2y_i - 1)f(x_i)}) + \eta D_{KL}(\pi(x_i)||\sigma(f(x_i))) \right]$$

where $f$ is linear function, $\sigma$ is the logistic function. Here $\pi$ is our "prior information" quantification, defined by:

$$\pi(x) = p(y = 1|x) = \begin{cases} 0.9; & \text{if } r(S_x) = 1 \text{ (true)} \\ 0.1; & \text{otherwise} \end{cases}$$

- **Inference**:

$$p(y = 1|x) = \sigma(f^*(x)); f^* = f + h_0$$

In this case, $h_0$ is the "prior term" defined as the inverse of logistic function of $\pi(x)$, that is $h_0(x) = \sigma^{-1}(\pi(x)) = \ln \left( \frac{\pi(x)}{1-\pi(x)} \right)$
Incorporating Prior Knowledge (Cont’d)

Experimental Result

- The model’s performance on varying number of training set is evaluated on a fixed testing set for fairness reason.
- Initially, when the number of training data is small, the Logistic Regression gives result the lowest F1 score.
- Adding prior knowledge at this stage helps increase the inference process, although the performance is still below the rule based model.
- As the size of training data grows, the performance of both two models (model Logistic Regression and Logistic Regression with Prior Knowledge) increases.
- The Logistic Regression with Prior Knowledge however gives the better performance at any training size over the Logistic Regression model.
- The rule based model shows constant score as this is a deterministic function.
- Once the training data is sufficient, the Logistic Regression with Prior Knowledge achieves highest score.
- We then leverage the trained Logistic Regression with prior knowledge to be applied into full data so that the index can further be constructed based on the inferred results.

<table>
<thead>
<tr>
<th>Aspect</th>
<th># of Training Data</th>
<th>Logistic Regression</th>
<th>Prior Knowledge (Rule Based Model)</th>
<th>Logistic Regression with Prior Knowledge</th>
</tr>
</thead>
<tbody>
<tr>
<td>Security Aspect</td>
<td>100</td>
<td>54.11</td>
<td>56.49</td>
<td>55.00</td>
</tr>
<tr>
<td></td>
<td>200</td>
<td>66.67</td>
<td>56.49</td>
<td>67.86</td>
</tr>
<tr>
<td></td>
<td>500</td>
<td>72.64</td>
<td>56.49</td>
<td>75.20</td>
</tr>
<tr>
<td></td>
<td>800</td>
<td>75.76</td>
<td>56.49</td>
<td>78.82</td>
</tr>
<tr>
<td></td>
<td>1016</td>
<td>78.17</td>
<td>56.49</td>
<td>81.36</td>
</tr>
<tr>
<td>Feature Aspect</td>
<td>100</td>
<td>50.39</td>
<td>56.14</td>
<td>57.65</td>
</tr>
<tr>
<td></td>
<td>200</td>
<td>51.81</td>
<td>56.14</td>
<td>74.90</td>
</tr>
<tr>
<td></td>
<td>500</td>
<td>62.50</td>
<td>56.14</td>
<td>77.67</td>
</tr>
<tr>
<td></td>
<td>800</td>
<td>64.15</td>
<td>56.14</td>
<td>77.87</td>
</tr>
<tr>
<td></td>
<td>1016</td>
<td>74.80</td>
<td>56.14</td>
<td>78.39</td>
</tr>
<tr>
<td>Design Aspect</td>
<td>100</td>
<td>45.24</td>
<td>55.43</td>
<td>52.38</td>
</tr>
<tr>
<td></td>
<td>200</td>
<td>60.32</td>
<td>55.43</td>
<td>64.43</td>
</tr>
<tr>
<td></td>
<td>500</td>
<td>63.25</td>
<td>55.43</td>
<td>67.95</td>
</tr>
<tr>
<td></td>
<td>800</td>
<td>63.45</td>
<td>55.43</td>
<td>69.39</td>
</tr>
<tr>
<td></td>
<td>1016</td>
<td>69.32</td>
<td>55.43</td>
<td>73.52</td>
</tr>
<tr>
<td>Customer Service Aspect</td>
<td>100</td>
<td>45.27</td>
<td>53.78</td>
<td>54.86</td>
</tr>
<tr>
<td></td>
<td>200</td>
<td>54.50</td>
<td>53.78</td>
<td>74.58</td>
</tr>
<tr>
<td></td>
<td>500</td>
<td>60.00</td>
<td>53.78</td>
<td>74.58</td>
</tr>
<tr>
<td></td>
<td>800</td>
<td>74.16</td>
<td>53.78</td>
<td>77.24</td>
</tr>
<tr>
<td></td>
<td>1016</td>
<td>79.07</td>
<td>53.78</td>
<td>82.78</td>
</tr>
<tr>
<td>Technology Aspect</td>
<td>100</td>
<td>40.00</td>
<td>55.85</td>
<td>47.50</td>
</tr>
<tr>
<td></td>
<td>200</td>
<td>55.81</td>
<td>55.85</td>
<td>65.21</td>
</tr>
<tr>
<td></td>
<td>500</td>
<td>63.64</td>
<td>55.85</td>
<td>70.61</td>
</tr>
<tr>
<td></td>
<td>800</td>
<td>65.16</td>
<td>55.85</td>
<td>75.30</td>
</tr>
<tr>
<td></td>
<td>1016</td>
<td>65.61</td>
<td>55.85</td>
<td>77.90</td>
</tr>
</tbody>
</table>
Result: Index (NB_PSP1)

In March and April 2019, the NB_PSP1 showed the increasing index on Security aspect, Feature aspect, as well as Customer Service aspect. In the period of July 2020 when the pandemic of COVID-19 had started to hit the country, the complaint and negative sentiments were about the Security, Feature and Customer Service aspect. The Technology and Design aspects, however, showed only once dramatic increasing curve in July 2019 and declined afterwards → it was identified that some improvement happened.
Result: Index (NB_PSP2)

During the period of 2019 until 2021 the number of overall reviews showed relatively no significant increase, except in September 2020. The index, however, started to increase at the time the pandemic of COVID-19 hit the country. It is identified that the app’s demand was highly increasing, and some possible troubles or unsatisfactory things were reported by user more than before. The most complained aspect was Technology aspect, 2020-Q2, 2021-Q1.

\[
\text{idx}_{o,t} = \frac{\sum_1^T f(x_{o,i,t})}{N_{o,t}} x \log N_{o,t}
\]
Conclusion and Future Directions

**Conclusion**

1. We conducted some experiments on the inference models: deterministic model (rule based), model-from-data (machine learning), and model-from-data incorporating prior knowledge. Prior knowledge can help to improve the inference process when the number of training data is limited.

2. Having demonstrating it onto 5 aspects, we suggest that this approach can be used to infer the user’s sentiment of mobile non-bank payment apps by some aspects in a big data quickly.

3. We also propose an index that is based on the inferred sentiment and the number of reviews. The series constructed by the index calculation represents the recent condition because the reviews can be scrapped at any time from the application’s store.

4. By automatically analyse the sentiment and monitoring the series periodically we can suggest that this approach can be used in order to timely monitor the nonbank-PSP performance, e.g. as a leading indicator so that any further systemic risk can be anticipated in advance.

**Future Directions**

1. Advancement on Text Normalisation Methods.

2. Enrich more various user’s expression.

3. Add more Human Languages.
Fintech in statistical classifications: suggestions and tentative figures in a central bank context

Ulf von Kalckreuth and Norman Wilson, Deutsche Bundesbank,
Celestino Giron and Urszula Kochanska, European Central Bank,
Enzo Buthiot and Yann Wicky, Bank of France,
Luis Ángel Maza and Román Santos, Bank of Spain
Fintech in statistical classifications: suggestions and tentative figures in a central bank context

Enzo Buthiot (Banque de France), Celestino Giron (ECB), Luis Ángel Maza (Banco de España), Ulf von Kalckreuth (Deutsche Bundesbank), Urszula Kochanska (ECB), Román Santos (Banco de España), Yann Wicky (Banque de France), Norman Wilson (Deutsche Bundesbank).^1

Abstract

Fintech happens where innovation takes place in the financial sector and where new methods and products emerge, are tested and made ready for the market. The results are shaping the financial industry as a whole. Central banks need to identify, describe and understand fintech activities.

Considering the central banks monitoring needs and following the targeted roadmap to construct Fintech statistics outlined in the central banks’ international initiative under the auspices of the IFC^2, we i) provide an overview of fintech activities, ii) discuss on how best reflect this reality in statistical classifications based on the example of the European Classification of Economic Activities (NACE), currently under review and iii) elaborate on the Fintech landscape in France, Spain and Germany,

Keywords: Fintech, NACE, statistical classification, monitoring

JEL classification: C81, G23, O30

^1 The authors thank Isabel Kerner (European Central Bank), Jana Stamer and Robert Kirchner (Deutsche Bundesbank) for important comments. The views expressed are those of the authors and do not necessarily reflect those of the affiliated institutions.

^2 The Irving Fisher Committee on Central Bank Statistics (IFC) Working Group on Fintech data, which published a report Towards monitoring financial innovation in central bank statistics, IFC Report No 12, July 2019
Contents

1. Introduction ....................................................................................................................................... 3
2. Overview of Fintech activities and how to classify them ................................................. 4
   Crypto-asset activities .................................................................................................................. 4
   Fintech related to financing and services auxiliary to financing ............................................. 5
   Fintech related to investment, asset management, and trade ................................................... 6
   Fintech related to payment services ....................................................................................... 7
   Digital-only banks, Quasi-banks and BankTech ................................................................. 8
   InsurTech and PensionTech ....................................................................................................... 8
   Tech facilitators ............................................................................................................................... 8
   SupTech and RegTech – borderline cases ............................................................................ 9
3. Fintech in statistical classifications from a central bank perspective: the case of NACE rev.2 .......................................................................................................................................... 9
4. Fintech Statistics at the Bank of Spain .................................................................................. 11
5. Experimental Data at the Deutsche Bundesbank ............................................................. 12
6. Fintech statistics at the Banque de France .......................................................................... 13
Sources ...................................................................................................................................................... 15
1. Introduction

In the last decade, innovation activity in the financing industry has, to a large degree, been moving to entities commonly labelled Fintech amid a wider phenomenon of decentralised finance in which financial functions migrate away from the traditional core of intermediaries, infrastructures, and instruments. Fintech refers to solutions for innovative products, services, and processes in the financial industry that could improve, complete and/or disrupt existing financial products and services. Decentralisation of financial intermediaries is evidenced by the emergence of Fintech players from outside the traditional financial sector as well as by finance embedded in non-financial digital platforms. As for the decentralisation of infrastructure, processes and the instruments, technology is enabling fundamentally different approaches to the provision of financial services and to the creation of new assets which do not follow the traditional asset-liability model. Fintech also apply to incumbent financial institutions that adopt similar business models. On the other hand, Tech facilitators embracing new technologies and supporting financial activities in terms of necessary infrastructure have also enabled innovation in the financial industry. However, in this paper, activities of Tech facilitators are not considered as financial as they do not constitute per se the provision of financial services.

There is no internationally accepted harmonised definition of Fintech for statistical or for other classification purposes. The most commonly used definition is that of the FSB defining Fintech broadly as “technologically enabled innovation in financial services that could result in new business models, applications, processes or products with an associated material effect on financial markets and institutions and the provision of financial services.” References to volatile and evolving concepts such as “innovation” and “technology” render this definition unsuitable for a statistical classification system which should be stable over time.

---

3 Menon R. (2021)

4 BigTech, technology firms, individuals, undefined groups of online users (“crowd”); e.g. as in the case of bitcoin, behind which there is no identifiable entity or person, or in decentralised applications (DAapps) on the Ethereum blockchain.

5 The term TechFin is often used for embedded finance, see e.g. https://payspacemagazine.com/fintech/the-difference-between-fintech-techfin/.

6 crypto network, self-executing smart contracts, non-custodian financial services where users maintain control over their assets at all times,

7 alternative currencies compared to those issued by central banks (non-fiat currencies by technology firms or individuals)

8 Nevertheless, beyond the context of NACE, Tech facilitators might be relevant in NCBs analysis as they constitute an important aspect of the value chain.


10 The term “Fintech” was coined from “finance” and “technology”, however defining Fintech as “financial technology” is incomplete.

11 von Kalckreuth and Wilson N. (2020)

12 Unified Theory of Acceptance and Use of Technology (UTAUT) postulated by Venkatesh et al. (2003) identifies the four main constructs “performance expectancy”, “effort expectancy”, “social influence” and “facilitating conditions” to predict the behavioural intention to use a certain technology.
This paper clarifies the scope of (current) Fintech activities to support their classification in statistical systems, thereby enabling their monitoring. Novel Fintech activities entailing financial intermediation and financial auxiliary functions should be identified and classified in an appropriate place in the statistical systems. Currently, many Fintech activities are not included in activities or product classifications entries for financial intermediation and auxiliary functions. The possibility of introducing new classification entries is discussed, taking as a point of departure the current NACE hierarchical structure K and distinguishing within the Fintech universe between the novel activities, old activities done in a new way, and activities which are not financial. The paper aims to serve as an input to any discussions on Fintech in the context of statistical classifications e.g. in the remaining stages of the current reviews of the NACE, the International Standard Industrial Classification of All Economic Activities (ISIC) and the classifications of products (CPC/PA). Importantly, this paper continues the central banks’ work on statistical classification of Fintech and largely completes the outlined road-map to construct fintech statistics envisaged in the IFC Report *Towards monitoring financial innovation in central bank statistics*\(^{13}\). Finally, the topic of this paper fits well in the context of the new Data Gap Initiative (DGI) currently discussed with its workplan covering recommendations under four main statistical and data priorities, of which priority three focuses on Fintech and financial inclusion\(^{14}\).

2. Overview of Fintech activities and how to classify them

This section describes business areas with technological innovations related to financial services\(^{15}\). They include different types of activities, namely (i) financial activities that are totally novel (ii) financial activities that are in principle already covered in NACE but their provision is enhanced by innovation, (iii) activities of Tech Facilitators\(^{16}\) and (iv) borderline cases.

**Crypto-asset activities**

“No liability crypto-asset (NLCA)\(^{17}\)” such as bitcoin, a prominent type of crypto-assets, can be defined as a new type of asset recorded in digital form and enabled by the use of cryptography that is not, and does not represent, a financial claim on, or a liability of, any identifiable entity\(^{18}\). Such NLCAs are truly new assets and have been

---

13 IFC (2019)

14 G20 Data Gaps Initiative (DGI-2) Progress Achieved, Lessons Learned, and the Way Forward; Prepared by the IMF Staff and FSB Secretariat 9 June 2022

15 Based on various reflections on input e.g. Deutsche Bundesbank, 2022, U. Kochanska Characterisation of the euro area fintech scene, Financial Integration and Structure in the Euro Area, ECB 2020.

16 That are not financial but for which automation, new delivery platforms, DLT, AI, and others provide the infrastructural foundation for innovation in the financial sector.

17 Such assets are also often labelled as Crypto-Assets Without Liabilities (CAWL) and unbacked crypto-assets.

18 Based on ECB (2019)
enabled through the emergence of the distributed ledger technology (DLT). The technology underpins i) the peer-to-peer transactions, ii) the decentralised way of validating transactions and, iii) when applicable, of issuing this type of crypto-asset. Stablecoin - another type of crypto-assets – that uses stabilisation mechanisms which can minimise price fluctuations\(^\text{19}\) may constitute a subset of NLCA\(^\text{20}\) when there is no liability attached. However, many stablecoins do have a liability attached. The crypto-assets ecosystem also encompasses Decentralised Finance (DeFi), in which financial applications are run by smart contracts and offer novel protocols, e.g. for trading, lending, and investing\(^\text{21}\). DeFi tokens very often do not have a liability attached either. Crypto-asset activities, especially related to NLCA, including emission, issuance, operating (validating), and services are considered novel and should warrant a dedicated entry in statistical classifications e.g. in K66 of the current NACE. Importantly, crypto-asset may be reused in other activities e.g. in payments.

In the context of crypto-assets it is worth mentioning tokenisation, which refers to the process of issuing a token on a distributed ledger (e.g. blockchain)\(^\text{22}\). Tokenisation typically covers securities or more broadly traditional financial instruments with liabilities attached, therefore this is not an entirely new activity. Nevertheless, considering the complexity of separating NLCA activities from activities related to crypto-asset with liabilities, they both could be lumped in the novel crypto-asset activities in K66 of the current NACE until the sub-segments are sufficiently big and distinguishable to treat them separately.

**Fintech related to financing and services auxiliary to financing**

This category covers Fintech activities supporting the access to funds, which includes crowdfunding and supply chain financing. First, **crowdfunding can be defined as the efforts by individuals or groups to fund their ventures by drawing on small contributions from a relatively large number of individuals (“crowd”) using the Internet**\(^\text{23}\). Crowdfunding draws inspiration from the concept of microfinance and crowdsourcing. It represents a unique form of fundraising where the demand side on the capital market (project proponents) are linked to the supply side (investors) through a crowdfunding intermediary (platform).\(^\text{24}\) Two main models of crowdfunding platforms\(^\text{25}\) can be distinguished (see Table 1).

---

20 Based on their design, stablecoins can be classified into four types: (i) tokenised funds; (ii) off-chain collateralised stablecoins; (iii) on-chain collateralised stablecoins; and (iv) algorithmic stablecoins. Stablecoins with a claim or liability attached share the function and characteristics of financial instruments, therefore would fall into the tokenised financial instruments category.
21 BIS (2021)
23 Mollick (2014)
25 There are also other ways to further subdivide crowding platforms, e.g. generalist vs. thematic platforms, or depending on the funding mechanism, e.g. all or nothing or keep it all.
Table 1. Crowdfunding models

<table>
<thead>
<tr>
<th>Investment models</th>
<th>Non-investment models</th>
</tr>
</thead>
<tbody>
<tr>
<td>- lending-based crowdfunding: funds are paid back, and funders have the right to receive a contractually agreed interest payment; includes peer-to-peer lending (P2P) and social lending (typically at local level)</td>
<td>- reward-based crowdfunding: funds are provided in exchange for non-monetary benefits (e.g. a small gift (reward), a pre-order)</td>
</tr>
<tr>
<td>- equity-based crowdfunding: funds are provided in exchange for company’s shares.</td>
<td>- donation-based crowdfunding: funds are provided for philanthropic or sponsorship reasons with no expectation of remuneration.</td>
</tr>
<tr>
<td>- other: funds are provided in exchange for company’s bonds e.g. mini-bonds, or in exchange of some hybrid instruments.</td>
<td></td>
</tr>
</tbody>
</table>

Second, supply chain finance (invoice financing, reverse factoring) constitutes an arrangement between a buyer, a supplier, and a financial intermediary where the creditworthiness of a buyer is used to improve the working capital position of a supplier\(^2\). Typically, such arrangements involve a large, financially strong company that is supplied by several SMEs, and a financial intermediary – often a bank. Fintechs in this domain operate as cloud-based software platforms and can enable “procure-to-pay” systems that incorporate both purchasing management and accounts payable functionalities. Fintechs may often take on some exposures to risk. In the future, some Fintechs may extend their services beyond financing the supply-chain, by offering related services such as procurement and supplier management. \(^2\)

Due to increasing importance both crowd-funding and new forms of supply-chain financing could considered novel and could be classified e.g. in K66.1X and K64.9X of the current NACE respectively.

Fintech related to investment, asset management, and trade

This category encompasses various activities facilitating investment activities covering a) social trading platforms, b) robo advice, c) personal financial management and d) other. **Social trading offers investing strategies that use copy trading or mirror trading, based on following the investment behaviours of peers and expert traders.** Social trading platforms facilitate connections within an online community of investors, in which users can fully observe and automatically, simultaneously, and unconditionally replicate investment strategies of other users based on relatively low costs\(^2\). **Robo advice constitutes an online automated investment platform that uses quantitative algorithms to manage investors’ portfolios.** Robo advice covers a wide range of digital (semi-)automatic investment platforms and services. Robo advice has evolved from online “manual” questionnaires and proposals to automated portfolio management using quantitative methods and algorithms to construct and rebalance portfolios. The latest robo advice systems cover the entire investment/portfolio management process, starting from the selection of the instrument universe and finishing with periodic portfolio rebalancing and appropriate performance reporting. **Personal financial management covers private financial**

\(^2\) [https://ec.europa.eu/growth/content/supply-chain-finance_en](https://ec.europa.eu/growth/content/supply-chain-finance_en)

\(^2\) Rogers D. et all (2016)

\(^2\) Reith R. et all (2020)
planning, in particular the administration and presentation of financial data (also from various accounts) using software or app-based services. To integrate the accounts of different providers into a personal financial management system the application programming interface (API) technology is frequently used. Personal financial management may overlap with some (Fintech) services in the payment and insurance service area (see below). Other activities in this section include e.g. online asset management platforms, deposit brokers and online trading platforms. Online based asset management platforms offer a combination of human and robo advice. Deposit brokers exploit differences in the interest rates from various countries e.g. in the EU and offer the opening and managing of accounts on one single website in a chosen/domicile country. While deposit brokers may fit well into the investment, asset management and trade section, they may often be considered among the Fintech activities related to banking as they can offer traditional banking products such as a cash account with certain IT functionalities. In the context of identifying novel activities the feature of robo-counterparty stands out and such activities could be classified in e.g. K66.X in the current NACE. The other activities can be considered as existing ones done in an innovative way; hence they should be covered in K66.30 or K66.19.

Fintech related to payment services

Very often blockchain and crypto-asset solutions are mentioned also in the payment service category as crypto-assets might be reused in other activities; however, such classification – if exclusive – does not capture the full scope of crypto-assets. Furthermore, the use of crypto-assets as a means of payments is largely contained within the crypto-asset ecosystem for the moment. Therefore, it is envisaged to analyse the various Fintech-related payment activities separately. Fintech activities in the payment services segment include e.g.: mobile payments, digital wallets, Peer-to-Peer payments and others (see Table 2). These payment activities are considered novel and could be classified in K66.1X: PayTech in the current NACE.

Table 2. PayTech examples

<table>
<thead>
<tr>
<th>Mobile payments (mobile finance)</th>
<th>Digital wallets (eWallets, cyberwallets)</th>
<th>Peer-to-Peer payments</th>
<th>Other such as:</th>
</tr>
</thead>
<tbody>
<tr>
<td>various functionalities that are handled via mobile phones, includes to make payments or bank transfers</td>
<td>systems in which both means of payment and payment information for various payment systems can be stored, thereby facilitating online payment.</td>
<td>electronic transfer of money</td>
<td>real-time payments, atomic payments, payments as you go, payments using QR codes or utilising Near Field Technology (NFT), payments initiation, novel cross-border remittances and transactions,</td>
</tr>
</tbody>
</table>

29 Dorfleitner G. et al (2017)
31 see Mallat N. (2007); Mallat N. et al. (2004); Merritt C. (2010)
32 Mjølsnes S. F. and Rong C. (2003); Mallat N. (2007)
33 P2P payments in the Fintech category would cover both emerging and mature markets. In this context it is worth recalling that there are three key differences on the reasoning behind P2P money transfers.
Fintech in NACE: Suggestions and tentative figures for Germany, France and Spain

| microfinance services, e-payments and loyalty cards, novel payroll services |

Digital-only banks, Quasi-banks and BankTech

Fintech activities related to banking are part of a trend towards the unbundling of financial services. Technological advances facilitated the emergence of a new type of bank that is fully digital and runs its operations online and/or through a mobile application – the digital-only bank (which also may be called neobank, nubank, online bank, internet-only bank, virtual bank, digital bank or challenger bank). Fintech banking activities might be provided also by entities without a banking licence; for the purposes of this note such entities are labelled quasi-bank. Finally, Fintech activities related to banking may cover embedded finance or the provision of tools for enhanced bank customer experience, fraud security, background checks or credit decisions, typically using the services of a bank or quasi-bank acting in the background. The activities of digital-only banks are considered to a large degree as existing activities done in an innovative way and existing NACE sub-sections (K64.19) already cover them in substance. On the other hand, the activities of quasi-banks and auxiliary activities are considered novel which could be classified as new K64.9X Quasi-banks and K66.1X Bank Tech respectively.

InsurTech and PensionTech

Fintech in this domain covers services that use, for example, big data and AI, chatbots, customisable insurance policies (e.g. pay as you go), or P2P- or crowd-surance. Fintech activities related to insurance and pension funding are considered largely existing activities done in an innovative way and could be included in NACE K66.29. Nevertheless, due to increasing importance InsurTech and PensionTech could earn a dedicated sub-entry eventually.

Tech facilitators

Tech facilitators provide infrastructure solutions, based e.g. on DLT, AI, the Internet of Things (IoT) and big data technologies. Some examples of infrastructural activities are the production of the hardware required for some Fintech activities (e.g. crypto-

when comparing emerging and mature markets: 1) cash vs. digital value stored in e.g. banks and credit cards as starting point, 2) need for a swift conversion/convertibility of stored value into physical cash and 3) lack of access to formal financial tools like banks; See Matsumoto M., Terrenghi L. (2016)

34 Matousek R., Xiang D. (2021)

35 Technologies that Tech facilitators leverage upon may include Cyber Security Technologies (Biometrics, cryptographic algorithms, etc.), Data Analytics Applications (based on machine learning and data science), Cloud and Software-as-a-Service (platform technology), NoSQL (Graph Databases, and other innovative ways of storing and accessing data), W3C Internet Standards, HTML, XML (XBRL), RDF, OWL, SWRL SQL Technologies, Data lakes, Hadoop; API Application Programming Interfaces, etc; see European Commission: Expert Group on Regulatory Obstacles to Financial Innovation (ROFIEG): 30 Recommendations on Regulation, Innovation and Finance, Belgium (2019). https://ec.europa.eu/info/files/191113-report-expert-group-regulatory-obstacles-financial-innovation_en
asset mining hardware) as well as software or apps. These activities that facilitate Fintech are neither strictly financial intermediation nor financial auxiliary services, as in many cases they are scalable and can be applied in non-financial sectors, therefore they should be classified in NACE outside section K.

**SupTech and RegTech – borderline cases**

SupTech services are dedicated to improving surveillance and analytical capabilities on the part of supervisors and regulators. SupTech encompasses technologies aimed at automating and streamlining administrative and operational procedures, digitising data and working tools, and improving data analytics. Some SupTech solutions enable regulators to prepare and transmit machine-readable and machine-executable regulatory documents to their regulated entities, which in turn could result in more automated regulatory compliance, lower costs and greater consistency in regulatory reporting. Other SupTech solutions are focused on achieving real-time risk alerts, thereby enabling supervisory teams to pivot attention towards pre-emptive rather than curative oversight, in turn possibly improving the resilience and stability of the broader financial system. **Aimed at regulated institutions, the use of RegTech improves compliance outcomes.** RegTech solutions encompass the areas of regulatory reporting, risk management, identity management & control, compliance, and transaction monitoring. RegTech activities cover, for example, automated data distribution and regulatory reporting through big data analytics; real-time reporting and cloud solutions; sanctions screening/watchlist filtering and regulatory reporting capabilities to comply with AML and CTF regulations; and Know Your Customer (KYC) procedures.

3. **Fintech in statistical classifications from a central bank perspective: the case of NACE rev.2**

This section provides a summary of the proposed classification of novel Fintech activities set out in the previous section referring to the coding framework of the current NACE.

**K – Financial and insurance activities**

- K64 – Financial service activities, except insurance and pension funding n.e.c.
  - K64.1 – Monetary intermediation
    - K64.11 – Central banking
    - K64.19 – Other monetary intermediation
  - K64.2 – Activities of holding companies

---

36 BIS on SupTech and RegTech: https://www.bis.org/about/bisih/topics/suptech_regtech.htm

K64.3 – Trusts, funds and similar financial entities
K64.9 – Other financial service activities, except insurance and pension funding
  K64.91 – Financial leasing
  K64.92 – Other credit granting
  Proposal K64.9X – New forms of supply chain financing activities
  Proposal K64.9X – Quasi-banking
  K64.99 – Other financial service activities except insurance and pension funding n.e.c.\(^{38}\)

K66 – Activities auxiliary to financial services and insurance activities
  K66.1 – Activities auxiliary to financial services, except insurance and pension funding
    K66.11 – Administration of financial markets
    K66.12 – Security and commodity contracts brokerage
    Proposal K66.1X – BankTech
    Proposal K66.1X – PayTech
    K66.19 – Other activities auxiliary to financial services, except insurance and pension funding
  K66.2 – Activities auxiliary to insurance and pension funding
    K66.21 – Risk and damage evaluation
    K66.22 – Activities of insurance agents and brokers
    K66.29 – Other activities auxiliary to insurance and pension funding
  K66.3 – Fund management activities
  Proposal K66.X Crowd-funding activities
  Proposal K66.X Robo financial activities
  Proposal K66.X Crypto-Assets activities n.e.c.

Some of the existing financial activities done in a novel way would not necessarily require new subitems in NACE. This applies to the Fintech activities related to security dealing platforms and asset management (in K66.30 or K66.19), digital-only banks (in K64.19), or InsurTech and PensionTech (in 66.29). For these activities, improving the explanatory notes to clarify the content of the sub-classes would facilitate the classification in cases of doubt and thereby harmonise and improve statistics. The explanatory notes should also elaborate that the borderline cases of SupTech and RegTech should ideally be considered as financial on account of supporting financial activities in K, and that the activities of Tech facilitators should be classified together with non-financial activities.

\(^{38}\) N.e.c.: not elsewhere classified
4. Fintech Statistics at the Bank of Spain

As in other countries, Spain does not have an official register of Fintech firms, since some of their corporate purposes do not need to be registered by a supervisory authority. This, together with the ongoing innovations in this area, has hindered the preparation of an exhaustive census of Fintech firms. Therefore, Banco de España (BdE) has to draw information from various public and private sources: the Spanish National Securities Market Commission (CNMV), business associations (the Spanish Fintech & InsurTech Association and the Spanish Crowdlending Association) and private consulting firms (Finnovating).

This Fintech project, counting now around five years of existence, produces a set of information based on the sample of identified Fintech firms that should be regarded as experimental statistics. Nevertheless, this dataset is considered good enough to obtain an initial characterization of Spanish Fintech firms.

The BdE study shows that the Fintech sector is a rapidly growing market in Spain. According to the most recent available data, the sample of Spanish Fintechs has grown from 77 in 2012 to 328 in 2021, registering an annual 17.5% increase in this nine-year period. Similarly, the number of employees rose from around 600 in 2012 to almost 5,000 in 2021, with an annual growth around 26%.

To assess the relevance of these figures, it should be borne in mind that they could underestimate the true importance of the Fintech activity in Spain for three reasons. First, it does not take into account the Fintech activity developed inside the banking industry; second, the census of Fintech firms is without any doubt incomplete; and third, it does not record the activity performed by non-resident Fintech firms in the Spanish market.

In relation to this last point, numerous non-resident firms providing Fintech services in Spain have been identified by our experimental investigations, but the lack of detailed financial information prevents their inclusion in our dataset with quantitative figures. As a matter of fact, the digital provision of Fintech services (often without a permanent establishment in the country where the firms are operating) and the cross-border characteristics of the Fintech industry (specially for BigTechs) demand a supranational perspective for the data collection on Fintech activity or, at least, a quite high coordination across countries.

All in all, the growth rates of Fintech activities in Spain, as well as in most developed countries, are expected to accelerate in the forthcoming years, thus increasing the relevance of this activity in the financial sector. Having this in mind, there is an urgent need to improve our statistical approach to this sector, including the development of a homogeneous treatment which would allow for reliable comparisons across economies.

In this sense, an adequate inclusion of Fintech activities in the economic activity classification should be considered as a fundamental prerequisite to produce better

---

39 Although not collecting information of Fintech credit activities, Banco de España has a non-exhaustive collection of Fintech firms and produces experimental statistics on them, quantifying their economic relevance and classifying in different Fintech activities. In addition, for a subset of these firms, annual detailed accounting information is available for research purposes at BELab (Banco de España’s Data Laboratory). Further details are available under the link: Banco de España - Economic analysis and research - What is BELab? - Content - Available microdata - Fintech no bancarias (FIN) (bde.es).
statistics. In the case of Spain, the main segment of this sector is the activity of lending and obtaining funding on online platforms (crowdfunding) - with 132 active companies, 2,220 employees and a turnover of more than 551 million euros - , followed by the provision of technology services (see Table 3).

Table 3. Size dimensions of suggested new classes in K for Spain

<table>
<thead>
<tr>
<th>New NACE Suggestion</th>
<th>No of Fintechs</th>
<th>No of Employees</th>
<th>Turnover (€)</th>
</tr>
</thead>
<tbody>
<tr>
<td>K66.X: crypto-assets</td>
<td>10</td>
<td>39</td>
<td>59,574,836</td>
</tr>
<tr>
<td>K66.X: Crowd-funding activities</td>
<td>132</td>
<td>2,220</td>
<td>551,244,418</td>
</tr>
<tr>
<td>K64.9X: New forms of supply-chain-financing</td>
<td>23</td>
<td>223</td>
<td>199,280,806</td>
</tr>
<tr>
<td>K66.X Robo financial activities</td>
<td>68</td>
<td>1,195</td>
<td>93,476,836</td>
</tr>
<tr>
<td>K66.1X: PayTech</td>
<td>55</td>
<td>877</td>
<td>91,773,910</td>
</tr>
<tr>
<td>Neobanks/BankTech</td>
<td>4</td>
<td>43</td>
<td>11,259,701</td>
</tr>
<tr>
<td>InsureTech</td>
<td>32</td>
<td>398</td>
<td>37,611,075</td>
</tr>
<tr>
<td>Suptech/RegTech</td>
<td>4</td>
<td>129</td>
<td>16,223,761</td>
</tr>
<tr>
<td>Totals</td>
<td>328</td>
<td>5,124</td>
<td>1,060,445,343</td>
</tr>
</tbody>
</table>

Note: BdE provisional data.

Analysing how these Fintech firms classify themselves in the national economic activity classification shows that very often they are included outside NACE section K (Financial Activities) and, unfortunately, they are spread across more than ten different groups. This result could be related to the lack of Fintech details in the current NACE, which makes both the identification of Fintech firms and a correct assessment of their activity very difficult. A more detailed sub classification in the section K of the NACE together with clear explanatory notes could solve these identification problems, making it easier to achieve an effective and more useful classification.

5. Experimental Data at the Deutsche Bundesbank

Statistical work on Fintech is notoriously difficult, not least because their activities do not fit current classification schemes and firms are hard to identify. Furthermore, there are no comprehensive reporting obligations that could be the basis of an encompassing data base. Hence, in a model project on Fintech data, Bundesbank Department of Statistics collects what is internally available and what can be obtained freely in electronic media, enhancing it with selected commercial information. This experimental data collection, a proof of concept, is labelled Fintech monitor.

The Fintech monitor has two main functions. First, it is to set up as a data hub for various Bundesbank departments working on Fintech activities. Thus, the Fintech monitor provides a platform to store and retrieve the collected pieces of information across the bank. Second, it is designed to become a basis for statistical data work. Accordingly, the Fintech monitor collects Fintech IDs, references and classification data, information on business models and basic data on commercial activity, such as turnover and number of employees.
As of October 2022, the Fintech monitor collects information on 1,227 companies with Fintech activities in Germany, of which 1,062 companies are active. Among the active companies, 939 ones are resident in Germany. It needs to be pointed out that important Fintech activities are carried out by BigTech companies. Typically, the commercial focus of BigTechs is outside finance, and they often offer Fintech services as a by-product or secondary activity. Furthermore, a larger part of Fintech activity is to be attributed to companies outside Germany. Thus, activities of Fintech companies in Germany will give only part of what is relevant.

The Fintech monitor allows to give a non-comprehensive and incomplete assessment on the number of active units located in Germany which would fall into the NACE categories suggested in this paper. Crypto-asset activities are the focus of 48 German companies. Looking at Fintech financing first, crowd-funding and new form of supply chain financing are offered by 148 and 25 companies respectively. Furthermore, the Fintech Monitor identifies 94 companies providing robo-financial activities. PayTech has expanded a lot in recent years - currently 112 companies have been identified in this segment. Finally, there are 18 BankTechs, 77 InsurTech/PensionTech companies and 24 companies that are either SupTech or RegTech (see Table 4).

Table 4. Identified active Fintech companies resident in Germany for suggested categories

<table>
<thead>
<tr>
<th>New NACE Suggestion</th>
<th>No. of Fintechs</th>
<th>No. of Employees</th>
<th>Turnover ($)</th>
</tr>
</thead>
<tbody>
<tr>
<td>K65.X: Crypto-assets</td>
<td>43</td>
<td>137</td>
<td>20,803,034</td>
</tr>
<tr>
<td>K65.1X: Crowd-funding activities</td>
<td>148</td>
<td>3150</td>
<td>690,196,527</td>
</tr>
<tr>
<td>K64.0X: New forms of supply-chain financing</td>
<td>25</td>
<td>515</td>
<td>133,234,880</td>
</tr>
<tr>
<td>K65.X: Robo financial activities</td>
<td>94</td>
<td>2,249</td>
<td>391,736,923</td>
</tr>
<tr>
<td>K65.1X: PayTech</td>
<td>112</td>
<td>4,240</td>
<td>806,180,218</td>
</tr>
<tr>
<td>K65.1X: InsurTech/PensionTech</td>
<td>18</td>
<td>1,058</td>
<td>68,293,399</td>
</tr>
<tr>
<td>SupTech/Regtech</td>
<td>24</td>
<td>2,372</td>
<td>317,331,452</td>
</tr>
<tr>
<td><strong>Totals</strong></td>
<td>548</td>
<td>14,887</td>
<td>2,543,422,853</td>
</tr>
</tbody>
</table>

Note: October 2022. All numbers are preliminary. Data on number of employees and turnover are to be considered lower bounds, because these data are not available for all companies.

All numbers displayed in table 4 should be interpreted with caution and are to be considered lower bounds. Not all of the relevant companies are known, and not from all the known companies there are data on the number of employees or turnover that could be displayed. At this point, no attempt was made to estimate or impute the missing data. Furthermore, much of the important activity is carried out by firms located outside Germany and doing business with German residents. Big crypto-asset exchanges and service providers are a good example of this issue, which is important when estimating a crypto-asset market size in Germany. It is clear that cross-border

---

40 Data sources used for this analysis: internal data, North Data (North Data Smarte Recherche), TheBanks.eu (The European Banks (thebanks.eu)) and CompanyHouse (Handelsregister- und Wirtschaftsinformationen - CompanyHouse)

41 All numbers are preliminary.

42 We consider as active firms those that are not in liquidation, and have not exited the market or closed their business.
cooperation among authorities is very important when trying to understand and capture the Fintech phenomenon and its dimensions.

6. Fintech statistics at the Banque de France

As for other central banks, there is no mandatory data collection for statistics on Fintechs in Banque de France (BdF). In order to set up a census on French Fintechs a twofold approach was followed. First, professional associations, with valuable expertise in the field, such as France FinTech or Finance Innovation have been approached for assessing the best practices. Secondly, knowledge from internal experts and regulators in charge of issuing licences were mobilised in order to link practices to the core business of National Central Banks (NCBs).

The establishment of an NCB compatible census opens the room to a data collection (number of employees, turnover, etc.) based on internal resources provided by a BdF’s department in charge of analysing enterprises data. The collaboration of the national statistical institute (INSEE) was also necessary for accessing the “SIRENE” database which is a comprehensive register for companies in France.

The characterisation of the activity of each Fintech entity was carried out by investigating their websites and by exchanging with the professional associations mentioned above. This step-by-step work has allowed an initial characterisation of French Fintechs according to the breakdown proposed to fulfil section K of the NACE (see Table 5).

Table 5. Size dimensions of suggested new classes in K for France

<table>
<thead>
<tr>
<th>New NACE Suggestion</th>
<th>Nº of Fintech</th>
<th>Nº of Employees</th>
<th>Turnover (£)</th>
</tr>
</thead>
<tbody>
<tr>
<td>K66.X: Crypto-asset activities</td>
<td>5</td>
<td>49</td>
<td>46 508 000</td>
</tr>
<tr>
<td>K66.X: Crowd-funding activities</td>
<td>28</td>
<td>707</td>
<td>38 387 000</td>
</tr>
<tr>
<td>K64.9X: New forms of supply-chain-financing</td>
<td>12</td>
<td>516</td>
<td>5 599 000</td>
</tr>
<tr>
<td>k66.X: Robo financial activities</td>
<td>12</td>
<td>232</td>
<td>4 647 000</td>
</tr>
<tr>
<td>K66.1X: PayTech</td>
<td>50</td>
<td>1 840</td>
<td>537 552 000</td>
</tr>
<tr>
<td>K66.1X: BankTech</td>
<td>20</td>
<td>652</td>
<td>39 892 000</td>
</tr>
<tr>
<td>InsureTech</td>
<td>36</td>
<td>732</td>
<td>23 039 000</td>
</tr>
<tr>
<td>SupTech/Regtech</td>
<td>21</td>
<td>601</td>
<td>26 372 000</td>
</tr>
<tr>
<td><strong>Totals</strong></td>
<td><strong>184</strong></td>
<td><strong>5 329</strong></td>
<td><strong>721 996 000</strong></td>
</tr>
</tbody>
</table>

It should be noted that the above figures are certainly underestimating the scale of Fintech companies in France. As a first caveat, it should be mentioned that the list of companies is not comprehensive, as it only corresponds to resident Fintechs and exclude foreign companies operating in France. In addition, a number of data points on turnover were missing. Moreover, the census of entities dates to early 2020, hence many enterprises have emerged and grown since then.

However, despite those caveats, the work done so far is an important and vital step to assess the importance of each activity segments.

Therefore, the PayTech sector appears to be the most preponderant in France with 50 active companies, 1 840 employees and a turnover of more than 530 million euros. This segment of Fintech activity is also the one that attracts the most investors, with
nearly €1.7 billion of funds raised since 2010. This sector expands extensively in France particularly by enlarging its service supplies. For example, the French unicorn *Lydia* that started with providing peer-to-peer payment solutions is now offering investments in crypto-assets and stocks, savings books, consumption loans, contactless mobile payments, online payments, virtual cards, etc.

The Fintech sector is expected to continue growing in the forthcoming years, since the investors’ appetite for such companies, especially in the form of foreign direct investments, remains strong. Due to their innovative services, the ease of access to their offers, their very competitive prices and with the ever-increasing digitalisation of our society, it is certain that Fintechs will take an ever more important place in the financial sector. It is therefore essential to be able to characterise the current activity of Fintechs in order to identify them more easily (hence our breakdown proposal based on the example for the sector K of NACE), as well as to collect information to monitor and analyse their activity and eventually to capture them in the NCBs statistical systems.

**Sources**

- BIS on Suptech and RegTech: [https://www.bis.org/about/bisih/topics/suptech_regtech.htm](https://www.bis.org/about/bisih/topics/suptech_regtech.htm)
- BIS (2021), DeFi risks and the decentralisation illusion, BIS Quarterly Review, December 2021, special feature by S. Aramonte, W. Huang and A. Schrimpf
- Dorfleitner G. et all (2017), FinTech in Germany, Cham: Springer

---

43 Microsoft PowerPoint - Semestre de la Fintech 2022 (fintech-metrix.com)
• Deutsche Bundesbank (2022), Fintech services from a central bank perspective, A segmentation as suggested for the Bundesbank Fintech Monitor, internal note
• Kochanska U. (2020), Characterisation of the euro area fintech scene, Financial Integration and Structure in the Euro Area, ECB
• Mallat N. et all (2004), Mobile Banking Services, Communications of the ACM 47(5):42-46
• Menon R. (2021), Decentralised finance and the future of money, BIS


• Reith R. et all (2020), Explaining the intention to use social trading platforms: an empirical investigation. J Bus Econ 90, 427–460

Fintech in statistical classifications: suggestions and tentative figures in a central bank context

25 August 2022

Enzo Buthiot (Banque de France), Celestino Giron (ECB), Luis Ángel Maza (Banco de España), Ulf von Kalckreuth (Deutsche Bundesbank), Urszula Kochanska (ECB), Román Santos (Banco de España), Yann Wicky (Banque de France), Norman Wilson (Deutsche Bundesbank).

11th Biennial IFC Conference
“Post-pandemic landscape for central bank statistics”
Overview

1. Motivation
2. Statistical classification of Fintech: the case of NACE rev.2
3. Experimental Fintech data at BdE, BuBa and BdF
4. Conclusions
1. Motivation

- In the last decade, **innovation activity in the financing industry** has largely been moving to entities commonly labelled **Fintech** amid a wider phenomenon of decentralised finance.
- There is **no internationally accepted harmonised definition of Fintech** for statistical or for other classification purposes.
- This paper **clarifies the scope of (current) Fintech** to support their classification in statistical systems, thereby enabling their monitoring.
- The possibility of introducing new classification entries is discussed, taking as a point of departure the current NACE hierarchical structure for economic activities and distinguishing within the Fintech universe between novel financial activities, old financial activities done in a new way, and activities which are not financial.
- **Context** → IFC Report *Towards monitoring financial innovation in central bank statistics* and the *New G-20 Data Gap Initiative (DGI) in preparation*.
- **Input** → to any discussions on Fintech in the context of statistical classifications e.g. NACE, ISIC and the classifications of products (CPC/PA).
2. Statistical classification of Fintech: the case of NACE rev.2

- **Crypto-asset activities**, especially related to No-Liability Crypto-Assets (NLCAs), including emission, issuance, operating (validating), and services are considered novel and should warrant a dedicated entry in statistical classifications e.g. in K66 of the current NACE. Importantly, crypto-asset may be used in other activities e.g. in payments.

- Fintech related to financing (or services auxiliary to financing) which includes **crowdfunding** and new **forms of supply chain financing**. Due to increasing importance, new entries are warranted e.g. in K66.1X and K64.9X of the current NACE respectively.

- Fintech related to investment, asset management, and trade covers: a) **social trading platforms**, b) **robo advice**, c) **personal financial management** and d) other e.g. online asset management platforms, deposit brokers and online trading platforms. Robo-advice stands out in terms of novelty and could be classified in e.g. K66.X of the current NACE. The other activities can be considered as existing ones done in an innovative way; hence they should be covered in K66.30 or K66.19.

- Fintech activities in the payment services segment include e.g.: **mobile payments, digital wallets, Peer-to-Peer payments and others such as real-time payment, atomic payments.** These payment activities are considered novel and could be classified in K66.1X: PayTech of the current NACE.
2. Statistical classification of Fintech: the case of NACE rev.2

• **Digital-only banks, Quasi-banks and BankTech.** The activities of digital-only banks are considered to a large degree as existing activities done in an innovative way and NACE sub-section (K64.19) already covers them in substance. The activities of quasi-banks and related auxiliary activities are considered novel and could be classified as new K66.1X Bank Tech and K64.9X Quasi-banks respectively.

• **InsurTech and PensionTech.** Fintech in this domain covers services that use e.g. big data and AI, chatbots, customisable insurance policies (e.g. pay as you go), or crowdsurance. Such activities are considered as existing activities done in an innovative way and could be included in NACE K66.29.

• **Tech facilitators** provide infrastructure solutions, based e.g. on DLT, AI, the Internet of Things (IoT) and big data technologies, however they are neither strictly financial intermediation nor financial auxiliary services, therefore they should be classified in NACE outside section K.

• **Borderline cases:** SupTech (dedicated to improve surveillance and analytical capabilities of supervisors and regulators) and RegTech (aimed at regulated institutions, improves compliance outcomes).
2. Statistical classification of Fintech: the case of NACE rev.2

K – Financial and insurance activities
K64 – Financial service activities, except insurance and pension funding n.e.c.
K64.1 – Monetary intermediation
K64.11 – Central banking
K64.19 – Other monetary intermediation
K64.2 – Activities of holding companies
K64.3 – Trusts, funds and similar financial entities
K64.9 – Other financial service activities, except insurance and pension funding
K64.91 – Financial leasing
K64.92 – Other credit granting
Proposal K64.9X – New forms of supply chain financing activities
Proposal K64.9X – Quasi-banking
K64.99 – Other financial service activities except insurance and pension funding n.e.c.

K66 – Activities auxiliary to financial services and insurance activities
K66.1 – Activities auxiliary to financial services, except insurance and pension funding
K66.11 – Administration of financial markets
K66.12 – Security and commodity contracts brokerage
Proposal K66. 1X – BankTech
Proposal K66. 1X – PayTech
K66.19 – Other activities auxiliary to financial services, except insurance and pension funding
K66.2 – Activities auxiliary to insurance and pension funding
K66.21 – Risk and damage evaluation
K66.22 – Activities of insurance agents and brokers
K66.29 – Other activities auxiliary to insurance and pension funding
K66.3 – Fund management activities
Proposal K66.X Crowd-funding activities
Proposal K66.X Robo financial activities
Proposal K66.X Crypto-Assets activities n.e.c.
### 3. Experimental Fintech data at BdE, BuBa and BdF

**Statistical work on Fintech is notoriously difficult**, not least because Fintech activities do not fit current classification schemes and firms are hard to identify. There are no comprehensive reporting obligations that could be the basis of an encompassing database, there are no official registers of Fintech firms.

<table>
<thead>
<tr>
<th>Output:</th>
<th>experimental statistics for a sample of identified Fintech firms (350 firms)</th>
<th>IDs, references and classification data, information on business models and basic data on commercial activity, such as turnover and number of employees (1227 companies with Fintech activities in Germany of which 939 resident companies)</th>
<th>number of employees, turnover, etc. based on internal resources used for analysing enterprises’ data; data from the so-called “SIRENE” database (a kind of comprehensive register for companies in France) thanks to the collaboration of the national statistical institute (INSEE), (184 companies)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Input:</td>
<td>various public and private sources: the Spanish National Securities Market Commission (CNMV), business associations (the Spanish Fintech &amp; InsurTech Association and the Spanish Crowdlending Association) and private consulting firms (Finnovating)</td>
<td>Bundesbank statistics collects what is internally available enhancing it with selected commercial information</td>
<td>1) professional associations, with valuable expertise in the field, such as <em>France FinTech</em> or <em>Finance Innovation</em> have been approached for assessing the best practices, 2) knowledge from internal experts and regulators in charge of issuing licences were mobilised in order to link practices to NCBs core business, 3) exploring Fintech’s websites to characterise the activity of each Fintech</td>
</tr>
</tbody>
</table>
3. Experimental Fintech data at BdE, BuBa and BdF

<table>
<thead>
<tr>
<th>New Nace Suggestion</th>
<th>Spain</th>
<th>Germany</th>
<th>France</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Number of Fintechs</td>
<td>Number of employees</td>
<td>Turnover (€ million)</td>
</tr>
<tr>
<td>K66.X: Crypto-asset activities</td>
<td>10</td>
<td>39</td>
<td>60</td>
</tr>
<tr>
<td>K66.X Crowd-funding activities</td>
<td>132</td>
<td>2,220</td>
<td>551</td>
</tr>
<tr>
<td>K64.9X: New forms of supply-chain-financing</td>
<td>23</td>
<td>223</td>
<td>199</td>
</tr>
<tr>
<td>K66.X: Robo financial activities</td>
<td>68</td>
<td>1,195</td>
<td>93</td>
</tr>
<tr>
<td>K66.1X: PayTech</td>
<td>55</td>
<td>877</td>
<td>92</td>
</tr>
<tr>
<td>K66.1X: BankTech</td>
<td>4</td>
<td>43</td>
<td>11</td>
</tr>
<tr>
<td>InsurTech/PensionTech</td>
<td>32</td>
<td>398</td>
<td>38</td>
</tr>
<tr>
<td>SupTech/RegTech</td>
<td>4</td>
<td>129</td>
<td>16</td>
</tr>
<tr>
<td>Total</td>
<td>328</td>
<td>5,124</td>
<td>1,060</td>
</tr>
</tbody>
</table>

Notes: ES – data for 2021, figures may underestimate the true importance the Fintech activity for three reasons: i) Fintech activity is also inside the banking industry, ii) incomplete sample, iii) missing info on activity performed by non-resident Fintech firms in the Spanish market (investigated, but lack detailed information); DE – based on preliminary, non-comprehensive assessment of Fintech entities resident in Germany, important Fintech activities are due to BigTechs and quite a large part of Fintech activity is attributed to companies outside of Germany, thus not included in this analysis, data compiled in August 2022; FR – data for 2020, figures underestimate the scale of Fintech phenomenon as the census is not comprehensive (resident Fintechs only, foreign companies operating in France are excluded) and a number of data on turnover were missing.
4. Conclusions

• **Fintech activities are expected to grow significantly** in the forthcoming years, also considering the substantial attraction of investors, thus increasing the relevance of this activity in the financial sector → there is an urgent need to improve our statistical approach to this phenomenon, including the development of a homogeneous treatment which would allow for reliable comparisons across economies.

• Fintech firms are often classified outside NACE section K Financial Activities (spread across more than ten different NACE sections). This could be related to the lack of Fintech details in the current NACE, which makes the identification of Fintech firms and a correct assessment of their activity very difficult → a more detailed Fintech classification in statistical systems e.g. the sub classification in the NACE section K together with clear explanatory notes, could solve these identification problems and make it easier to achieve an effective and more useful classification.

• This paper clarifies the scope of (current) Fintech activities and provides examples on how they could be classified in the current NACE section K. Furthermore, it aims to support any future discussion on Fintech classification in statistical systems.
11th Biennial IFC Conference on “Post-pandemic landscape for central bank statistics”
BIS Basel, 25-26 August 2022

Issues in reflecting digital assets in the Flow of Funds Accounts¹

Yoshiko Sato,
Bank of Japan

¹ This presentation was prepared for the conference. The views expressed are those of the author and do not necessarily reflect the views of the BIS, the IFC or the central banks and other institutions represented at the event.
Issues in reflecting digital assets in the Flow of Funds Accounts¹

Yoshiko Sato²

Abstract

This paper proposes an idea of how digital assets could be included in the Flow of Funds Accounts. Digital assets which are used as a means of payment, including CBDCs, stablecoins, and other types of crypto-assets, can have a potential impact on the financial system and are therefore expected to be reflected in an appropriate, feasible and consistent manner within the existing statistical framework of the Flow of Funds Accounts. Reviewing international discussions on the statistical treatment of these digital assets, this paper looks at the issues regarding the definition and classification of these instruments and explains difficulties in the collection of source data.

Keywords: System of National Accounts, Flow of Funds Accounts, digital assets, crypto-assets, stablecoins, CBDCs

¹ I thank Yosuke Tada at the Cabinet Office, Akira Otani, Ko Nakayama, Tomohiro Sugo, Yuhei Shimizu, Yuichi Fujiwara, Makoto Saito, Haruko Kato, Shinya Nakamura, Kuniko Moriya and other colleagues at the Bank of Japan for their valuable comments. The views expressed in this article are those of the author and do not necessarily reflect those of the Bank of Japan.

² Research and Statistics Department, Bank of Japan  yoshiko.satou@boj.or.jp
## Contents

Issues in reflecting digital assets in the Flow of Funds Accounts .............................................. 1

1. Digital assets in the context of the Flow of Funds Accounts .............................................. 3
   1.1 Introduction .................................................................................................................... 3
   1.2 Criteria for financial assets ......................................................................................... 4
   1.3 Guidance note ............................................................................................................. 4
   1.4 Proposed classification ............................................................................................... 6

2. Reflecting digital assets in the Flow of Funds Accounts .................................................... 9
   2.1 Stablecoins ................................................................................................................ 9
   2.2 CBDCs.................................................................................................................... 9
   2.3 Security tokens ......................................................................................................... 10

3. Difficulties in the collection of source data ..................................................................... 12

References ............................................................................................................................. 13
1. Digital assets in the context of the Flow of Funds Accounts

Recent years have been an evolutilional age for digital assets. There is significant attention on digital assets, including crypto-assets—particularly on how they behave, who participate in transactions, and where they are actually held. However, with no common definition of digital assets, no consensus has been made on what should or should not be included in statistics. Against this backdrop, international organizations are working together with national compilers of macroeconomic statistics to update the system of national accounts (SNA)—towards the 2025 SNA—on how to include digital assets in macroeconomic statistics.

The Flow of Funds Accounts are compiled based on the SNA and serve as essential source data for compiling financial accounts in the SNA. Digital assets which are used as a means of payment, including CBDCs, stablecoins, and other types of crypto-assets, can have a potential impact on the financial system and are therefore expected to be reflected in an appropriate, feasible and consistent manner within the existing statistical framework of the Flow of Funds Accounts. Focusing on the Flow of Funds Accounts, this paper discusses the issues regarding the definition and classification of these instruments and explains difficulties in the collection of source data. Following a quick overview of digital assets, the first section introduces the international discussions about the statistical treatment of these digital assets and proposes an appropriate classification. The second section presents an idea of how digital assets could be recorded in the Flow of Funds Account. The last section explains difficulties in the collection of source data.

1.1 Introduction

While there is no common definition of the word of digital assets, this paper looks at a broad scope of digital assets, rather than focusing only on newer products such as crypto-assets. This approach has the advantage of enabling classification by clarifying the differences to and similarities of new products with the existing digital assets.

Before considering how to classify digital assets in the Flow of Funds Accounts, it is worthwhile looking at the common attributes of digital assets, and then reviewing the statistical standard on the treatment of financial assets.

Several common attributes are observed (Table 1). Some digital assets have an issuer (while others do not). They may be account-based or token-based in representing value and ownership. Some digital assets are exchanged on platforms using distributed ledger technology (DLT) and others are exchanged on platforms without DLT. Some digital assets are to be redeemed at a fixed rate of a certain asset (typically fiat currency) and others are variable at the time of redemption. Some may be used as a means of payment and others may not.

In the operation, some may not have any governing entity and/or transactions are conducted without permission (distributed); others may have a governing entity and/or permission is required in transactions (centralized). One of the novel features is the cryptographic validation technique. It is closely related to the above mentioned attributes such as DLT usage and governance/operation. Sometimes, the digital assets are classified in terms of whether the assets are issued by public or private
entities or, in a slightly different sense, whether the redemption is partially or fully guaranteed by the government.

<table>
<thead>
<tr>
<th>Attributes</th>
<th>Examples</th>
</tr>
</thead>
<tbody>
<tr>
<td>Issuer</td>
<td>Existence or absence of issuer. Liability of the issuer corresponding to the asset.</td>
</tr>
<tr>
<td>Type</td>
<td>Claim-based or object-based. Account-based or token-based.</td>
</tr>
<tr>
<td>DLT usage</td>
<td>Use of Distributed Ledger Technology (DLT).</td>
</tr>
<tr>
<td>Value</td>
<td>Redemption value is guaranteed by the issuer at a fixed rate of a certain asset (typically fiat currency), or variable at the time of redemption.</td>
</tr>
<tr>
<td>Mean of payment</td>
<td>Playing as a means of payment.</td>
</tr>
<tr>
<td>Governance / operation</td>
<td>Centralized or distributed. Permissioned or permission-less.</td>
</tr>
<tr>
<td>Cryptography</td>
<td>Use of cryptographic validation technique.</td>
</tr>
<tr>
<td>Public / private</td>
<td>The issuer is public or private entity. The redemption is partially or fully guaranteed by the government.</td>
</tr>
</tbody>
</table>


1.2 Criteria for financial assets

What are the attributes that matter in the compilation of the Flow of Funds Accounts? In other words, what kind of digital assets should be regarded as financial assets? Some of the crypto-assets seem to be used as a medium of exchange and as such one may want to classify them into financial assets. However, the question is not straightforward.

Financial assets in the SNA, in principle, should have a counterpart liability. The 2008SNA, the current version of the SNA, defines an asset as a store of value representing a benefit or series of benefits (¶3.5). Benefits are exchanged by means of payments. From this, a financial claim, and hence a liability, can be defined (¶3.32). A liability is established when one unit (the debtor) is obliged, under specific circumstances, to provide a payment or series of payments to another unit (the creditor) (¶11.5). ... Whenever either of these types of liability exists, there is a corresponding financial claim that the creditor has against the debtor. A financial claim is the payment or series of payments due to the creditor by the debtor under the terms of a liability (¶11.7). This means that in recording the Flow of Funds Accounts, which are based on the current SNA, all financial assets (the creditors) entail corresponding liabilities (the debtors), with an exception of monetary gold.

Recognizing this current SNA criteria for financial assets, all digital assets which seem to act as a means of payment will not necessarily be categorized in financial assets. Among the attributes mentioned above, special attention should be paid to the existence of issuer (liability).

1.3 Guidance note

As part of the updating process from the current 2008SNA towards 2025SNA, guidance notes for prioritized issues have been developed by the Intersecretariat Working Group on National Accounts (ISWGNA), receiving its mandate from the
United Nations Statistical Commission (UNSC). One of the issues included in the guidance notes is the recording of crypto-assets in macroeconomic statistics and this note calls for global consultation on this point. The guidance note proposes grouping crypto-assets into three broad categories, to discuss the criteria of classifying them into financial or non-financial assets.

- Crypto-assets designed to act as a general medium of exchange
  - with a corresponding liability:
    - issued by a monetary authority (e.g., central bank digital currencies (CBDCs) that qualify as crypto-assets)
    - not issued by a monetary authority (e.g., stablecoins with a claim on the issuer)
  - without a corresponding liability (CAWLM) (e.g., crypto-assets such as Bitcoin)
- Crypto-assets that only act as a medium of exchange within a platform or network (i.e., payment tokens)
  - with a corresponding liability
  - without a corresponding liability (CAWLP)
- Security tokens (which always have a counterpart liability)
  - Debt security crypto-assets (e.g., Bond-i issued by the World Bank); this also includes utility tokens that provide the holders future access to goods or services.
  - Equity crypto-assets
  - Derivative crypto-assets (i.e., derivative contracts that rely on cryptography and that can be exchanged peer-to-peer even if the underlying asset is not a crypto-asset).

Firstly, it is important to point out that the CBDCs are not "crypto-assets" (G7 (2021)). Recent literature defines cryptocurrencies or crypto-assets, which are used interchangeably, as "private digital assets with their own currency unit of account, such as Bitcoin and Ethereum. Cryptocurrencies do not represent a claim on a central bank, which makes them different from CBDCs" (Kosse and Mattei (2022)). Since CBDCs are not crypto-assets, it is appropriate to use a wider terminology of "digital assets" instead of "crypto-assets."

Secondly, the above categories in the guidance note puts the highest priority on whether the digital assets act as a medium of exchange. If they act as a medium of exchange, then the scope of the exchange is needed: do they only act within a specific platform or are they used more generally? How they are used “generally” depends on the technologies available at the time of evaluation. In fact, there may be digital assets developed in future that can be exchangeable across platforms. Therefore, it is appropriate to avoid setting statistical classifications based on a medium of exchange.

Foreseeing diverse opinions on the treatment of crypto-assets, the guidance note presents three recording options; treating crypto-assets without a corresponding liability (CAWLM) as any one of “financial”, “non-financial”, or “hybrid” assets. In the hybrid case, they are recorded in the newly created hybrid account. As a current status, these recording options have not yet reached a consensus. Discussions are still ongoing for international agreement.
considering the possibility of future technologies enabling a variety of other types of crypto-assets to emerge.

Finally, as closely related to the above, given the volatile value of crypto-assets, they will not act as a unit of account even if they act as a medium of exchange or a store of value. It lends support to the view that the medium of exchange attribute should not be the highest priority in deciding the assets as financial assets.

1.4 Proposed classification

This paper proposes a classification of digital assets which align with the current SNA (Chart 1). The proposed criteria that the digital assets are classified as financial or non-financial depends on whether there is a corresponding liability, irrespective of whether and how generally they act as a means of payment. Among the digital assets, there may be those issued by public sector and others by private sector. Crypto-assets are different from CBDCs but both of them are digital assets.

**Stablecoins:** There are a variety of so-called stablecoins observed in the market but they have no robust definition. The common feature is that they are designed to achieve stable value (G7WG(2019)) or aim to maintain stable value (FSB(2022)). For statistical purposes, this paper defines stablecoins as only those with a corresponding liability. This treatment is in line with the definition of Boar and Wehrli (2021) which requires stablecoins to have an identifiable issuer as a key criterion. The methodologies to maintain stable value are either (1) backed (collateralized) or (2) unbacked (uncollateralized) but algorithmically controlled. In case of (1), collaterals may be a specific asset or a basket of assets. Furthermore, collaterals may be in some cases held on the issuers’ balance sheet, and in other cases segregated from the issuers’ balance sheet and administered, for instance, in trust accounts. In case of (2), only those which have a corresponding liability should be treated as financial assets.

**Other types of crypto-assets used as a means of payment with a corresponding liability:** This category contains crypto-assets which are not called stablecoins but have a corresponding liability and are used as a means of payment. As a means of payment, the value of the assets is expected to have low volatility. For that reason, at this moment, although this category is conceptually possible, there may not be many digital assets classified to it.

**Prepaid payment instruments:** Prepaid payment instruments are the digital assets that already exist, typically as prepaid card or e-money. Some countries have reflected them in their statistics. They are currently issued based on traditional techniques but if they could be issued in the form of crypto-assets, they would become close to the above category of other types of crypto-assets used as a means of payment with a corresponding liability.

---

4 The classification is based on currently existing digital assets and may vary according to the future market development.

5 It is important to recognize that tying the value to underlying assets does not always mean collateralizing the underlying assets. Moreover, collateralizing is not the same as assuring legal claim of the holders to the underlying assets.
Bank deposits: Bank deposits are in a broad sense one forms of digital assets. 6

Central bank digital currency (CBDCs): As is described in CPMI-MC (2018), CBDC is a digital form of central bank money that is different from balances in traditional reserve or settlement accounts. Since CBDCs are liabilities of central banks, they are not crypto-assets.

Security tokens: Security tokens are digital assets that represent negotiable financial claim on the issuer such as debt securities, equity, and derivatives. In most cases, they are issued or managed using DLT. Given these roles, they will not be expected to act as a means of payment.

Digital assets without a corresponding liability: These are assets which according to our proposed classification—which prioritizes existence or non-existence of corresponding liability—are not to be classified as financial assets. Further research is not provided in this paper. Bitcoin and Ethereum might be possible examples of other types of crypto-assets used as a means of payment without a corresponding liability. Non fungible tokens (NFT) also might be an example of digital assets without a corresponding liability.

6 In this classification, Bank deposits are particularly indicated in light of its role for payment instruments. Securities in book entry transfer system may also be digital assets in the same breadth of digital assets definition.
Proposed digital assets classification

<table>
<thead>
<tr>
<th>Digital assets</th>
<th>Used as a means of payment</th>
<th>With a corresponding liability</th>
</tr>
</thead>
<tbody>
<tr>
<td>Stablecoins</td>
<td></td>
<td></td>
</tr>
<tr>
<td>USDT, USDC, Terra etc.</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Other types of crypt-assets</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Prepaid payment instruments</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Bank deposits</td>
<td></td>
<td></td>
</tr>
<tr>
<td>CBDCs</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Security tokens</td>
<td></td>
<td></td>
</tr>
<tr>
<td>NFT</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Other types of crypt-assets</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Bitcoin, Ethereum etc.</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

1 USD: Tether, USDC: USD Coin, Terra: UST(Terra), NFT: Non-fungible tokens

Stablecoin definitions in the literature

<table>
<thead>
<tr>
<th>Sources</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>G7WG(2019)</td>
<td>Stablecoins are digital tokens that typically transact on a distributed ledger and rely on cryptographic validation techniques to be transacted, with the goal of achieving stable value relative to fiat currencies.</td>
</tr>
<tr>
<td>G30(2020)</td>
<td>A stablecoin is a digital currency. The entity issuing a stablecoin attempts to reduce its price volatility by pegging its value to some external asset or basket of assets like fiat money or exchange-traded commodities.</td>
</tr>
<tr>
<td>Boar, C and A Wehrli (2021)</td>
<td>Private digital tokens that have an identifiable issuer or represent a claim and/or underlying assets.</td>
</tr>
<tr>
<td>FSB(2022)</td>
<td>Stablecoins are a category of crypto-assets that aim to maintain a stable value with reference to a specified asset (typically US dollars), or basket of assets, and provide perceived stability when compared to the high volatility of unbacked crypto-assets.</td>
</tr>
</tbody>
</table>
2. Reflecting digital assets in the Flow of Funds Accounts

This chapter presents an idea of how digital assets could be recorded in the Flow of Funds Accounts based on the abovementioned proposed classification of digital assets. Though not exhaustive, some numerical examples are given for stablecoins, CBDCs, and security tokens (Table 3-5).

2.1 Stablecoins

Stablecoins are defined as having a corresponding liability (Table 3). It is assumed that the stablecoins are issued by Financial corporations (FC). Households or Non-financial corporations (HH • NFC) initially hold currency and purchase stablecoins in exchange for their currency. Here, the transaction of stablecoins is by convention to be recorded under the instrument name of Deposits money. If they are backed by collaterals, the underlying assets may be a single asset or a basket of assets. If the collaterals are financial assets, i.e., Currency, Deposits, and CP, they are recorded in the balance sheets. On the other hand, if the collaterals include crypto-assets without a corresponding liability, these crypto-assets do not appear in the Flow of Funds Accounts.

The underlying assets may either be held by the FC on its balance sheet, or by the third party in the trust account (TC) segregated from the issuer's balance sheet. The financial claim of the FC on this TC is by convention represented as Trust beneficiary certificates.

The recording of algorithmically controlled stablecoins having issuers but without specific collaterals lacks sufficient information to elaborate. Further research is necessary on the mechanism and relationship of players involved.

2.2 CBDCs

The CBDC example (Table 4) draws on the accounting framework employed by the flow of funds representation of CPMI-MC (2018). The HH • NFC initially hold currency and deposits, with a counterpart liability of the central bank (CB) and the depository corporation (DC) respectively. Along with this, the DC holds reserve balances with CB, represented under the name of Deposit with CB.

At the time of issuance, the CB issues CBDC to the DC against the Deposit with CB. When HH • NFC purchases CBDC, there are two options, either to exchange it with an equivalent amount of currency, or to withdraw their deposit. In the former case, the currency returned to the DC is simultaneously deposited to the CB. In the latter case, the deposit withdrawn from the DC are, other things being equal, not

---

7 All described here are conceptual sketch, completely independent of any statistical and policy decision.
8 Stablecoins may not be redeemable if the issuer goes bankrupt. Depending on the legal arrangement, HH • FC may directly claim ownership of the underlying assets in trust account (bankruptcy remote).
9 Table 3 illustrates an algorithm type UST(Terra) issued by the FC (Terra) in the center column and the native token Luna issued by the FC in the left, assuming the two FCs are different. Exact information is required.
replenished automatically. The resulting size of the DC’s balance sheet differs between the two cases, by the amount of CBDC issued.

2.3 Security tokens

Security tokens have issuers and take on the same characteristics of the existing financial instruments (Table 5). They will be unarguably recorded under respective transaction names such as Debt securities, Equity, and Derivatives, according to characteristics of each token.

The example assumes the issuers to be FC, and HH • NFC purchase the debt security tokens, equity security tokens, and derivatives security tokens issued by the issuer in exchange for their holding currency.

### Stablecoins

With collateral on the issuers’ balance sheet

<table>
<thead>
<tr>
<th></th>
<th>FC</th>
<th>HH • NFC</th>
</tr>
</thead>
<tbody>
<tr>
<td>Initial holding</td>
<td></td>
<td>Currency 1,000</td>
</tr>
<tr>
<td>After purchase (single asset)</td>
<td>Currency 1,000</td>
<td>Deposits money 1,000</td>
</tr>
<tr>
<td>After purchase (a basket)</td>
<td>Currency 400 Deposits 300 CP 300</td>
<td>Deposits money 1,000</td>
</tr>
</tbody>
</table>

With collateral segregated from the issuers’ balance sheet

<table>
<thead>
<tr>
<th>After purchase (a basket)</th>
<th>Trust beneficiary certificates 1,000</th>
<th>FC</th>
<th>HH • NFC</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>Deposits money 1,000</td>
<td>Deposits money 1,000</td>
</tr>
</tbody>
</table>

Without collateral (algorithm type) - Provisional idea -

<table>
<thead>
<tr>
<th>After purchase</th>
<th>FC</th>
<th>FC</th>
<th>HH • NFC</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Deposits money 1,000</td>
<td>Deposits money 1,000</td>
<td>Deposits money 1,000</td>
</tr>
</tbody>
</table>

1 The table illustrates the case in which only financial corporations (FC) are allowed to issue stablecoins. This is not always the case, for example, there are cases when non-financial corporations are allowed to issue stablecoins.

2 Assets (left hand side) and liability (right hand side) are shown in each sector.

Source: Author
### CBDCs

<table>
<thead>
<tr>
<th></th>
<th>CB</th>
<th>DC</th>
<th>HH • NFC</th>
</tr>
</thead>
<tbody>
<tr>
<td>Initial holding</td>
<td>Currency 700</td>
<td>Deposits 1,000</td>
<td>Currency 700</td>
</tr>
<tr>
<td></td>
<td>Deposit with CB 500</td>
<td>Deposit with CB 500</td>
<td></td>
</tr>
<tr>
<td>Issuance</td>
<td>Currency 700</td>
<td>Deposits 1,000</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Deposit with CB 300</td>
<td>Deposit with CB 300</td>
<td></td>
</tr>
<tr>
<td></td>
<td>CBDC 200</td>
<td>CBDC 200</td>
<td></td>
</tr>
<tr>
<td>Withdrawal (against currency)</td>
<td>Currency 500</td>
<td>Deposits 1,000</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Deposit with CB 500</td>
<td>Deposit with CB 500</td>
<td></td>
</tr>
<tr>
<td></td>
<td>CBDC 200</td>
<td>CBDC 200</td>
<td></td>
</tr>
<tr>
<td>Withdrawal (against deposits)</td>
<td>Currency 700</td>
<td>Deposits 800</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Deposit with CB 300</td>
<td>Deposit with CB 300</td>
<td></td>
</tr>
<tr>
<td></td>
<td>CBDC 200</td>
<td>CBDC 200</td>
<td></td>
</tr>
</tbody>
</table>

1. Assets (left hand side) and liability (right hand side) are shown in each sector.  

### Security tokens

<table>
<thead>
<tr>
<th></th>
<th>FC</th>
<th>HH • NFC</th>
</tr>
</thead>
<tbody>
<tr>
<td>Initial holding</td>
<td>Currency 1,000</td>
<td></td>
</tr>
<tr>
<td>After purchase</td>
<td>Currency 1,000</td>
<td>Debt securities 500</td>
</tr>
<tr>
<td></td>
<td>Debt securities 500</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Equity 300</td>
<td>Equity 300</td>
</tr>
<tr>
<td></td>
<td>Derivatives 200</td>
<td>Derivatives 200</td>
</tr>
</tbody>
</table>

1. Assets (left hand side) and liability (right hand side) are shown in each sector.  
   Source: Author
3. Difficulties in the collection of source data

The previous chapters discuss the issues regarding the definition and classification of digital assets and propose a classification which aligns with the context of the SNA. It should be noted that even if the appropriate classification is applied in the statistical standard, the compilation of the Flow of Funds Accounts will face difficulties in the collection of source data.

One of the difficulties is the identification of the holders’ sector. In any financial asset, the Flow of Funds Accounts records the amount issued in one sector and simultaneously the same amount purchased in the counterpart sectors. The difficulty is particularly relevant to crypto-assets. Since most crypto-assets work on the public DLT platforms, transactional data are publicly available. Nevertheless, because of their anonymous nature, identification of holders is impossible by design. Moreover, transactions occur between addresses. It is difficult to aggregate addresses even at the sectoral level.\textsuperscript{10}

Consequently, if the holders’ sector remains unknown, the Flow of Funds Accounts will rely on estimation. Some academic research sheds light on the ownership of crypto-assets. Makarov and Schoar (2021) showed that by clustering addresses, individual investors collectively control 8.5 million bitcoins, almost half the bitcoins in circulation by the end of 2020, while the balances held by intermediaries comprise only 5.5 million bitcoins, about one-third of Bitcoin in circulation. Currently though, the result of the research is too far from being utilized as a reference in the estimation of the Flow of Funds Accounts. More detailed information on the ownership structure of crypto-assets is required in future studies.

Another stream related to this issue is the use of private big data in the statistics. As is well known, official statistics around the world are increasingly likely to use private big data through recent digitalization developments. Against this backdrop, some academic research focuses on developing methodologies to make use of big data while assuring anonymity. The statistics in which these studies have interest are basically population estimates, different types of statistics to the SNA, but the aim is the same—safely protecting confidentiality in data sources. The stream of using big data from private and public data sources may provide a hint in compiling holders’ sector.

Taking into account the rapidly proliferating nature of crypto-assets, another pragmatic difficulty for compilers is to consider what crypto-assets should be reflected in the statistics. Key questions here should include: are they of relative importance, widely used and expected to stay in the position for a reasonably long period of time? Are they measurable in the first place?

In summary, given the specific nature of crypto-assets, compilers might have to confront making a more pragmatic decision to allow for implementation, by striking a balance between feasibility and ideal coverage in the statistics.

\textsuperscript{10} The Financial Stability Board (2022) indicates, “it is difficult to aggregate and analyse such data, especially as many transactions occur ‘off-chain’, rather than on the DLT ledger, and at entities that do not report off-chain data, or through complex protocols and smart contracts”.
References


G7 United Kingdom (2021): Public Policy Principles for Retail Central Bank Digital Currencies (CBDCs), October.


Issues in reflecting digital assets in the Flow of Funds Accounts

11th Biennial IFC Conference
BIS, Basel, 25-26 August 2022

Yoshiko Sato
Bank of Japan
Motivation

- No consensus on what digital assets should or should not be included in statistics.
- A guidance note (GN) on the recording of crypto-assets in macroeconomic statistics has been developed, as an updating process towards the 2025 SNA.
- The Flow of Funds Accounts (FOF) — as the essential source data for financial accounts in the SNA — needs to reflect digital assets in an appropriate, feasible and consistent manner.
- This paper presents an idea for the classification of digital assets used as a means of payment (CBDCs, stablecoins, and other types of crypto-assets), with some numerical examples for discussion.
Attributes that matter in the classification of financial assets

• Several common attributes of digital assets are observed in literature from academia and international organizations.

<table>
<thead>
<tr>
<th>Attributes</th>
<th>Examples</th>
</tr>
</thead>
<tbody>
<tr>
<td>Issuer</td>
<td>Existence or absence of issuer. Liability of the issuer corresponding to the asset.</td>
</tr>
<tr>
<td>Type</td>
<td>Claim -based or object -based. Account -based or token -based.</td>
</tr>
<tr>
<td>DLT usage</td>
<td>Use of Distributed Ledger Technology (DLT).</td>
</tr>
<tr>
<td>Value</td>
<td>Redemption value is guaranteed by the issuer at a fixed rate of a certain asset (typically fiat currency), or variable at the time of redemption.</td>
</tr>
<tr>
<td>Mean of payment</td>
<td>Playing as a means of payment.</td>
</tr>
<tr>
<td>Governance / operation</td>
<td>Centralized or distributed. Permissioned or permission-less.</td>
</tr>
<tr>
<td>Cryptography</td>
<td>Use of cryptographic validation technique.</td>
</tr>
<tr>
<td>Public / private</td>
<td>The issuer is public or private entity. The redemption is partially or fully guaranteed by the government.</td>
</tr>
</tbody>
</table>


• Financial assets in the SNA, in principle, should have a counterpart liability (¶11.5, ¶11.7).

• Recognizing this current SNA criteria for financial assets, all digital assets which seem to act as a means of payment will not necessarily be categorized in financial assets. Among the attributes, special attention should be paid to the existence of liability.
Proposed digital assets classification

Digital assets with a corresponding liability are to be treated as financial assets.

The GN’s options:
Crypto-assets without a corresponding liability, designed to act as a general medium of exchange (e.g., Bitcoin) are
- “Financial” assets?
- “Non-financial” assets?
- “Hybrid” assets – to be recorded in the newly created hybrid account?

...Discussions are still ongoing.
Reflecting digital assets in the Flow of Funds Accounts: **Stablecoins**

- An idea of recording -

<table>
<thead>
<tr>
<th>Initial holding</th>
<th>FC</th>
<th>HH+NFC</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>Currency 1,000</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>With collateral on the issuers’ balance sheet</th>
</tr>
</thead>
<tbody>
<tr>
<td>FC</td>
</tr>
<tr>
<td>Currency 400</td>
</tr>
<tr>
<td>Deposits 300</td>
</tr>
<tr>
<td>CP 300</td>
</tr>
<tr>
<td>HH+NFC</td>
</tr>
<tr>
<td>Deposits money 1,000</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>After purchase</th>
<th>FC</th>
<th>HH+NFC</th>
</tr>
</thead>
<tbody>
<tr>
<td>With collateral on the issuers’ balance sheet</td>
<td></td>
<td></td>
</tr>
<tr>
<td>FC</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Deposits money 1,000</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>With collateral segregated from the issuers’ balance sheet</th>
</tr>
</thead>
<tbody>
<tr>
<td>TC</td>
</tr>
<tr>
<td>Currency 400</td>
</tr>
<tr>
<td>Deposits 300</td>
</tr>
<tr>
<td>CP 300</td>
</tr>
<tr>
<td>FC</td>
</tr>
<tr>
<td>Trust beneficiary certificates 1,000</td>
</tr>
<tr>
<td>HH+NFC</td>
</tr>
<tr>
<td>Deposits money 1,000</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Without collateral (algorithm type) - provisional idea -</th>
</tr>
</thead>
<tbody>
<tr>
<td>FC</td>
</tr>
<tr>
<td>Deposits money 1,000</td>
</tr>
</tbody>
</table>

There are a variety of so-called stablecoins.

With a corresponding liability: appear in the FOF
Without a corresponding liability: not appear in the FOF

- Represented in “Deposits money” by convention
- Assumed to be issued by FC

Different ways of recordings, depending on the methodologies used to maintain stable value.

- With/without collateral
- On/off issuers’ balance sheet
- Algorithmically controlled
Reflecting digital assets in the Flow of Funds Accounts: **CBDCs**

- An idea of recording -

<table>
<thead>
<tr>
<th></th>
<th>CB</th>
<th>DC</th>
<th>HH•NFC</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Initial holding</strong></td>
<td>Currency 700</td>
<td>Deposits with CB 500</td>
<td>Deposits 1,000</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Deposits with CB 500</td>
<td>Deposits 1,000</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Issuance</strong></td>
<td>Currency 700</td>
<td>Deposits with CB 300</td>
<td>Deposits 1,000</td>
</tr>
<tr>
<td></td>
<td></td>
<td>CBDC 200</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>Deposits with CB 300</td>
<td>Deposits 1,000</td>
</tr>
<tr>
<td></td>
<td></td>
<td>CBDC 200</td>
<td></td>
</tr>
<tr>
<td><strong>Withdrawal</strong></td>
<td>Currency 500</td>
<td>Deposits with CB 500</td>
<td>Deposits 1,000</td>
</tr>
<tr>
<td>(against currency)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>Deposits with CB 500</td>
<td>Deposits 1,000</td>
</tr>
<tr>
<td></td>
<td></td>
<td>CBDC 200</td>
<td></td>
</tr>
<tr>
<td><strong>Withdrawal</strong></td>
<td>Currency 700</td>
<td>Deposits with CB 300</td>
<td>Deposits 800</td>
</tr>
<tr>
<td>(against deposits)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>Deposits with CB 300</td>
<td>Deposits 800</td>
</tr>
<tr>
<td></td>
<td></td>
<td>CBDC 200</td>
<td></td>
</tr>
</tbody>
</table>

Currently, only a very limited number of countries have introduced CBDCs.

A case of retail CBDCs is described.

There are two options for HH•NFC to withdraw CBDCs against:

- currency
- deposits

CB: Central bank, DC: Depository corporations
Difficulties in the collection of source data

• Identification of holders’ sector.
  - Particularly relevant to crypto-assets. Since most crypto-assets work on the public DLT platforms, transactional data are publicly available. Nevertheless, because of their anonymous nature, identification of holders is impossible by design.
  - The Flow of Funds Account could rely on estimation. Although some academic research has shed light on the ownership of crypto-assets, more detailed information is required.
  - Experiences of using big data safely while protecting confidentiality may provide a hint.

• Striking a balance between feasibility and ideal coverage in the statistics.
  - With lack of source data, compilers might ask themselves “what crypto-assets should actually be reflected?”
Thank you for your attention

Yoshiko Sato
yoshiko.satou@boj.or.jp
Research and Statistics Department
Bank of Japan
Digitally enhanced macroeconomic statistics manuals: the quest for methodological serviceability and compilation synergies

Celestino Girón,
European Central Bank

---

1 This presentation was prepared for the conference. The views expressed are those of the author and do not necessarily reflect the views of the BIS, the IFC or the central banks and other institutions represented at the event.
Digitally enhanced macroeconomic statistics manuals: the quest for methodological serviceability and compilation synergies

Celestino Giron\textsuperscript{1,2}

Abstract

The serviceability of the system of international economic statistics manuals is hampered by the absence of digital versions that facilitate the navigation from one part of the system to the other. Moreover, the service from, access to, and navigation through the system are currently hindered by the presence of frequent overlapping areas across handbooks. We propose a concentric, overlap-free design for the distribution of methodological content across digital manuals, where the fundamental conceptual building blocks are all contained in a central body to which extensions are added for specific statistical domains avoiding any duplication of content. We discuss the advantages of such a model for the development and diffusion of methodological knowledge and the quality of official statistics.


Contents

Digitally enhanced macroeconomic statistics manuals: the quest for methodological serviceability and compilation synergies .............................................. 1

1. Introduction ....................................................................................................................................... 2

2. Building a cross-referencing architecture for the statistical standards ...................... 3

3. Concentric network model for handbooks on economic statistics ............................. 4

4. 2025 SNA and BPM7 ...................................................................................................................... 6

5. Conclusion .......................................................................................................................................... 7

\textsuperscript{1} European Central Bank (ECB).

\textsuperscript{2} Elaborating on a concept note prepared by the same author in cooperation with Daniel Suriany (ECB) and benefiting from discussion with staff from the OECD and Eurostat.
1. Introduction

Methodological handbooks constitute basic building blocks of the economic statistics architecture. They provide descriptive models of the economic reality designed to become useful, objective and comparable tools for economic analysis and policy making. Their development has been the result of a constant worldwide collaborative intellectual effort, which has been a major contribution to our understanding of the economic reality.

The central element of the methodological ecosystem is the UN System of National Accounts (SNA), which provides an all-encompassing accounting framework for the recording of value, its generation, distribution and accumulation. A number of other manuals address specific macroeconomic statistics consistently with the SNA, like the IMF Government Finance Statistics Manual, or provide more detailed guidance on specific issues, like the UN Handbook on National Accounting: Financial Production, Flows and Stocks in the System of National Accounts.

Special status regarding its relation with the SNA has the Balance of Payment Manual (BPM), which addresses largely the same economic phenomena as the SNA, the description of all economic flows of all agents in an economy, but there consolidating out interactions with other agents in the same economy. Moreover, the European System of Accounts (ESA) constitutes a European version of the SNA suitable for legal enforcement in the European Union.

The methodological prescriptions in the different manuals are interrelated, in a sort of conceptual network that grows in complexity as more material is added, in particular if more specific documentation providing guidance, clarification or interpretation of the main manuals and developed in different contexts are also considered as part of the ecosystem. Navigating through this network, even if restricted within the realm of a single manual, might be as challenging as understanding the statistical concepts, classifications and methodological treatments themselves.

The system would benefit from a digital infrastructure that would facilitate access to and use of that body of knowledge through tools that bridge the navigation difficulty. Although Hypertext and HTML - the obvious solutions for establishing cross-reference links in a digital environment - were for long available when the last SNA was finalised -2008-, and more than ten years has elapsed since, no version of the system is currently in place that makes use of this technology. Developing a network of hyperlinks across the various manuals would increase the serviceability of the system to users, be them producers of statistics or users of statistical information.

The ongoing review of the system of methodological handbooks is an opportunity to take into account the inherent network nature of the system in the first place, and to develop as a consequence a content structure that facilitates the implementation of navigation features. The Communication Task Team, set up to support the review process in relation to communication issues, is examining different elements of a design of digital versions of the manuals. Moreover, the editorial teams working on both the updated SNA and BPM are currently developing tables of content and outlines of the manuals that consider to the needs stemming from the digitally enhanced handbooks.

This note contributes to the current debates by proposing a concept of concentric topology for the digital manuals that maximises the advantages of the
digital manuals by separating drastically shared methodological concepts from specific concepts, which would lie in different digital repositories, while preserving the conceptual interconnections between them. Apart from easing the cross-referencing across the ecosystem, a network or concentric topology for the system of manuals would also contribute to avoiding overlaps and conflicts between the various methodological prescriptions, and even to facilitate the drafting process itself.

This note presents this concept of concentric typology. Section 2 discusses on models of digital access and cross-referencing via hyperlinks. Section 3 proposes a concentric model for the system of methodological manuals. Section 4 reflects on the current process of review of the Economic Accounting statistical Standards in the light of the discussion in this paper. Section 4 concludes.

2. Building a cross-referencing architecture for the statistical standards

A natural choice for developing connectivity across standards is the use of HTML-based technology to describe content, so that “hyperlinks” can be used to create a “network” of interrelated content. Hyperlinks in HTML are powering the navigation in the internet, allowing interlinkages within and across web pages and so being the essential instrument for the internet to become a network, a web of resources. The economic statistics standards have all the features of a logical network, with for instance concepts in financial accounts that derive in details belonging to monetary statistics which in turn have implications for balance of payments cross-border financial transactions. Their logical model fits into what HTML and hyperlinks can offer.

Developers of a HTML-based hypertext network decide for a more or less granular, sparse network depending on a cost-benefit assessment of the work needed in each case. For the economic statistics standards, it seems reasonable that, as a minimum requirement, any textual reference within a given manual (e.g., the usual calls from one SNA paragraph to another SNA paragraph to provide the reader with additional detail or extensions) is implemented as a hyperlink within the corresponding internet resource.

A second layer of interconnectivity might consist in adding cross-references from chapter to chapter across manuals, irrespective of whether the methodological prescriptions strictly require a direct link across the chapters. This would imply adding textual information compared with the traditional, non-digital manual versions to indicate and inform on the cross-manual navigation functionalities.

A third layer would also include links from paragraph to paragraph across manuals (and perhaps additional linking within the same manual), which, apart from requiring the introduction of additional text similar as (and in a larger scale than) in the case of the chapter-to-chapter referencing, would entail a great deal of additional conceptual work to ensure an appropriate mapping of the standards across handbooks. At the same time, these conceptual difficulties can be considerably eased if the design of the methodological content already takes into account the need to embed a network in it.

Different degrees of granularity and sparsity can be envisaged for this third layer of connectivity, ranging from just including cross-referencing paragraphs that deal
with the very same economic phenomena to more encompassing links that relate topics which present some, but not necessarily direct, connection.

In addition to support a network across official standards, HTML linking might be used to provide access from the standards to other methodological material that clarify or interpret them in different contexts (like compilation guides, clarification notes or research notes). Finally, HTML hyperlinks and the HTML capacity to embed different media resources might be used to enhance communication, explanatory or training material related to the standards. This extension, in particular if the additional material is provided as part of the standards’ network, would require a careful design of the web pages hosting the manuals, which would likely yield a look-and-feel substantially different from the traditional flat physical or PDF versions of the manuals.

3. Concentric network model for handbooks on economic statistics

Recognising the knowledge network structure of the system of economic statistics, irrespective of its implementation in a specific set of handbooks, is key for improving its serviceability via navigation facilities. At the same time, the specific way in which the methodological content is distributed across the handbooks, i.e. the design of the system of manuals itself, might reinforce such network quality, and/or make it simpler from a topological point of view, and so facilitate the implementation of the navigation functionalities using the HTML technology. We propose here a concentric model for the organization of content to empower the system as a network of resources and facilitate their implementation through HTML technology.

Although the current system of manuals already presents features of a concentric network having the SNA at its core, the handbooks present as well numerous overlaps that make the interlinks often intricate and obscure. This considerably complicates the implementation of the navigation solutions sketched in Section 2. Figure 1 below (left panel, 1.1) provides a graphical representation of the relationships between manuals when overlaps are allowed, showing intersection areas of Venn diagrams that represent overlapping methodological requirements.

A more efficient model for the distribution of content across manuals is presented in Panel 1.2 (right panel) where the central role of the core manual in not compromised by intersections with other manuals. The core body would be the repository of the central methodological requirements, including all fundamental principles, while the other elements would here only contain extensions to the core body for specific representation and analytical needs.

From the point of view of navigation across the network, this layout simplifies considerably the flow of references and facilitates its implementation via hyperlinks (this is illustrated in Figure 1.2 as arrows that only flow from the extensions to the core, as opposed to the two-direction, multi-handbook arrows represented in Figure 1.1).

Apart from easing the implementation of navigation facilities, this topology brings clarity by avoiding the confusion potentially caused by overlapping requirements. The differences and inconsistencies that can currently be found in the
standards would be avoided by construction if an overlap-free model as sketched in 1.2 is followed.

Moreover, the concentric topology provides methodological navigation certainty. In a context like that, methodologists and compilers easily find the area where they should look for the specific guidance they are looking for, diminishing the risk that relevant guidance is available in some other part of the system that they might be unaware of.

The model in Figure 1.2 would also facilitate the drafting and amendment of the handbooks by clearly delineating the various areas to be prepared. A natural division of labour arises where separate editorial teams would be in charge of the various extensions, each of them presumably under the leadership of an international organisation. The design without overlaps avoids unintentional inconsistencies and the duplication of efforts.

At the same time, the core body could also be distributed across leading editors, always avoiding overlaps, ideally allocating to the same team those extensions and core sub-areas more closely related to one another. Thus, for example the team and international organisation in charge of the extension for balance of payments purposes would at the same time be leading the drafting of the elements in the core body more closely related to that domain, like for instance the residency principle or the treatment of income for foreign direct investment links. The more cross-cutting aspects would also be distributed across editors and international organisation after mutual agreement, with the strict requirement that no overlap arises.
4. The new standards: 2025 SNA and BPM7

The current process of review of the Economic Accounting Statistical Standards offers the opportunity to take into account the advantages (and costs) of digitally enhanced manuals since the beginning. The Communication Task Team (CMTT), one of the technical groups set up to provide support to the review process for both the SNA and the BPM, has the development of digitally enhanced manuals as part of its remits.

The development of digital manuals is not only, and even not fundamentally, an IT issue. As argued in this paper, the design of the distribution of methodological content across and within the manuals is key for taking full advantage of what the HTML technology can bring to the standards. In particular, the development of a content topology similar to the one presented in Section 3 would both facilitate the deployment of the digital versions and result in a system of interlinked methodological prescriptions of high serviceability.

The review processes are currently entering in the drafting phase, and the various stakeholders are already discussing on the corresponding outlines and associated distribution of contents. The outlines are being designed with a high degree of cooperation between the SNA and BPM editors, and as agreed at the beginning of the review process, there is a firm intention to share text across manuals, with minimum divergence between them. This goes in the direction of a concentric topology, the shared text constituting the content of the core body.

However, to fully embrace the topological concept, the shared text should present no discrepancy at all, and therefore be prepared with a more ambitious objective than the current intention to have “minimum divergence”. At the same time, the shared text should have a certain degree of internal cohesion, so that the core body can be seen as a separate methodological manual and lie in a separate digital repository. Inter alia, the implementation of a separate repository requires then that the outlines under development consider separate chapters only containing common text and avoiding chapters that cover both common and specific text.

The review processes have focused on certain thematic areas which have been attracting the interests of methodologists over the last years. While this approach made sense in the review phase in order to identify the changes needed in the standards to correctly cope with emerging economic trends, it is dubious that a theme-oriented approach should be followed for the development of the manual outlines.

There is the temptation to include thematic chapters in the manuals to cover areas like globalisation or digitalisation, which are not per se methodological domains, but economic trends (which indeed might require certain changes in methodology, but do not constitute separate conceptual fields). This would inevitably lead to overlaps, just the opposite to what a concentric topology both requires and tries to avoid. This should be avoided.

This doesn't mean that there is no value in developing specific material on economic trends, but that those should rather be interpretations, basically of compilation nature, of methodological prescriptions that strictly follows a concentric topology. Such compilation guidance should rather be developed in separate manuals which as such would be part of the wide concentric topology an reside in the more external layers of it.
At the same time, it should be noted that not all priority areas followed in the reviews’ research phase refer to economic trends. Thus, discussions on wellbeing and sustainability rather try to widen the scope of standard macroeconomic statistics and as such the theme constitutes a new methodological corpus that certainly deserves a separate methodological repository in the inner layers of the concentric topology.

5. Conclusion

The traditional way to disseminate statistical methodology, based on long, thick volumes devoted to the specific domains, falls short of the needs of the methodologist and compiler nowadays, and above all, of the possibilities that new technologies bring.

HTML-based resources interlink concepts in an agile way, facilitating the navigation through a given body of knowledge, and allowing moving from the general to the specific and the other way around at the will of the one accessing the information. The Economic Accounting Statistical Standards constitute an ideal field to apply this concept.

It is important to insist that when we talk about HTML-powered or digitally enhanced statistical standards we do not merely refer to the provision of additional digital versions of traditional manuals that are otherwise the same as they used to be before HTML technologies were available. We rather talk about a new paradigm in the design and distribution of methodological content that is geared by two objectives: facilitating the expression of the methodological content via HTML technologies and, more importantly, exploiting the benefits that such technology brings to the maximum in order to eliminate some of the drawbacks that the system of standards currently has. This includes the presence of overlapping and potentially contradicting methodological prescriptions, and the difficulties to promptly find all prescriptions relevant for a specific case study.

The ongoing reviews of the 2008 SNA and BPM7 is an opportunity to incorporate this concept in the design of macroeconomic statistics methodology for the forthcoming 15 years. We entertain in this context a concentric topological model where common principles and concepts would lie in a central repository accessible to all users of the system, and specific separate extensions would cover specific statistical domains. No overlap would be present throughout the inner layers of the topology.

In order to approximate this concentric model, some practices are suggested to the manuals’ drafting teams. First, common, identical texts should be developed whenever possible, and such texts must constitute a coherent body so that they can all reside in a single repository. Second, chapters in the manuals should avoid sharing common and specific text to facilitate the separation of the two subsets and reinforce the overlap-free concentric structure. Finally, thematic content would rather lie in compilation guides, and not in the main manuals, to thus further contribute to avoiding the presence of conflicting prescriptions.
Digitally enhanced macroeconomic statistics manuals: the quest for methodological serviceability and compilation synergies

Celestino Girón

11th Biennial IFC Conference on Post-pandemic landscape for central bank statistics

26 August 222
What are we discussing?

On the surface (and partially wrong): ways to express traditional methodological manuals through digital resources, and, in doing so, increase serviceability by facilitating navigation.

In reality: how to distribute methodological content across and within manuals seeking for maximizing serviceability. This is…

• Independent of (navigation) technologies, but

• Technologies make innovative answers to this question feasible
<table>
<thead>
<tr>
<th><strong>What are we after?</strong></th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>No overlaps</strong> within and across manuals</td>
</tr>
<tr>
<td><strong>Structured multi-level organisation</strong>, from the general to the specific, from core principles to domain extensions</td>
</tr>
<tr>
<td><strong>Unidirectional referencing</strong> within and across manuals, direct and indirect</td>
</tr>
</tbody>
</table>
What is proposed?

Concentric overlap-free, unidirectional topology, entailing a hierarchical structure with:

1. Core principles
   2. Specific domains
   3. Extensions
   4. Interpretations, clarifications
   5. Thematic manuals

… other material …
What is proposed?

What we have now
What is proposed?

Concentric topology
Separate more general content from more specific content when designing the content of the manuals

Develop common text for common methodological questions

Develop single, shared digital repositories for common content

Avoid thematic chapters in the more central manuals
Thank you for your attention!
11th Biennial IFC Conference on “Post-pandemic landscape for central bank statistics”
BIS Basel, 25-26 August 2022

Measuring payment system policy credibility using machine learning

Muhammad Abdul Jabbar, Okiriza Wibisono and Alvin Andhika Zulen,
Bank Indonesia

---

1 This presentation was prepared for the conference. The views expressed are those of the authors and do not necessarily reflect the views of the BIS, the IFC or the central banks and other institutions represented at the event.
Measuring Payment System Policy Credibility Using Machine Learning

Okiriza Wibisono¹, Muhammad Abdul Jabbar², Alvin Andhika Zulen³

Abstract

In this digital age, payment system plays a crucial role in ensuring the smooth functioning of economic and financial activities. There have been many developments in the payment system landscape, such as digital / wallet-based payments, QR for payment, open APIs, and fast payment systems. For some consumer segments, all these progress can be daunting so as to hinder adoption. On the other hand, some consumers may use these payment facilities without being aware of how best to protect their funds and data. In order to realize the benefits while managing the risks of payments in the digital era, central banks need to understand how the public perceives the payment system ecosystem and its related policies. In our previous research, we have developed machine learning methodology to measure central bank monetary policy credibility. In this paper, we apply a similar machine learning methodology using news data to measure Bank Indonesia's payment system policy credibility. In our case, the index measures public sentiment on 5 aspects of payment system policy: infrastructure, industry conduct, regulation / policy formulation, entry policy, and supervision. The index is aimed at helping central banks in formulating better payment system policy, e.g. by informing better policy communication to the public and as feedback for policy formulation process.

Keywords: payment system, policy credibility, natural language processing

JEL classification: E42, C88

The views and results expressed here are those of the authors and do not necessarily represent Bank Indonesia.

¹ Department of Statistics, Bank Indonesia. email: okiriza_w@bi.go.id
² Department of Statistics, Bank Indonesia, email: muhammad_abdul@bi.go.id
³ Department of Statistics, Bank Indonesia, email: alvin_az@bi.go.id
1. Background

In this digital age, payment system plays a crucial role in ensuring the smooth functioning of economic and financial activities. For example, during 2021, SMS/mobile and internet banking in Indonesia reached almost 8 billion transactions in volume, growing 57% year on year (Bank Indonesia Payment Systems and Financial Market Infrastructure Statistics – Table 7). There have been many developments in the payment system landscape in Indonesia and globally, such as digital/wallet-based payments, QR for payment (Beck et al., 2022), open APIs (BIS-BCBS, 2019), and fast payment systems (BIS-CPMI, 2021). The COVID-19 pandemic and the ensuing mobility and contact restrictions further accelerate noncash payments worldwide (BIS, 2021). These progresses are all welcome, since digital financial inclusion in payments is shown to boost economic growth by as much as 2.2 percentage points (Khera, 2021). In this regard, Bank Indonesia is well-prepared to navigate various developments in payments and related sectors, having launched our Indonesia Payment System Blueprint 2025 back in 2019.

In order to realize the benefits while managing the risks of payments in the digital era, there is a need to build awareness and ascertain consumers that adopting innovative payments technology can be safe and beneficial for them. For some consumer segments, recent innovations in payments can be daunting so as to hinder adoption. Various studies have shown that consumers’ adoption of new technology, such as mobile payments, are driven by the consumers’ perceived risk of said technology. This can include perceived financial risk, privacy risk, performance risk, psychological risk, and time risk. These risks in turn can be affected by perceived technological uncertainty, information asymmetry, regulatory uncertainty, and service intangibility (Yang, 2015).

Other line of research has shown that mobile payments adoption are also driven by the way consumers obtain information. (Suoranta, 2004) find that for adopting new technology, experienced users are more compelled by interpersonal communication by service providers. Non-users and less experienced users, on the other hand, receive information more from mass media, which is related to this paper’s topic. Thus, understanding how the public perceives their payment system policy and the payment ecosystem more broadly can help central banks in formulating payment system policy.

For this purpose, Bank Indonesia used to regularly conduct survey to external stakeholders to measure policy credibility, including payment system policy credibility. The Policy Credibility Survey is based on 6 aspects of credibility: formulation, independence, communication, accountability, coordination, and effectiveness.

In practice, the survey method (in general) has several weaknesses:

1. Survey fatigue: respondents experiencing burnout if surveyed repeatedly, so that surveys cannot be carried out too often (especially since the pool of economists or other stakeholders as respondents can be quite limited);
2. Desirability bias: respondents giving a response that is favorable for the surveyor, which is the central bank, hence the results are less objective;
3. Recency bias: respondents generally providing responses based on recent policies and/or events, hence the survey results are very dependent on the execution time; and
4. Survey cost & time.
Based on these considerations, we develop a payment system policy credibility measurement (indexes) by utilizing Big Data Analytics. The indexes are constructed from text mining of public perceptions toward payment system and payment system policy credibility that are reported in news media. This paper draws largely from the methodology in (Wibisono, 2022). We apply a similar machine learning methodology using news data to measure Bank Indonesia’s payment system policy credibility. The resulting index measures public sentiment on 5 aspects of payment system policy: infrastructure, industry conduct, regulation / policy formulation, entry (policy), and supervision (these aspects are aligned with various payment system functions within related department in Bank Indonesia). The index is aimed at helping central banks in formulating better payment system policy, e.g. by informing better policy communication to the public and as feedback for policy formulation process.

The paper is organized as follows. In section 2, we provide literature reviews on Bank Indonesia’s Policy Credibility Survey and text mining for policy analysis. In section 3, we discuss the data and methodology. In section 4, we provide a summary of the results and evaluation of the model. In section 5, we conclude the paper and offer some thoughts for future works.

2. Literature Review

2.1 Bank Indonesia’s Policy Credibility Survey

From 2013 to 2018, Bank Indonesia conducted Bank Indonesia’s Policy Credibility Survey, a semi-annual survey to measure policy credibility for all 3 sectors of policy: monetary, macroprudential, and payment system. The survey was aimed to provide a measure for policy credibility that is objective, accurate, reflecting broad view of stakeholders (including general public), and available timely. The survey was used to determine the effectiveness of policy communication as well as feedback for formulating future policy communication strategies.

The target respondent of the survey was approximately 1,000 respondents in 20 major cities in Indonesia, consisting of government personnel, bankers, industry players, academics, and general public. The survey measured 6 aspects of policy credibility, from which our indexes are derived:

1. Formulation: whether our policies are formulated carefully according to their objectives
2. Independence: whether we formulate our policies independently, without intervention from any party
3. Communication: whether our policies are well-communicated to the public
4. Accountability: whether our policies are well accounted for
5. Coordination: whether we always coordinate well with the government
6. Effectiveness: whether our policies are effective in achieving their objectives

While most of the aspects above are based on the literature on central bank credibility, especially monetary policy credibility, they are perhaps less relevant for measuring payment system policy credibility, especially independence and coordination aspects. Thus, in this paper we use different credibility aspects, which we describe in section 3.
2.2 Text mining for policy analysis

Text data have been widely used for research in economics and finance. Nowadays, text mining algorithms are growing rapidly along with the adoption of big data and machine learning. These algorithms can automatically “read” and “extract” relevant information from texts, such as person’s name, topics, and sentiment. Compared to manual approach, text mining allows us to make use of much larger text data faster, including news, social media, and press releases. Applied to news media, text mining has the potential to complement survey indicators by extracting and quantifying public’s opinions and sentiments contained in the news.

As an example related to central banks, Sahminan (2008) identified keywords that reflect a tight, neutral, or loose monetary policy inclination in the press release statement of Bank Indonesia over the period from January 2004 to December 2007. (Tobben et al., 2017) developed the Hawkish-Dovish (HD) index that measures media’s perception of ECB communications using two methods: semantic orientation (SO) and support vector machine (SVM). These are based on co-occurrences of strings and machine learning classification algorithm.

Recently, we developed text mining methodology for measuring monetary policy credibility from news (Wibisono, 2022). We use news that are relevant to monetary policy; specifically, sentences that contain any of the related keywords: “inflation”, “monetary”, “exchange rate”, “current account”, “policy rate”, “BI Rate”, “BI 7-Day Reverse Repo Rate”, and their variations in writing. Using methodology similar to what we describe in the next section, the unstructured news sentences are passed into machine learning model and the outputs are aggregated into policy credibility indexes.

3. Methodology

The overall methodology for constructing the policy credibility indexes starting from source data is depicted in Figure 1.
3.1. Data

**Source:** News data serves as the main input for constructing the policy credibility indexes. We use news data from Bank Indonesia's Cyber Library, which is a curated internal repository of news articles related to economic and financial topics. There are more than 30 local news media, with an average of about 850 articles daily, although the number of news media and news articles can vary from month-to-month. The news data are available on a daily basis since 1999, but the news data that we use in this paper span from January 2013. The news are in Bahasa Indonesia (Indonesian language).

**Filtering:** We filter out news that are not relevant for constructing the payment system policy credibility index. Specifically, we only keep news sentences that contain any of the keywords related to payment systems. In total there are more than 100 keywords used, examples of which are: "payment system", "BI-RTGS", "BI-FAST" (our fast payment system), "internet banking", "mobile banking", "e-money", "payment service providers", "card payments", "credit cards", "debit cards", "merchant discount rate", "transfer fee", "QRIS" (our standard for QR payment), and "Open API". Furthermore, the sentence or its previous/next sentence must mention "BI" or "Bank Indonesia".

After an initial run of our model development, we expanded the list of keywords to include those that pertain to payment system regulation, entry policy, and supervision. These are important functions within Bank Indonesia and thus we want to be able to monitor news about these three topics as well. Example keywords are “consumer protection”, “payment service licensing”, “supervisory technology”, “fraud supervision”, and “cyber security”.

3.2. Policy Credibility Index

3.2.1. Annotation

A random sample of the filtered news sentences are manually annotated to construct training data for “teaching” machine learning classification models. Each sentence is labelled with 5 information, representing public’s perception on the payment system credibility aspects. The possible labels for each aspect are positive, negative, or irrelevant, indicating whether the perception/sentiment contained in the sentence are positive, negative, or unrelated to the aspect.

The 5 aspects and their short descriptions are:

1. Payment system infrastructure: policy, developments, and conduct (e.g. reliability, safety, efficiency) of payment system infrastructures, both those that are operated by Bank Indonesia such as BI-RTGS and BI-FAST, or by the industry such as National Payment Gateway (GPN)
2. Payment system conduct by industry: conduct of payments services by the industry, how Bank Indonesia’s payment system policies are implemented by the industry
3. Payment system regulation: whether Bank Indonesia’s payment system policies are well-formulated and their effectiveness in achieving their intended objectives
4. Payment system entry policy: effectiveness and efficiency of payment system entry and licensing activities

5. Payment system supervision: Bank Indonesia’s supervision of the payments industry, e.g. related to payment service and payment infrastructure providers, consumer protection

Annotation is done by the authors and subject matter experts on payment system policy within Bank Indonesia. Prior to annotation, we write out the guidelines on how to annotate the news sentences including specific examples, so that the result is more consistent across annotators. Each sentence is annotated by 2-3 annotators to minimize bias.

A total of 8,556 sentences were annotated. Example annotated sentences for each aspect are provided in Appendix A.

We initially targeted for 5,000 annotated sentences, but additional set of sentences were annotated since we found that there were few sentences that have negative sentiment label, and also few sentences that are relevant to entry policy and supervision aspects. Furthermore, the label distributions after additional annotation are still heavily imbalanced towards positive label, so we resort to machine learning modeling technique that deals with imbalanced data as described in section 3.2.3.

### Distribution of annotated sentences

<table>
<thead>
<tr>
<th>Credibility aspect</th>
<th>Positive</th>
<th>Negative</th>
<th>Irrelevant</th>
</tr>
</thead>
<tbody>
<tr>
<td>Infrastructure</td>
<td>561 (6.6%)</td>
<td>24 (0.3%)</td>
<td>7,971 (93.1% of total)</td>
</tr>
<tr>
<td>Industry conduct</td>
<td>1,745 (20.4%)</td>
<td>200 (2.3%)</td>
<td>6,611 (77.3% of total)</td>
</tr>
<tr>
<td>Regulation</td>
<td>1,423 (16.6%)</td>
<td>80 (1.0%)</td>
<td>7,053 (82.4% of total)</td>
</tr>
<tr>
<td>Entry policy</td>
<td>105 (1.2%)</td>
<td>27 (0.3%)</td>
<td>8,424 (98.5% of total)</td>
</tr>
<tr>
<td>Supervision</td>
<td>179 (2.1%)</td>
<td>42 (0.5%)</td>
<td>8,335 (97.4% of total)</td>
</tr>
<tr>
<td><strong>TOTAL</strong></td>
<td>4,013 (9.4%)</td>
<td>373 (0.9%)</td>
<td>41,789 labels (97.7%)</td>
</tr>
</tbody>
</table>

In general, there are very few sentences with negative label. For payment system entry policy and supervision aspects, almost all sentences are irrelevant.

#### 3.2.2. Data preprocessing

Each filtered sentence (not necessarily annotated) as described in section 3.1 is transformed from textual format into tabular-numeric so that it can be processed by machine learning algorithms, following the steps below:

1. Sentence cleansing: lowercasing, replacing synonyms, abbreviations, numbers, and common names in the sentence;
2. Tokenization: splitting sentence into words/ tokens, removing rarely occurring terms;
3. N-gram tf-idf vectorization: creating n consecutive words (n-gram) with term frequency-inverse document frequency as features (Juan, 2003);

4. Word embedding vectorization: creating vector as numeric representation for each sentence through the average of its words’ word embedding vector (Mikolov et al., 2013);

An example is shown in Figure 2 below.

3.2.3. Model training

From the preprocessed and annotated sentences, we train machine learning models to classify each sentence into one of possible labels (positive/negative/irrelevant). So in total we have 5 sets of models representing the 5 aspects.

**Handling imbalanced labels:** As can be seen from the annotation results (Table 1), the class distributions are quite imbalanced, with most sentences in the news being irrelevant (even after filtering with payment system-related keywords). For relevant sentences, the distributions are also heavily imbalanced towards positive class. Considering these imbalances, we carry out the classification in 2 stages for each credibility aspect:

1. Classifying whether the filtered sentence contains sentiment about payment system (either positive or negative, vs. irrelevant); and

2. For relevant sentences (those that contain sentiment), classifying the sentiment in the sentence (positive vs. negative).

In addition, we also applied synthetic-minority oversampling technique (SMOTE) to remedy the label imbalances (Chawla, 2002).

**Cross validation:** We use 5-fold cross validation: the sentence features generated in section 3.2.2 and resampled with SMOTE are divided into training and validation sets, with approximately 80% in training set (the exact percentages differ across credibility aspects to accommodate the different label distributions). We repeat this 5 times, so we have 5 different training and validation sets, to get a more robust measure of model accuracy.

**Model training:** The sentences in the training sets are input into machine learning algorithms. We experimented with various machine learning algorithms: logistic regression, k-nearest neighbors, support vector machine (SVM), naïve bayes,
decision tree, random forest, XGBoost, and deep learning – long short-term memory (LSTM). To obtain the best accuracy, each algorithm is constructed with various hyperparameter settings.

In total we will have 10 machine learning models, since we have 2 machine learning stages (classifying relevance and classifying sentiment) for each of the 5 aspects.

**Evaluation:** The resulting models are given the sentences in the validation sets, to measure their accuracy on unseen data. We use macro-average F1 score across labels as evaluation metric:

\[
F_1 = \frac{\text{average}(F_1_l)}{l \in \{\text{positive, negative, irrelevant}\}}
\]

\[
F_1_l = \text{precision}_l \times \text{recall}_l
\]

\[
\text{precision}_l = \frac{\# \text{true positive}_l}{\# \text{predicted positive}_l}
\]

\[
\text{recall}_l = \frac{\# \text{true positive}_l}{\# \text{annotated positive}_l}
\]

### 3.2.4. Text classification and index calculation

**Aspect index:** Having obtained the classification models for each credibility aspect, we apply the models to the whole (filtered) news sentences in Cyber Library, including those that were not annotated. For each time period (quarterly or yearly), we tabulate the number of sentences classified as positive or negative for each credibility aspect.

The index for each aspect is calculated as the net balance of the number of positive and negative sentences in each time period.

\[
\text{index}_{aspect,t} = \frac{\# \text{positive}_{aspect,t} - \# \text{negative}_{aspect,t}}{\# \text{positive}_{aspect,t} + \# \text{negative}_{aspect,t}}
\]

The indexes have the characteristics of a net balance index, as below:

1. **Range of index:** [-100%,100%].
2. The index will be close to 100% if there are more news sentences with positive sentiment on the policy credibility aspect. Conversely, the index will be close to -100% if there are more news with negative sentiment.
3. Positive index means more news sentences with positive sentiment on the policy credibility aspect, compared to negative sentences. Conversely, negative index means more news with negative sentiment than positive ones.
4. Zero index means equal number of news sentences with positive sentiment and negative sentiment on the policy credibility aspect.
5. If \(\text{index}_{t1} > \text{index}_{t2}\) then the proportion of news with positive sentiment on the policy credibility aspect is greater in \(t1\) than in \(t2\). Conversely, if \(\text{index}_{t1} < \text{index}_{t2}\) then the proportion of news with negative sentiment on the policy credibility aspect is greater in \(t1\) than in \(t2\).

**Overall index:** The 5 indexes are weighted-averaged to obtain the overall payment system policy credibility index, for each period (quarterly or annually). The weights are based on the aspects’ number of sentences. In addition, we apply a
threshold on the number of relevant sentences on each aspect: there must be at least 4 sentences in a quarter. The threshold value is chosen by median number of sentences in each aspect over the whole data. Aspects with 3 or less sentences in a quarter will be excluded from the overall index.

The full formula for the overall index is below, where 1 denotes the thresholding (whether there are more relevant sentences than the threshold).

\[ \text{index}_t = \sum_{\text{aspect}} 1_{\text{aspect},t} \times (\#\text{positive}_{\text{aspect},t} + \#\text{negative}_{\text{aspect},t}) \times \text{index}_{\text{aspect},t} \]

Besides on historical data, we also calculate the index calculation for ongoing periods, without the need for more manual annotation as the sentence classification models have been trained.

4. Result and Discussion

4.1 Machine learning model evaluation

Below is the best result for each credibility aspect, averaged across the validation sets.

<table>
<thead>
<tr>
<th>Credibility aspect</th>
<th>Best model combination</th>
<th>A: Relevance classification F1</th>
<th>B: Sentiment classification F1</th>
<th>End-to-end F1 (A*B)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Infrastructure</td>
<td>logistic regression &amp; logistic regression</td>
<td>75%</td>
<td>86%</td>
<td>64.5%</td>
</tr>
<tr>
<td>Industry conduct</td>
<td>logistic regression &amp; XGBoost</td>
<td>72%</td>
<td>78%</td>
<td>56.2%</td>
</tr>
<tr>
<td>Regulation</td>
<td>logistic regression &amp; logistic regression</td>
<td>73%</td>
<td>74%</td>
<td>54.0%</td>
</tr>
<tr>
<td>Entry policy</td>
<td>logistic regression &amp; random forest</td>
<td>75%</td>
<td>88%</td>
<td>66.0%</td>
</tr>
<tr>
<td>Supervision</td>
<td>XGBoost &amp; decision tree</td>
<td>68%</td>
<td>89%</td>
<td>60.5%</td>
</tr>
<tr>
<td>Overall (average)</td>
<td>-</td>
<td>72.6%</td>
<td>83.0%</td>
<td>60.2%</td>
</tr>
</tbody>
</table>

Several observations that can be made from the above results are:

- Logistic regression algorithm is the most accurate in majority of cases. This suggests that the relationship between sentence features and credibility labels is mostly linear, or that more (annotated) data is needed to be able to extract more accuracy from using nonlinear algorithms.

- Classifying sentiment (positive vs. negative) is relatively easier than classifying whether the sentence contains sentiment in the first place (83.0% vs. 72.6% averaged macro-F1). Thus, further model development can be focused on improving relevance classification.

- Payment system industry conduct and regulation aspects have lowest end-to-end F1 (due to lowest sentiment F1). This is somewhat unexpected, since these aspects have the largest share of negative sentences (less imbalance), although the numbers are still quite small. We posit that this result may be due to the evaluation metric used, and needs to be further inspected.

- End-to-end F1 (60.2% average) is acceptable, but may warrant further improvement to ensure more robustness of the resulting credibility indexes.
4.2 Index results

The resulting indexes are presented in Figure 3. The lines represent each credibility aspect, and the bar is the overall (threshold-weighted-averaged) index.

Referring to the figure, we comment on some of the indexes’ movements:

- Most indexes are always positive, which means that there are more positive sentences about payment system than negative sentences in the news. The overall index is also increasing over time (86.1% in 2017 to 96.7% in 2021; annual numbers not shown in figure).

- Only supervision index ever reaches negative value. However, we note that supervision index has high volatility, and to less extent, entry policy index as well. This reflects the small number of sentences relevant to these indexes as described in Table 1, even after we applied SMOTE algorithm during model training. This is also one of our considerations for thresholding and weighting the aspect indexes when calculating the overall index.

- The indexes have average pairwise correlation of 13.4% (30.7% if excluding the volatile supervision index), although industry conduct and regulation has 75% correlation. On a high level, the low average correlation suggests that the aspect indexes contain different information that can be relevant for payment system policy analysis.

Example issues classified correctly by the model in 2022 (index numbers not shown in Figure 3) are in Table 3 below.
### Example issues identified in 2022 by aspect

<table>
<thead>
<tr>
<th>Credibility aspect</th>
<th>Example issues</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Infrastructure</strong></td>
<td>Positive news about our newly implemented fast payment system, BI-FAST, its features and advantages</td>
</tr>
<tr>
<td><strong>Industry conduct</strong></td>
<td>Positive news about wider acceptance and use of QRIS (QR Indonesia Standard), and about the increase in QRIS transaction limit</td>
</tr>
<tr>
<td><strong>Regulation</strong></td>
<td>Positive news about Bank Indonesia’s payment system policy, which is aimed at accelerating payment system digitalization to further integrate the digital economy and support national economic recovery</td>
</tr>
<tr>
<td><strong>Entry policy</strong></td>
<td>Positive news about payment system providers licensing by Bank Indonesia</td>
</tr>
<tr>
<td><strong>Supervision</strong></td>
<td>N/A (index cannot be calculated since the number of relevant sentences is less than threshold)</td>
</tr>
</tbody>
</table>

5. **Conclusion & Future Works**

5.1. **Conclusion**

We develop a methodology for measuring Bank Indonesia’s payment system policy credibility by utilizing news articles data and machine learning-based technique. From the out-of-sample evaluation results, we achieve an average F1-score of 60.2%. The methodology is largely drawn from our previous research on monetary policy credibility index from news (Wibisono, 2022).

The resulting policy credibility index shows a positive trend, which means that news about payment systems in recent years is more positive. The machine learning models also seem to be able to capture relevant payment system developments from news, such as BI-FAST implementation and wider adoption of QRIS. On the other hand, some of the aspect indexes, i.e. entry policy and supervision, are highly volatile, which mostly likely is due to the small number of sentences relevant to these aspects.

5.2. **Future works**

Some possible research directions include:

- **Reidentification of credibility aspects**: In our current methodology, we have 5 policy credibility aspects: payment system infrastructure, industry conduct, regulation, entry policy, and supervision. As described before, these are aligned with the different functions performed by related department in Bank Indonesia. However, there are at least 2 issues with the current aspect grouping: (1) entry policy and supervision aspects have very few relevant sentences, and (2) industry conduct and regulation aspects show high correlation. Further assessment may be needed to find the best set of aspects that are relevant to payment system policy-making, have low pairwise index correlations, and whose sentences are relatively easy to classify.

- **Model improvement**: As noted before, end-to-end average F1-score of the machine learning models is 60.2%. On average, this means that the models have 40% probability to miss relevant sentences or misclassify the sentiment (positive vs. negative) of relevant sentences. For more robust policy credibility indexes, there may be a need to improve the accuracy of the models. Several options to improve accuracy are: collect more annotated data (e.g. through additional
keywords) and use nonlinear algorithms in conjunction with the larger data. We can also try to divide the keywords into specific keywords for each aspect to reduce irrelevant sentences, since we observe that the number of irrelevant sentences (Table 1) can be as few as 1.5% per aspect.

- **Econometric analysis**: This paper focuses on developing the policy credibility indexes, from source data, sentence filtering, annotation, preprocessing, machine learning modeling & evaluation, and finally calculating the indexes. Our proposition is that the indexes, at the highest level, provide a measure of how well payment systems and payment system policy is covered in the news. This may impact the tendency for consumers to use/adopt innovative payments services, as well as impact their trust on the central bank’s payment system policy. It will be interesting to analyze the significance of the payment system policy credibility indexes, e.g. as explanatory variable in the framework of mobile payments adoption presented in (Yang, et al., 2015) or analysis of drivers of digital financial inclusion presented in (Khera, et al., 2021).

- **Updating of topics/keywords**: Compared to other policy areas of central bank, e.g. monetary and macroprudential policy, it can be argued that developments in payment systems are faster, technological or otherwise. Thus it is a challenge to keep up with relevant issues about payments in the news, let alone quantify them as positive or negative. Methods in textual trend detection can be considered as a solution for this (Kontosthatis, et al., 2004).
References


## Appendix A: Example annotated sentences

**Example annotated sentences from news**

<table>
<thead>
<tr>
<th>No.</th>
<th>Sentence*</th>
<th>Credibility aspect</th>
<th>Label</th>
</tr>
</thead>
<tbody>
<tr>
<td>1.</td>
<td><em>Bank Indonesia strengthens payment system infrastructures and promotes noncash payments.</em></td>
<td>Infrastructure</td>
<td>Positive</td>
</tr>
<tr>
<td>2.</td>
<td><em>Bank Indonesia detected that on Tuesday there have been network disturbance in their BI-RTGS, BI National Clearing System (SKNBI) and BI Scripless Securities Settlement System (BI-SSSS).</em></td>
<td>Infrastructure</td>
<td>Negative</td>
</tr>
<tr>
<td>3.</td>
<td><em>Since the standardization of electronic payments using QRIS, about 4 million MSMEs have adopted QRIS [for accepting payments].</em></td>
<td>Industry conduct</td>
<td>Positive</td>
</tr>
<tr>
<td>4.</td>
<td><em>QRIS hampers payment system efficiency since there is requirement to become interoperable and interconnected.</em></td>
<td>Industry conduct</td>
<td>Negative</td>
</tr>
<tr>
<td>5.</td>
<td><em>Bank Indonesia’s payment system policy is aimed at accelerating payment system digitalization to further integrate the digital economy and support national economic recovery.</em></td>
<td>Regulation</td>
<td>Positive</td>
</tr>
<tr>
<td>6.</td>
<td><em>Regulation on fintech companies by Bank Indonesia and by Financial Services Authority have not been able to provide legal certainty for consumers.</em></td>
<td>Regulation</td>
<td>Negative</td>
</tr>
<tr>
<td>7.</td>
<td><em>We [Bank Indonesia] are reviewing [payment system] licensing system so that it can be more efficient.</em></td>
<td>Entry policy</td>
<td>Positive</td>
</tr>
<tr>
<td>8.</td>
<td><em>There are complaints from the industry that the fintech licensing system [by Bank Indonesia] is complex and costly in terms of time required.</em></td>
<td>Entry policy</td>
<td>Negative</td>
</tr>
<tr>
<td>9.</td>
<td><em>Bank Indonesia closely supervises payment service providers in order to ensure consumer protection in the payment system and digital economy.</em></td>
<td>Supervision</td>
<td>Positive</td>
</tr>
<tr>
<td>10.</td>
<td><em>Bank Indonesia must recover consumer’s loss in this incident of payment system failure.</em></td>
<td>Supervision</td>
<td>Negative</td>
</tr>
</tbody>
</table>

*) Translated by the authors from Bahasa Indonesia to English. Notes in [square bracket] are added for clarity.
Measuring Payment System Policy Credibility Using Machine Learning

Okiriza Wibisono, Muhammad Abdul Jabbar, Alvin Andhika Zulen

11th IFC Biennial Conference
25-26 August 2022

The views and results expressed here are those of the authors and do not necessarily represent Bank Indonesia.
Payment system plays a crucial role in ensuring the smooth functioning of economic and financial activities. More so in this digital age.

Public’s perception on the payment system ecosystem may impact their adoption of new developments in payments.

Previous approach to measure public perception on our payment system policy credibility: semiannual survey to stakeholders (e.g. economists, academics, government, general public).

This research: Utilizing Big Data Analytics – text mining to gather public perception regarding payment system ecosystem and its related policies.

Methodology largely based on our previous use case on measuring monetary policy credibility (Wibisono, 2022).
Data & Overview of Methodology

News articles

Source: Cyber Library (internal repository of curated economic and financial news)

~30 domestic news (in Bahasa Indonesia)
~850 articles daily

Whole corpus: since Jan 1999
Training data: Jan 2013 – Sep 2021

Example keywords for filtering the news:
- payment system
- BI-FAST
- BI-RTGS
- SKNBI
- internet banking
- mobile banking
- e-money
- payment service providers
- card payments
- credit cards
- debit cards
- transfer fee
- EDC
- Open API
- QRIS
- consumer protection
- payment service licensing
- supervisory technology
- fraud supervision
- cyber security

Payment system infrastructure:
- policy, developments, and conduct (e.g. reliability, safety, efficiency) of payment system infrastructures, both those that are operated by BI or by the industry

Payment system conduct:
- conduct of payments services by the industry, how BI’s payment system policies are implemented by the industry

Payment system regulation:
- whether BI's payment system policies are well-formulated and effective in achieving their intended objectives.

Payment system entry policy:
- effectiveness and efficiency of payment system entry and licensing activities

Payment system supervision:
- BI’s supervision of the payments industry, e.g. related to payment service and payment infrastructure providers, consumer protection

5 Credibility Aspects

Overall index
Methodology – Index Construction

1. Annotation

A sample of filtered sentences are annotated as training data for ML classification models.

- Annotated by authors and domain experts within BI.
- Guidelines incl. examples
- 2-3 annotators per sentence
- Most sentence turns out irrelevant for the aspect

2. Data preprocessing

Each sentence is transformed from text into tabular-numeric format for training ML models.

- Sentence cleansing
- Tokenization
- Remove sparse terms
- N-gram vectorization
- Word embedding

3. Model training

ML model is trained for classifying sentences into pos/neg labels, for each aspect.

- 5-fold CV, 80-20 train-validation split
- SMOTE to handle imbalanced labels
- 2-step classification: relevant vs irrelevant, positive vs negative
- Best average macro F1: 60.2%
- Best algorithm: logistic regression

4. Index calculation

- The ML models are applied to all sentences, to construct monthly indexes.
- The overall index is a weighted average of the 5 component indexes based on number of sentences.
- Any component with 3 or less sentences in a quarter is excluded.

\[
index_{aspect, t} = \frac{\#positive_{aspect, t} - \#negative_{aspect, t}}{\#positive_{aspect, t} + \#negative_{aspect, t}}
\]

\[
index_t = \sum_{aspect} \delta_{aspect, t} \times (#positive_{aspect, t} + #negative_{aspect, t}) \times index_{aspect, t}
\]
<table>
<thead>
<tr>
<th>No</th>
<th>Sentence</th>
<th>Credibility Aspect</th>
<th>Label</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Bank Indonesia strengthens payment system infrastructures and promotes noncash payments.</td>
<td>Infrastructure</td>
<td>Positive</td>
</tr>
<tr>
<td>2</td>
<td>Bank Indonesia detected that on Tuesday there have been network disturbance in their BI-RTGS, BI National Clearing System (SKNBI) and BI Scripless Securities Settlement System (BI-SSSS).</td>
<td>Infrastructure</td>
<td>Negative</td>
</tr>
<tr>
<td>3</td>
<td>Since the standardization of electronic payments using QRIS, about 4 million MSMEs have adopted QRIS [for accepting payments].</td>
<td>Industry conduct</td>
<td>Positive</td>
</tr>
<tr>
<td>4</td>
<td>QRIS hampers payment system efficiency since there is requirement to become interoperable and interconnected.</td>
<td>Industry conduct</td>
<td>Negative</td>
</tr>
<tr>
<td>5</td>
<td>Bank Indonesia's payment system policy is aimed at accelerating payment system digitalization to further integrate the digital economy and support national economic recovery.</td>
<td>Regulation</td>
<td>Positive</td>
</tr>
<tr>
<td>6</td>
<td>Regulation on fintech companies by Bank Indonesia and by Financial Services Authority have not been able to provide legal certainty for consumers.</td>
<td>Regulation</td>
<td>Negative</td>
</tr>
<tr>
<td>7</td>
<td>We [Bank Indonesia] are reviewing [payment system] licensing system so that it can be more efficient.</td>
<td>Entry policy</td>
<td>Positive</td>
</tr>
<tr>
<td>8</td>
<td>There are complaints from the industry that the fintech licensing system [by Bank Indonesia] is complex and costly in terms of time required.</td>
<td>Entry policy</td>
<td>Negative</td>
</tr>
<tr>
<td>9</td>
<td>Bank Indonesia closely supervises payment service providers in order to ensure consumer protection in the payment system and digital economy.</td>
<td>Supervision</td>
<td>Positive</td>
</tr>
<tr>
<td>10</td>
<td>Bank Indonesia must recover consumer's loss in this incident of payment system failure.</td>
<td>Supervision</td>
<td>Negative</td>
</tr>
</tbody>
</table>
### Result – ML model

<table>
<thead>
<tr>
<th>Credibility aspect</th>
<th>Best model relevance</th>
<th>Best model sentiment</th>
<th>A: Relevance classification F1</th>
<th>B: Sentiment classification F1</th>
<th>End-to-end F1 (A*B)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Infrastructure</td>
<td>logistic regression</td>
<td>logistic regression</td>
<td>75%</td>
<td>86%</td>
<td>64.5%</td>
</tr>
<tr>
<td>Industry conduct</td>
<td>logistic regression</td>
<td>XGBoost</td>
<td>72%</td>
<td>78%</td>
<td>56.2%</td>
</tr>
<tr>
<td>Regulation</td>
<td>logistic regression</td>
<td>logistic regression</td>
<td>73%</td>
<td>74%</td>
<td>54.0%</td>
</tr>
<tr>
<td>Entry policy</td>
<td>logistic regression</td>
<td>random forest</td>
<td>75%</td>
<td>88%</td>
<td>66.0%</td>
</tr>
<tr>
<td>Supervision</td>
<td>XGBoost</td>
<td>decision tree</td>
<td>68%</td>
<td>89%</td>
<td>60.5%</td>
</tr>
<tr>
<td><strong>Overall (average)</strong></td>
<td></td>
<td></td>
<td><strong>72.6%</strong></td>
<td><strong>83.0%</strong></td>
<td><strong>60.2%</strong></td>
</tr>
</tbody>
</table>

Some observation about the results:

- **Logistic regression** algorithm is the most accurate in majority of cases.
- Classifying sentiment (positive vs. negative) is relatively **easier** than classifying whether the sentence contains sentiment in the first place (83.0% vs. 72.6% averaged macro-F1).
- Payment system **industry conduct and regulation aspects** have lowest end-to-end F1 (due to lowest sentiment F1). This is somewhat unexpected, since these aspects have the largest share of negative sentences (less imbalance).
- **End-to-end F1** (60.2% average) is **acceptable**, but may warrant further improvement to ensure more robustness of the resulting indexes.
Some observation about the indexes:

- **Most** of the indexes are always positive, which means that there are more positive sentences about payment system than negative sentences in the news.

- The overall index is also **increasing** over time (86.1% in 2017 to 96.7% in 2021; annual numbers not shown in figure).

- **Only supervision** index ever reaches negative value. However, we note that supervision index has high volatility, and to less extent, entry policy index as well.

- The indexes have average pairwise **correlation** of 13.4% (30.7% if excluding the volatile supervision index), although industry conduct and regulation has 75% correlation.

- **Example** recent trends captured by the models: positive news about our newly implemented payment system (BI-FAST), positive news about wider use of QRIS.
Conclusion

1. Developed a machine learning methodology for measuring public’s perception of payment system policy credibility by utilizing news data.

The resulting index shows positive trend: according to the models, news about payment systems in Indonesia in recent years is more positive. But supervision index is highly volatile due to small number of relevant sentences.

2. The models seem to be able to capture relevant developments, such as BI-FAST implementation and wider adoption of QRIS.

Future Works

1. Reidentification/redefinition of the credibility aspects.

2. Model accuracy improvement e.g. by annotating more data, or using specific keywords for each aspect.

3. Econometric analysis (econometric effect of the indexes on macro indicators).

A timely estimation of local investment trends using administrative data

Michele Loberto,
Bank of Italy

---

1 This presentation was prepared for the conference. The views expressed are those of the authors and do not necessarily reflect the views of the BIS, the IFC or the central banks and other institutions represented at the event.
A timely estimation of local investment trends using administrative data

Michele Loberto (Banca d’Italia)

Abstract

Investments in transport equipment are one of the most volatile components of the Italian National Accounts, and they are highly correlated with overall business investments in capital goods. By aggregating selected components of motor vehicle registrations coming from administrative sources, it is possible to approximate national trends in quarterly investments in transport equipment very well. Based on this evidence, we build provincial-level quarterly indicators of business investments based on granular data on motor vehicle registrations, disaggregated by vehicle and owner type. We report significant heterogeneity in investment trends across provinces.

Keywords: business investments, vehicle registrations, administrative data.

JEL classification: E01, E22.

Contents

1. Introduction ....................................................................................................................2

2. Data ..........................................................................................................................3

3. Commercial vehicle registrations and business investments ..................................4
   Exploring provincial developments in commercial vehicle registrations ...............6

4. The impact of Covid-19 epidemics on commercial vehicle registrations ...............9

5. Conclusions ..............................................................................................................10

References .....................................................................................................................11
1. Introduction

Central banks need timely information about the state of the economy to calibrate monetary policy decisions. Indeed, it is difficult to make sound quantitative assessments of the impact of macroeconomic shocks when only aggregate, low-frequency data are available. Moreover, the importance of having high-frequency macroeconomic data available has undoubtedly become much more evident since the outbreak of the Covid-19 pandemic.

In this paper, we exploit a new dataset of granular administrative data for 2019-2021 to nowcast business investments in Italy. In particular, we use data on motor vehicle registrations to track the evolution of investments in transport equipment, which are one of the most volatile components of the Italian National Account and strongly correlate with overall business investments in capital goods.

So far, car registrations have been widely used to assess household expenditure trends on durable goods. Here, we show that aggregating selected components of motor vehicle registrations makes it possible to obtain precise estimates of national trends in quarterly investments in transport equipment as well. In particular, we exploit the information on the type of owner, i.e., whether it is a private individual or a company. This information is crucial to identify and remove the component of car registrations that should be accounted for as households’ expenditure on durable goods, as they may show different trends compared to registrations by companies. Moreover, cars account for the largest and most volatile share of commercial vehicles, and keeping track of cars is key to obtaining a good estimate of investment trends.

Based on this evidence, we exploit the spatial granularity of motor vehicle registrations to build quarterly indicators of business investments at the provincial (NUTS-3) level. These indicators fill a severe statistical gap in terms of spatial granularity of Italian National Accounts. In Italy, business investment estimates are available only up to the NUTS-2 level and with a three-year lag.

We report significant heterogeneity in investment trends across provinces. Looking at the overall evolution of registrations during the pandemic, the percentage variation in vehicle registrations ranges from -37.0 percent to +10.4 percent. In addition, we find that the evolution of investments has been weaker in the provinces most affected by the epidemic, as measured by Covid-19-related hospitalizations. In particular, a 10 percent increase in hospitalizations is associated with a 0.32 percent reduction in the number of vehicle registrations. Therefore, given the significant variance of hospitalization rates across provinces, epidemiological conditions account for a significant fraction of the observed heterogeneity in commercial vehicle registrations.

This paper is organized as follows. Section 2 describes the administrative dataset. Section 3 assesses the correlation between our indicators of commercial vehicle registrations and business investments, and we analyze trends at the provincial level. Section 4 investigates the correlation between investment trends and epidemiological conditions. Section 5 concludes.

---

1 We thank the Ministry of Sustainable Infrastructures and Mobility for providing the data and S. Mocetti for his collaboration during the data acquisition process. We also thank L. Bartiloro, S. Fabiani, A. Rosolia, G. Zevi and R. Zizza for their useful comments.
2. Data

We use granular data on motor vehicles registrations in Italy provided by the Ministry of Sustainable Infrastructures and Mobility. Our dataset includes the monthly number of registrations by vehicle type at the municipal level, from January 2019 to June 2021. The vehicle types are as follows: i) cars, ii) buses, iii) light commercial vehicles, and iv) trucks.\(^2\) Registrations are assigned to municipalities based on the primary address of the owner.

![Figure 1. Car registrations by owner](source)

Source: Ministry of Sustainable Infrastructures and Mobility.

Regarding cars, we know the type of owner, i.e., whether it is a private individual or a company. This information is crucial to identify and remove the component of car registrations that should be accounted as households’ expenditure in durable goods, as they may show different trends compared to registrations by companies (see Figure 1).\(^3\) In our definition commercial vehicles include buses, light commercial vehicles, trucks and cars owned by companies.

Cars account for the largest and most volatile share of commercial vehicles (Figure 2).\(^4\) In 2019, cars accounted for 79 percent of commercial vehicle registrations. The second most important category is light commercial vehicles (18 percent).

\(^2\) Commercial vehicles are classified according to their mass in tons. The mass of light commercial vehicles should be less than 3.5 tons.

\(^3\) In 2019, the share of cars registered by private individuals was about 57 percent.

\(^4\) Unfortunately, we do not have data about the monetary value of different type of vehicles. It is very plausible that the share of cars over total expenditure would be lower, and that of other segments (e.g. trucks) higher.
3. Commercial vehicle registrations and business investments

Our goal is to use commercial vehicle registrations to timely track trends in business investments at a granular level. Similar to Chetty et al. (2020), we test the representativeness of these statistics through a comparison with a publicly available benchmark, possibly at the aggregate level. Unfortunately, official statistics on investment at a geographical level are scarce and very lagging.\(^5\)

Gross fixed investments in transport equipment in the National Accounts – which include depreciation – are the most appropriate benchmark for this comparison.\(^6\) Investments in transport equipment are also one of the most volatile components of the National Accounts, and therefore they are a challenging benchmark. Moreover, investments in transport equipment are highly correlated with overall business investments in machinery, equipment and weapons (Figure 3).\(^7\) Therefore, they provide timely and valuable information on trends in business investment.

\(^5\) ISTAT disseminates annual investment estimates only up to the regional level and with a three-year lag. Last available data refer to 2018. Moreover, the breakdown for asset type is not available.

\(^6\) New vehicle registrations are the main input used by ISTAT to estimate investments in transport equipment. Differently from us, they retrieve data from UNRAE (Unione Nazionale Rappresentanti Autoveicoli Esteri), and there may be minor misalignments compared to the data disseminated by the Ministry of Sustainable Infrastructures and Mobility. Moreover, ISTAT exploits information on transactions of used cars, production and external trade of transport equipments. See ISTAT (2015a) and ISTAT (2015b).

\(^7\) Investments in transport equipment are a subcomponent of investments in machinery, equipment and weapons (their share was about 18 percent during 2015-2020).
Vehicle registrations have a strong seasonal component. Unfortunately, our time series starts in January 2019 and is too short for applying any seasonal adjustment procedure. Therefore, we compare the year-on-year percentage changes of quarterly commercial vehicle registrations and investment in transport equipment since the first quarter of 2020 (Figure 4).\textsuperscript{8} Although this period features unprecedented volatility in macroeconomic variables, changes in commercial vehicle registrations

\textsuperscript{8} The comparison is based on Quarterly National Accounts released by ISTAT on August, 31, 2021.
approximate quite well those of National accounts investments. We detect significant differences only in the second quarter of both 2020 and 2021: in the second quarter of 2020, registrations fell more than investment; as a result, they grew more on an annual basis in the corresponding period of 2021.

A more robust validation would require the availability of longer time series. However, the ability of commercial vehicle registrations to capture the large swings in investment in transport equipment indicates that these statistics are very useful in tracking investment patterns.

Exploring provincial developments in commercial vehicle registrations

Based on the evidence from the previous section, we use commercial vehicle registrations to assess the heterogeneity in investment across provinces. Turning from national to provincial trends, we must point out two potential pitfalls. First, we observe new vehicles registrations but have no information on transactions of existing vehicles. This introduces a measurement error, because a sale of an existing vehicle between two provinces should be accounted as an investment for the destination province and a divestment of the province of origin. Second, registrations are assigned to municipalities based on the primary address of the owner. Instead, it would be ideal if they were assigned based on the location of the production unit that uses the asset in its production process.

Figure 5. Distribution of commercial vehicle registrations across provinces

![Figure 5](image)

Source: Ministry of Sustainable Infrastructures and Mobility.

---

9 Investment in transport equipment also includes lumpy large expenditures on other goods (ships, planes, and trains). The dynamics of these expenditures can explain the differences between the trends in investment and registrations in specific quarters.

10 Up to our knowledge, there is no estimate of trading volumes in the market for existing commercial vehicles in Italy.
Keeping these issues in mind, we analyze the distribution of vehicle registrations across provinces. Taking into account the ratio of number of registrations to the number of active enterprises,\(^{11}\) we find that the provinces of Aosta, Bolzano and Trento are outliers, as in these provinces many car rental companies have their headquarters (Figure 5.a).\(^ {12}\) Based on National Accounts rules, vehicle registrations by rental companies must be recorded as in investments in the provinces where these companies are located. However, in the following analysis we will remove these provinces from our sample, as their vehicle registrations would reflect mostly national rather than local factors (Figure 5.b).\(^ {13}\)

We find large heterogeneity in quarterly year-on-year growth rates across provinces; the dispersion increased in the first half of 2021 (Figure 6), although the median year-on-year percentage change was similar across NUTS-1 regions (Figure 7).

![Figure 6. Heterogeneity in quarterly trends (year-on-year percentage change)](image)

In 2020, commercial vehicle registrations declined in all provinces (Figure 8.a). The steepest declines occurred in South Sardinia, Gorizia, Lecce and Trieste, where the fall was larger than -50 percent. On the other side, registrations were most resilient in Avellino, Enna and Verbania, where the variation was smaller than -15 percent. The median decline was -34.4 percent.

\(^{11}\) Data on active enterprises are available from the Statistical register of active enterprises (ASIA - Enterprises) maintained by ISTAT.

\(^{12}\) In Aosta, Bolzano and Trento registration taxes are the lowest among all provinces, and for this reason they account for a very large share of rental car registrations (about 75 percent in 2019).

\(^{13}\) The province of Florence also has a high ratio of vehicle registrations to population, which is explained by the presence of a large rental company in Scandicci. To eliminate this bias, we removed all observations related to Scandicci from the sample.
The overall evolution of registrations during the pandemic was also very mixed (Figure 8.b). We computed the percentage variation in the first half of 2021 compared to the same period in 2019 to avoid the 2020 base effect. Considering the 5th and the 95th percentiles of the distribution, the percentage variation ranges from -37.0 percent to +10.4 percent. The median change is -18.3 percent.

Source: Ministry of Sustainable Infrastructures and Mobility.
4. The impact of Covid-19 epidemics on commercial vehicle registrations

It is difficult to make a timely quantitative assessment of the impact of macroeconomic shocks when only aggregate, low-frequency data are available. For this reason, as discussed by Nakamura and Steinsson (2018), exploiting spatial heterogeneity may be a strategy to identify causal effects in empirical macroeconomics.\(^{14}\) For example, Chetty et al. (2020) use a large and spatially granular dataset to identify in real-time how COVID-19 epidemics affected the economic activity in the US.

In this section, we use our dataset to find preliminary insights about the relation between local epidemiological conditions and business investments in Italy.

Figure 9. Vehicle registrations (cumulative percentage change between 2019H1 and 2021H1) and hospitalizations

Although an identification of the transmission channels would be key to assess the causal link, the preliminary evidence is that the evolution of commercial vehicle registrations has been weaker in the provinces most affected by the epidemic, as measured by Covid-19 related hospitalizations (Figure 9). To provide a quantitative assessment of the impact, we estimate the following panel regression model on quarterly data between the first quarter of 2019 and the second quarter of 2021:\(^{15}\)

\[
Y_{i,t} = \alpha_i + \gamma_t + \beta \text{Hosp}_{i,t} + \epsilon_{i,t}
\]

\(^{14}\) Quoting Nakamura and Steinsson (2018): “The use of regional data typically multiplies the number of data points available by an order of magnitude or more. It also allows for difference-in-difference identification and makes possible the use of a powerful class of instrumental variables: differential regional exposure to aggregate shocks.”

\(^{15}\) We remove from the sample Gorizia and South Sardinia, as they are outliers (Figure 8.b). Our results would hold also including these small provinces, although the magnitude of the estimated effect of the pandemic on commercial vehicle registrations would slightly decrease.
In this regression, $Y_{i,t}$ is the logarithm of commercial vehicle registrations in province $i$ during quarter $t$, and $Hosp_{i,t}$ is the logarithm of the number of Covid-19 related hospitalizations per 100,000 inhabitants. We control for structural differences between provinces by including fixed effects $\alpha_i$, while the time dummies $\gamma_t$ absorb common trends across provinces. The parameter $\beta$ measures the elasticity of commercial vehicle registrations to hospitalizations.

Figure 10. Quarterly hospitalizations across provinces

![Quarterly hospitalizations across provinces](image)

Source: ISTAT, Ministry of Sustainable Infrastructures and Mobility and Italian National Institute of Health.

We find that the estimate of parameter $\beta$ is both statistically and quantitatively significant: a 10 percent increase in the number of hospitalizations is associated with a 0.32 percent reduction of the number of vehicle registrations. Therefore, given the significant variance of the empirical distributions of the quarterly hospitalization rates across provinces (Figure 10), epidemiological conditions account for a significant fraction of the observed heterogeneity in commercial vehicle registrations.

Although this result is robust to different specifications of the regression model, further work is needed to better understand the causal link between the epidemic and investments in transport equipment.

5. Conclusions

This paper uses vehicle registrations data to analyze the heterogeneity in local investment patterns. In addition, by exploiting spatial granularity and heterogeneity

---

16 We consider hospitalizations because they are less sensitive to differences in testing capacity across time and regions. Since all variables are in logs, we must add 1 to the number of hospitalizations.

17 This estimate is statically significant at the 1 percent level (the standard error is 0.007). The $R^2$ of the regression is 0.98.

18 Considering the empirical distributions of the quarterly hospitalization per 100,000 population across provinces, the 75th percentile is about seven times the 25th percentile (145.3 and 19.6, respectively).
in the spread of the Covid-19 epidemic, it presents quantitative evidence of its potential impact on business investments.

As further developments of this preliminary evaluation, we will analyze longer time series to better assess how commercial vehicle registrations fit the investments in transport equipment from national accounts. Moreover, we will apply seasonal adjustment algorithms to better interpret local patterns. Subsequently, we will investigate the feasibility to analyze trends in vehicle registrations at a lower level of granularity than the province, e.g., the labor market areas (Sistemi locali del lavoro).

A further research avenue will be the analysis of data on car registrations by private individuals. These data will allow a timely analysis of the propensity of households to purchase durable goods up to the municipal level and on a monthly basis.

References


ISTAT (2015b), I conti economici trimestrali - Principali elementi informativi, report.

A timely estimation of local investment trends using administrative data

Michele Loberto (Banca d’Italia)
11th IFC Biennial Conference, August 26, 2022
Introduction

- Investments in transport equipment are very volatile and they are highly correlated with overall business investments in capital goods.

- This paper:
  - Nowcasting business investments in transport equipment in Italy
  - Building timely quarterly indicators of business investments at the NUTS-3 level

- We exploit timely granular administrative data on motor vehicles

- We find that:
  - By aggregating selected components of motor vehicle registrations, we can approximate trends in investments in transport equipment very well
  - There is significant heterogeneity in investment trends across NUTS-3 regions
  - The evolution of investments has been weaker in the provinces most affected by the Covid-19 pandemic
Data

- Monthly number of registrations by vehicle type and owner at the municipal level, from January 2019 to June 2021, provided by the Ministry of Sustainable Infrastructures and Mobility
- We exploit the information about the owner to filter out car registrations from natural persons
- Cars account for the largest and most volatile share of commercial vehicles

**Figure 1. Car registrations and total commercial vehicles registrations**

- a) Car registrations by owner  
  \( indices \ 2019=100 \)
- b) Commercial vehicles registrations by type  
  \( thousands \)

*Source: Ministry of Sustainable Infrastructures and Mobility.*
Vehicle registrations and business investments

- We define our indicator of commercial vehicle registrations as the sum of registrations of buses, light commercial vehicles, trucks and cars owned by companies.

- We validate our indicator against gross fixed investments in transport equipment in the National Accounts.

Figure 2. Commercial vehicles registrations and investments

a) Investments in machinery, equipment and weapons and in transport equipment (millions of euro; chain linked seasonally adjusted)

b) Investments in transport equipment and commercial vehicles registrations (year-on-year percentage change)

Source: ISTAT and Ministry of Sustainable Infrastructures and Mobility.
Estimating regional trends in investments

- We compute NUTS-3-level indicators of commercial vehicle registrations to proxy investments
- We find large heterogeneity in quarterly year-on-year growth rates across provinces
- The dispersion increased in the first half of 2021. The median year-on-year percentage change was similar across NUTS-1 regions.

Figure 4. Heterogeneity in trends across provinces

- a) Heterogeneity in quarterly trends (year-on-year percentage change)
- b) Heterogeneity in half-yearly trends (year-on-year percentage change)

Source: Ministry of Sustainable Infrastructures and Mobility.
Heterogeneity in regional trends

- The overall evolution of registrations during the pandemic was very mixed

Source: Ministry of Sustainable Infrastructures and Mobility.
The impact of the pandemic on vehicle registrations

To provide an assessment of the impact of the pandemic on vehicle registrations, we estimate the following panel regression model on quarterly data between the first quarter of 2019 and the second quarter of 2021:

\[ Y_{i,t} = \alpha_i + \gamma_t + \beta Hosp_{i,t} + \epsilon_{i,t} \]

In this regression, \( Y_{i,t} \) is the logarithm of commercial vehicle registrations in province \( i \) during quarter \( t \), and \( Hosp_{i,t} \) is the logarithm of the number of Covid-19 related hospitalizations per 100,000 inhabitants.

We find that the estimate of parameter \( \beta \) is both statistically and quantitatively significant: a 10 percent increase in the number of hospitalizations is associated with a 0.32 percent reduction of the number of vehicle registrations.

Given the significant variance of the empirical distributions of the quarterly hospitalization rates across provinces, epidemiological conditions account for a significant fraction of the observed heterogeneity in commercial vehicle registrations.
Conclusions

• This paper shows that administrative data on motor vehicles allow:
  • Nowcasting business investments in transport equipment
  • Building timely quarterly indicators of business investments at the NUTS-3 level

• We find that:
  • By aggregating selected components of motor vehicle registrations, we can approximate trends in investments in transport equipment very well
  • There is significant heterogeneity in investment trends across NUTS-3 regions
  • The evolution of investments has been weaker in the provinces most affected by the Covid-19 pandemic

• Next steps:
  • analyze longer time series and apply seasonal adjustment algorithms to better interpret local patterns
  • analyze trends in vehicle registrations at a lower level of spatial granularity
Using the press as a real-time economic confidence indicator

Juan Pablo Cova M, Hugo Peralta V and María del Pilar Cruz N,
Central Bank of Chile

---

1 This presentation was prepared for the conference. The views expressed are those of the authors and do not necessarily reflect the views of the BIS, the IFC or the central banks and other institutions represented at the event.
Using the press as a real-time economic confidence indicator

María del Pilar Cruz N, Hugo Peralta V and Juan Pablo Cova M, Central Bank of Chile.

Abstract

Through the application of textual border analysis tools, this work presents the construction of a high-frequency indicator, generated in real time, based on the computerized reading of the news from the main print media from January 2015 to December 2020 in Chile. This indicator captures the emotional tone of economic and opinion news by making use of an extensive—and novel—purpose-built dictionary in Spanish. This lexicon of words was subjected to challenging statistical tests of robustness, complemented by tests of predictive precision. The latter were carried out by comparing the degree of similarity between a classification by automated means and another by manual means in a random sample. The economic application shows that the constructed indicator has a high correlation with confidence indicators based on surveys, and a high predictive capacity in the face of shock phenomena hitting the economy.

Keywords: in Text mining, nowcasting, dictionary.

JEL classification: C22, C53, C82.

1 The authors wish to thank Professor Stephen Hansen for his valuable comments during the internal seminar organized by the Central Bank of Chile "Unstructured Data Analysis for Macroeconomics and Monetary Policy", held on 13 September 2021. We also thank Pablo Filippi and Carlos Medel for their valuable insights to this research, Consuelo Edwards for her editing services, Observatorio Tecnológico for their support in the application of text mining tools, Cristián López for his support in hardware services, Valentina Vásquez for her research assistance, as well as Alejandra Rozas, Carlos Madeira, Carlos Medel, Consuelo Edwards, Daniel Pérez, Erika Arraño, Felipe Lozano, Joaquín Pérez, José Manuel Salinas, Josué Pérez, Juan Pablo Ugarte, León Sanz, Luis Felipe Villanueva, Marcos Correa, Mario González, Miguel Musa, Natalia Oyanedel, Pablo Filippi, Sebastián Becerra, Sebastián Silva, whose valuable contribution made it possible to generate a set of manually tagged news articles. The opinions and conclusions expressed in this document are the sole responsibility of the authors and do not necessarily represent the views of the Central Bank of Chile. Emails: mcruz@bcentral.cl, pilarcruz@yahoo.com, hperalta@bcentral.cl, jcova@bcentral.cl. This is a shortened version of the paper, prepared for this Conference. The full version can be found here.
Table of contents

Using the press as a real-time economic confidence indicator............................................ 1

1. Introduction....................................................................................................................................... 3

2. Synthesis of related research...................................................................................................... 3

3. Characterizing the news to create a sentiment index.............................................................. 5

4. Constructing a dictionary to read the news........................................................................... 6
   4.1. Methodology for creating a dictionary........................................................................ 6

5. Constructing a set of manually tagged news........................................................................... 9

6. Constructing a dictionary to read the news........................................................................... 9

7. IS-News results and relationship with economic indicators ........................................... 10

8. Conclusions...................................................................................................................................... 15

References................................................................................................................................................ 17
1. Introduction

Since the invention of the printing press, the written media have been an essential mechanism of information, thanks to their ability as a credible source for keeping up to date judgments and visions about the future.

The explosive development of text information by written media has accelerated the need to identify their contents in order to integrate them effectively with economic phenomena. In response to these challenges, computational text analysis has emerged with great force in recent years, through which relevant information can be extracted from unstructured texts and be transformed into structured information. The increase in computer processing capacity has made it possible to create algorithms applicable to language and to considerably reduce processing time and costs.

This technical progress, has in turn, led to important innovations in nowcasting models. Official activity statistics take time to generate and compile, so their availability has lags. The literature in this field (Banbura, Giannone, & Reichlin (2011)) has so far been oriented towards predicting the trajectory of economic variables through the combined use of activity indicators with business and consumer survey indicators. However, more recently, studies such as Kalamara, Turrell, Redl, Kapetanios, & Kapadia (2020) and Thorsrud (2016) have proven that the performance of these models can be improved by considering the news in the press, due to its real-time availability and its predictive value in periods of economic stress.

The main contribution of this work focuses precisely on generating a high-frequency indicator built in real time based on the computerized reading of the news in Chile, with predictive capacity about the evolution of economic conditions, the business cycle and the trajectory of confidence levels in the economy. It uses a database of about 935 thousand pieces of news contained in six major newspapers in the country for the period 2015-2020.

Another contribution is the generation of a Spanish dictionary built from the same news database, which is extensive in the number of tagged words, comparable to its English similes, complex in its variety of grammatical forms, tested on a sample of manually tagged news articles and robust to the statistical testing on news.

The paper is organized as follows. Section 2 contains a synthesis of the related research; section 3 presents a characterization of the IS-News and its transformation into a database suitable for text mining; section 4 presents the construction of the dictionary for reading the news; section 5 presents the construction of a manually tagged set of news; section 6 describes the construction of the IS-News indicator; section 7 shows the results of the IS-News and their relationship with economic indicators, to close with the main conclusions in section 8.

2. Synthesis of related research

Sentiment Analysis is a text mining tool that allows to obtain qualitative data in real time, without resorting to population surveys. It is a field of research that performs a computational treatment of opinions and feelings contained in texts. Its specific application on printed news is relatively recent and its accelerated diffusion is
explained by its contribution to economic analysis and modelling, its ability to anticipate changing conditions and its low production costs compared to surveys (Shapiro, Sudhof, & Wilson (2020)).

From these text analysis tools, two areas of research can be identified that address different objectives: one that detects “intensity” and one that detects “tone”. Intensity measures are based on the number of times certain words appear in the text, and has been extensively disseminated by Baker, Bloom, & Davis (2016) through the Economic Policy Uncertainty (EPU) index. This index counts the number of times words such as uncertainty or recession appear in the analyzed text, similar to what has been done by Altig et al. (2020), Cerda, Silva, & Valente (2016) and Becerra & Sagner (2020). The tone measures, on the other hand, called “Sentiment Analysis” capture the underlying sentiment of optimism or pessimism in texts, using broad tagged lexicons and also automated or machine learning tools. The research area analyzing this paper detects the tone of the texts in the news.

Due to the relevance of using databases with a marked tonal polarity, an important part of sentiment analysis research uses documents with a high subjective and judgmental content, as is the case with blogs, social networks, or product reviews. The press, meanwhile, does not contain this same density of tone as it tries to give the impression of objectivity (Balahur et al. (2013)). In these types of texts, the recount of judgments and polarity is often found in the form of third-party opinions or else in news stories that call upon the opinion of third parties. The editorial line may choose to emphasize or moderate the final texture of the message, but ultimately results in a discourse structure that is more complex to analyze. Therefore, success in capturing the polarity of a news article depends on correctly identifying the sentiment implicit in it (“opinion mining”), isolating the sign of the news piece itself (Saberi & Saad (2017)).

Sentiment analysis saw its first publications in the early 2000s and experienced very accelerated growth a few years later. Using Google Scholar and Scopus citation counts, Mäntylä, Graziotin, & Kuutila (2018) note that in 2000 there were only 37 publications in this field, rising to around 7,000 in 2016. The authors conclude that what enabled this dizzying boom has been the possibility of analyzing huge volumes of texts with computational text mining tools. Some of the research that seems most relevant to us is reviewed next.

Work published by the Federal Reserve of San Francisco (Shapiro, Sudhof, & Wilson (2020)) develops a time series that captures sentiment derived from news stories drawn from economic newspapers between 1980 and 2015. By generating a model that combines various internationally recognized labelled lexicon, they show that daily news sentiment is a good predictor of survey-based confidence indices.

Research by Larsen & Thorsrud (2015) confirms the widespread belief that changes in expectations caused by news is an important autonomous driver of economic fluctuations. Using the main Norwegian business newspapers, the authors identify the topics with the highest predictive power, from which they construct an aggregate index and show that unexpected changes in the index cause significant and persistent fluctuations in markets, especially in credit ones.

Finally, Cruz, Peralta, & Ávila (2020) used computational linguistics to analyze the Business Perceptions Report of the Central Bank of Chile (IPN) and generated an index with high correlations with business and economic confidence indicators.
3. Characterizing the news to create a sentiment index

The formulation of the IS-News is based on the reading and processing of the main printed daily newspapers in the country, of national and international coverage, namely: *El Mercurio, La Tercera, Pulso, La Segunda, Diario Financiero*, and *Estrategia*. Regional print media were excluded because their coverage is not nationwide, as well as those with lower periodicity (weekly or bimonthly), because the index is based on high-frequency data. The database for the construction of the index comes from information generated by *NexNews*.

**Figure 1**: News by media and sections, 2015-2020 (filtered base, % of total news by media)

The number of news articles in the selected media totals 935 thousand for the period under study (between 2015 and 2020). Two events of great impact on the Chilean economy are included in this period: the social crisis of October 2019 and the arrival of the pandemic in March 2020. This base was purged (see section 3.1), reducing its size to 417 thousand news articles, which include the sections Economy and Business, National, International, Politics, Current Affairs and Opinion (Figure 1).

In this new base, the “Economy and Business” and “Opinion” sections are the most relevant, as they account for an average of 56% of the total news, while the “National”, “International” and “Politics” sections each represents around 12%. By media, news in *El Mercurio* and *La Tercera* account for 60% of the total base, while *Diario Financiero, La Segunda, Pulso*, and *Estrategia* make up the remaining 40% (Figure 1).

---

2 In late 2016, this paper began publishing only on digital media, so from then on it is deleted from the database.

3 Relevant media to construct databases for regional sentiment indexes.
4. Constructing a dictionary to read the news

The most used techniques to perform Sentiment Analysis (SA) are, on the one hand, Machine Learning techniques, which encompass a wide range of statistical models capable of learning from massive databases and generate text tone predicting models. This technique’s predictive potential makes it particularly advantageous, but its quality depends on the training set, which needs to be voluminous and complex to learn about the unit lexicons of the language, as well as about both simple and complex sentences.

A second group of commonly used techniques to perform SA contains those that use previously tagged lexicons. This approach is based on a selection of terms that reflect a clear sentiment orientation and can be tagged. Unlike Machine Learning, it is not based on algorithms, but on semantic dictionaries containing terms classified with valences. Words with positive valences are used to collect desired states, and those with negative valences are used to collect undesired ones. In this way words that are contained in the tagged lexicon are detected and added together according to the corresponding polarity. These are the techniques used in this research.

4.1. Methodology for creating a dictionary

The methodology for creating a dictionary in Spanish consisted of several sequential stages, all of them referenced to news databases in Spanish. No other words from other Spanish dictionaries or from translated English dictionaries were included. In other words, the tagged terms were obtained entirely from the printed news themselves, with the purpose of having a domain-specific dictionary.

Briefly, the stages were the following:

- Use of the Python spaCy library, which helps identify the grammatical class$^4$ and the lemma of each word, as well as recording the publication where the news appears. With this processing, the size of the original database is multiplied roughly by 100$^5$, because each word contains a new set of linguistic information.

- The base was subdivided by randomly selecting$^6$ news articles equivalent to 10% of the total base (47 thousand news), in the form of a sampling without replacement$^7$. The processing of the first random sample of news pieces identified a total of 9,596 terms and the second random sample, when compared with the first, added barely 325 new terms (3.4%). For this reason, a third sampling was ruled out due to its low probability of adding any new terms. Thus, with 90 thousand randomly collected news articles, the universe of terms originating the tagged dictionary was covered.

---

$^4$ For example, verbs, adjectives, adverbs, nouns, and determiners, plus names of entities or locations.

$^5$ The complete news base, including functionalities added by spaCy, is around 100 GB in size.

$^6$ For the sample to be representative, the random choice includes the dimension of the media, sections, and dates, and it is made in a stratified way, that is, preserving their proportions in the full base.

$^7$ Sampling without replacement is that which is carried out without taking back to the base those news articles that were chosen to build the sample.
• Identification of the terms (verbs, adjectives, and adverbs) in the random samples is done with the functions of lemmatization and computational POS tagging by spaCy. The complete universe of unique words found in the two samples was 9,921, of which 7,616 were selected for tagging, as 2,305 were removed because of their low frequency.

• Upon completing the process described above, the next step is to tag the set of unique words with positive, negative, or neutral tone by each member of the research team. Following international recommendations (Balahur et al. (2013)), the tagging criteria are standardized, in the sense that only those words with a clear tonality are appraised, and that those terms with two or more diverging meanings or with a diffuse connotation are tagged as neutral.

• Upon completing of the manual tagging process, it was decided to maintain the tags of the words that posted 100% coincidence while, in the cases that presented partial coincidences, it was decided to maintain them only if the discrepant valuation was not the opposite.

• For the rest, where the divergence was bigger, the team proceeded to re-evaluate word by word, analyzing the contexts in which they were used.

The outcome of this methodology was the generation of a tagged dictionary with valences different from zero, based on printed news, with a total of 5,419 lemmas with their inflections, out of which 374 are unique lemmas (257 verbs, 67 adjectives and 50 adverbs) (figure 2).

---

**Figure 2:** Main tagged complex dictionaries (with positive or negative valence)

<table>
<thead>
<tr>
<th>Dictionary</th>
<th>Feature Space</th>
<th>Language</th>
<th>Lexicon Size (Tagged Words)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Hu-Lui (HL)</td>
<td>Movie Reviews</td>
<td>English</td>
<td>6,789</td>
</tr>
<tr>
<td>IS News</td>
<td>News Articles</td>
<td>Spanish</td>
<td>5,419</td>
</tr>
<tr>
<td>Harvard General Inquirer (GI)*</td>
<td>General English</td>
<td>English</td>
<td>4,206</td>
</tr>
<tr>
<td>Loughran-McDonald (LM)*</td>
<td>Financial Reports</td>
<td>English</td>
<td>2,683</td>
</tr>
<tr>
<td>Business Perception Report</td>
<td>Economic and Business Reports</td>
<td>English</td>
<td>774</td>
</tr>
<tr>
<td>Financial Stability Report</td>
<td>Economic and Business Reports</td>
<td>Spanish</td>
<td>565</td>
</tr>
<tr>
<td>Central Bank of Chile (FSR CBCh)</td>
<td>Economic and Business Reports</td>
<td>Spanish</td>
<td>361</td>
</tr>
</tbody>
</table>

* These dictionaries contain additional categories to positive and negative, which means that words can belong to several other different categories such as “degree of uncertainty”, “power”, “strength”, among others. In the LM there are about 1,553 words in other categories which gives rise to a total of 4,236 words tagged in all categories. The whole dictionary contains 26 categories and a total of 11,788 words tagged in all categories. Note that a same word can be counted several times; the word “about” is counted seven times because it is tagged in seven different categories.

---

8 Words with very low frequency, present in fewer than six pieces of news in the sample, may represent an error or be irrelevant in the indicator’s calculation.

9 The dictionary can be found on [here](URL).

10 Inflections are conjugations, plurals, verb tenses, and others. Each verb in the dictionary has associated an average of 20 inflections, for example: aumentar, vs aumentado, aumentando, aumentó, aumentara, aumentaria, and so on.
The extension of this dictionary compares favourably with other English-language dictionaries, popular in SA, such as the Harvard General Inquirer (GI), which has a length of 4,206 words and is a general-purpose dictionary developed by the Harvard University. It also compares positively with the Loughran-McDonald (LM), which is a little smaller than the GI, with a total of 2,683 words, and which has the particularity of being for the specific domain of economics and finance. The Hu-Lui (HL) dictionary, on the other hand, has a total of 6,789 terms, but because it is created from magazines with movie reviews, it limits its use to the economic-financial area.

The analysis of the result of negative and positive valences in each of the aforementioned dictionaries reveals that negative tags in IS-News represent 66% of the total number of words with non-neutral valences, similar to what is found in the other two Spanish dictionaries (IPN and IEF), and in the HL (figure 3). The prevalence of negative over positive tags is consistent with the findings of Reis et al. (2015) (section 2).

![Figure 3: Tagged words in selected dictionaries (with positive or negative valence)](image)

Yet, it is necessary to bear in mind some limitations associated with its effectiveness to perform Sentiment Analysis (Mechulam Burstin & Salvia Varela (2018)). For instance, correctly identifying and tagging the tone of the words, paying attention to the contexts in which they are used. Indeed, a given word may have a neutral sentiment orientation in an economic context, but negative in a scientific or legal one, or vice versa. Using a sample of corporate financial reports (10-K\(^{12}\)) in the United States between 1994 and 2008, Loughran & McDonald (2011) show that nearly three-quarters of the words identified as negative by the GI dictionary are typically not considered negative in financial contexts, such as liability, depreciate, tax, or cost.

---

\(^{11}\) This comparison must take into consideration that the derivations or inflected forms in English are less numerous than in Spanish.

\(^{12}\) 10-K is a yearly financial report of companies listed in the U.S. stock exchange, containing more details than the one that is delivered every year to stockholders, and is a requirement of the U.S. SEC.
5. Constructing a set of manually tagged news

With the objective of evaluating the predictive capacity of the IS-News dictionary, a manual tagging procedure is being performed on a set of news articles, like the one by Shapiro, Sudhof, & Wilson (2020). This procedure is carried out by surveying a group of 23 CBCh researchers, who manually classify 840 news articles according to the polarity of sentiment they detect from their reading. The 840 news articles are randomly selected from the database used by the research.

Figure 4 shows the histogram of manual classifications, automated news ratings and a comparison with that obtained by a similar manual tagging exercise in Shapiro, Sudhof, & Wilson (2020). The results show that the highest frequency of ratings occurs in neutral news (particularly in the automated rating), although with a bias towards negative categories consistent with international evidence, which shows a greater emphasis on reporting negative news.

**Figure 4:** Compared histogram if news set tagging (Is-News tagged manually vs automated)

![Histogram of manual classifications, automated news ratings and a comparison with that obtained by a similar manual tagging exercise in Shapiro, Sudhof, & Wilson (2020).](image)

* Adjusted to histogram scale in Shapiro et al. (2020)

6. Constructing a dictionary to read the news

The central objective of this research is to construct a time series, of daily frequency, based on printed news, using the text mining method called Sentiment Analysis.

Two methodologies commonly found in international research are used to calculate the series. One, which we have termed IS-News⁻⁰⁻⁻⁻⁻⁻⁻⁻⁻⁻⁻⁻⁻⁻⁻⁻⁻⁻⁻⁻⁻⁻⁻⁻⁻⁻⁻⁻⁻⁻⁻⁻⁻⁻⁻⁻⁻⁻⁻⁻⁻⁻⁻⁻⁻⁻⁻⁻⁻⁻⁻⁻⁻⁻⁻⁻⁻⁻⁻⁻⁻⁻⁻⁻⁻⁻⁻⁻⁻⁻⁻⁻⁻⁻⁻⁻⁻⁻⁻⁻⁻⁻⁻⁻⁻⁻⁻⁻⁻⁻⁻⁻⁻⁻⁻⁻⁻⁻⁻⁻⁻⁻⁻⁻⁻⁻⁻⁻⁻⁻⁻⁻⁻⁻⁻⁻⁻⁻⁻⁻⁻⁻⁻⁻⁻⁻⁻⁻⁻⁻⁻⁻⁻⁻⁻⁻⁻⁻⁻⁻⁻⁻⁻⁻⁻⁻⁻⁻⁻⁻⁻⁻⁻⁻⁻⁻⁻⁻⁻⁻⁻⁻⁻⁻⁻⁻⁻⁻⁻⁻⁻⁻⁻⁻⁻⁻⁻⁻⁻⁻⁻⁻⁻⁻⁻⁻⁻⁻⁻⁻⁻⁻⁻⁻⁻⁻⁻⁻⁻⁻⁻⁻⁻⁻⁻⁻⁻⁻⁻⁻⁻⁻⁻⁻⁻⁻⁻⁻⁻⁻⁻⁻⁻⁻⁻⁻⁻⁻⁻⁻⁻⁻⁻⁻⁻⁻⁻⁻⁻⁻⁻⁻⁻⁻⁻⁻⁻⁻⁻⁻⁻⁻⁻⁻⁻⁻⁻⁻⁻⁻⁻⁻⁻⁻⁻⁻⁻⁻⁻⁻⁻⁻⁻⁻⁻⁻⁻⁻⁻⁻⁻⁻⁻⁻⁻⁻⁻⁻⁻⁻⁻⁻⁻⁻⁻⁻⁻⁻⁻⁻⁻⁻⁻⁻⁻⁻⁻⁻⁻⁻⁻⁻⁻⁻⁻⁻⁻⁻⁻⁻⁻⁻⁻⁻⁻⁻⁻⁻⁻⁻⁻⁻⁻⁻⁻⁻⁻⁻⁻⁻⁻⁻⁻⁻⁻⁻⁻⁻⁻⁻⁻⁻⁻⁻⁻⁻⁻⁻⁻⁻⁻⁻⁻⁻⁻⁻⁻⁻⁻⁻⁻⁻⁻⁻⁻⁻⁻⁻⁻⁻⁻⁻⁻⁻⁻⁻⁻⁻⁻⁻⁻⁻⁻⁻⁻⁻⁻⁻⁻⁻⁻⁻⁻⁻⁻⁻⁻⁻⁻⁻⁻⁻⁻⁻⁻⁻⁻⁻⁻⁻⁻⁻⁻⁻⁻⁻⁻⁻⁻⁻⁻⁻⁻⁻⁻⁻⁻⁻⁻⁻⁻⁻⁻⁻⁻⁻⁻⁻⁻⁻⁻⁻⁻⁻⁻⁻⁻⁻⁻⁻⁻⁻⁻⁻⁻⁻⁻⁻⁻⁻⁻⁻⁻⁻⁻⁻⁻⁻⁻⁻⁻⁻⁻⁻⁻⁻⁻⁻⁻⁻⁻⁻⁻⁻⁻⁻⁻⁻⁻⁻⁻⁻⁻⁻⁻⁻⁻⁻⁻⁻⁻⁻⁻⁻⁻⁻⁻⁻⁻⁻⁻⁻⁻⁻⁻⁻⁻⁻⁻⁻⁻⁻⁻⁻⁻⁻⁻⁻⁻⁻⁻⁻⁻⁻⁻⁻⁻⁻⁻⁻⁻⁻⁻⁻⁻⁻⁻⁻⁻⁻⁻⁻⁻⁻⁻⁻⁻⁻⁻⁻⁻⁻⁻⁻⁻⁻⁻⁻⁻⁻⁻⁻⁻⁻⁻⁻⁻⁻⁻⁻⁻⁻⁻⁻⁻⁻⁻⁻⁻⁻⁻⁻⁻⁻⁻⁻⁻⁻⁻⁻⁻⁻⁻⁻⁻⁻⁻⁻⁻⁻⁻⁻⁻⁻⁻⁻⁻⁻⁻⁻⁻⁻⁻⁻⁻⁻⁻⁻⁻⁻⁻⁻⁻⁻⁻⁻⁻⁻⁻⁻⁻⁻⁻⁻⁻⁻⁻⁻⁻⁻⁻⁻⁻⁻⁻⁻⁻⁻⁻⁻⁻⁻⁻⁻⁻⁻⁻⁻⁻⁻⁻⁻⁻┅┅┅┅┅┅┅┅┅┅┅┅┅┅┅┅┅┅┅┅┅┅┅┅┅┅┅┅┅┅┅┅┅┅┅┅┅┅┅┅┅┅┅┅┅┅┅┅┅┅┅┅┅┅┅┅┅┅┅┅┅┅┅┅┅┅┅┅┅┅┅┅┅┅┅┅┅┅┅┅┅┅┅┅・・・・・
To create an algorithm to detect polarity for each news article, which adds words with positive and negative tags. Each text is scrolled through, identifying verbs, adjectives, or adverbs, and then looked up in the dictionary. Sentiment words can take the values 1 (positive tone) or -1 (negative tone), modifying words can take values of 1.5 (intensifies the tone) or 0.5 (attenuates the tone) and negation words, a quotient of -1, which reverses the polarity of the next three contiguous words.

To assign tags at the level of each sentence, for which those with assigned polarity are added together, modifiers and negations are applied, and the result is divided by the total number of words in the news article, to avoid longer stories having a greater weight on the sentiment index.

By contrast, Pointwise Mutual Information (PMI\textsuperscript{13}) methodology provides a measure of the probability that a given word is associated with a given sentiment. Thus, the value assigned manually to each word according to the dictionary is adjusted based on the usage detected throughout the news corpus. For example, if a positive word is found more frequently in positive news, its value will be amplified, whereas if it is more present in negative news, its value will be diminished. Consequently, the application of the PMI algorithm means that the sentiment associated with a news item depends not only on the individual tag of each word, but also on its relative frequency in the rest of the text.

7. IS-News results and relationship with economic indicators

To evaluate the usefulness of the IS-News high frequency indicator, we explore its ability to anticipate confidence indicators such as the Economic Perception Index (IPEC\textsuperscript{14}), the Monthly Business Confidence Index (IMCE\textsuperscript{15}) and their respective sub-indexes, the Business Confidence Index (ICE\textsuperscript{16}), as well as various activity data, in order to identify those with which it is best related. Also, we explore its capacity to anticipate shocks in the economy, whose occurrence over time is also examined.

In addition, we identify which IS-News calculation formula delivers a higher predictive value, i.e., whether the one that uses exclusively the dictionary (IS-News\textsuperscript{Dicc}) or the one that uses the dictionary-based corpus lexicon (IS-News\textsuperscript{PMI}). The results for the period 2015-2020 show the following:

\textsuperscript{13} Re-weighting with PMI is an approximation to methods known as vector space models.

\textsuperscript{14} Economic Perception Index (IPEC in Spanish) produced by GfK Adimark Chile. It is calculated by the monthly application of a structured questionnaire to a sample of 1100 persons 18 years old or more, residing in the main cities of Chile.

\textsuperscript{15} Monthly Indicator of Business Confidence (IMCE in Spanish) produced by ICARE and Adolfo Ibañez University. It is a synthetic index that is applied to 607 companies in four sectors (Industry, Mining, Trade and Construction) reflecting the weighted sum of those indicators for each sector.

\textsuperscript{16} Business Confidence Index (ICE in Spanish) produced by Del Desarrollo University. It is measures economic perception from businessmen point of view. It is elaborated from surveys to approximately 300 general managers, business owners or executives, through telephone surveys or emails.
The Pearson correlation coefficient\(^\text{17}\) presents, without exception, significantly higher values when IS-News is calculated using the PMI methodology. This is consistent with the findings of Shapiro, Sudhof, & Wilson (2020) in which the predictive accuracy of the indicator using PMI metrics is better than that generated with any other of the available on-the-shelf dictionaries.

Correlations rise almost widely when they are lag behind confidence and activity indicators in about four weeks, and they tend to decline when the lag grows to two months, or they are measured contemporaneously.

The highest correlations of the IS-News\(^\text{PMI}\) were found, with a one-month lag, with the Economic Perception Index and its two subindexes, “situation to purchase home items” and “perception of the firms’ current economic situation”, which yield Pearson coefficients of 83%, 92%, and 85%, respectively (figure 7). In other words, the IS-News\(^\text{PMI}\) has the best fit with indicators that reflect the population’s perception of the present and future economic situation as well as of the propensity to consume.

Regarding the Economic Perception Index, the IS-News succeeds in capturing with intensity the two shocks occurring in the period analyzed, i.e., the social outbreak of October 2019 and the arrival of the pandemic in Chile in March 2020. Figure 8 focuses specifically on the period in which these shocks occurred (the complete series is presented in figure 7), showing that both events begin to manifest in the IS-News approximately in the chronological month prior to the one that appears in the survey results. The Pearson coefficient for this specific period rises to 92% for the Economic Perception Index and to 95% for the sub-index “situation to buy household items”. Something similar occurs during the recovery phase, where the IS-News recovers its level before the confidence indicators do so.

\(^{17}\) A measure of the strength of the linear dependence between two quantitative random variables, irrespective of the scale on which they are displayed. It is a good metric when the samples are large, and its distribution follows a normal curve.
When the IS-NewsPMI is analyzed relative to indicators linked to business confidence surveys, the correlations are weaker than with the Economic Perception Index (IPEC). In fact, the correlation coefficient measured against the Monthly Business Confidence Index (IMCE) and its sectoral indexes, ranges between 50% and 60% and the Business Confidence Index (ICE) stands at 73%. Only a couple of exceptions can be noted in some very specific IMCE sub-indexes in which the Pearson coefficient is higher, one being “general current situation of the firm” in the trade sector, and another being “expected costs” in the...
construction sector in which, for both, the correlation coefficient climbs to 81% (figure 9).

**Figure 9:** IS-News\textsuperscript{PMI} vs business confidence indicators (IMCE): 2015-2020 (contemporaneous series; Pearson’s R calculated using IMCE+1)

- When measuring the correlations of IS-News\textsuperscript{PMI} with economic activity indexes, it appears that they are also somewhat lower than those observed with the confidence-in-the-economy index. The Pearson coefficient marks 78% with the headline Imacec, 79% with the non-mining Imacec, and 83% with the services Imacec, all of them displaced one month ahead with respect to the IS-News. In the particular case of the Trade Imacec, the correlation coefficient becomes very low (46%), which has to do with the impulse on consumption during the economy’s recovery phase of the pandemic (second half 2020), which is factored into neither the IS-News nor the Economic Perception Index (IPEC) and sub-indexes, which show a rather minor rebound (figure 10).
As with the confidence indicators, the IS-NewsPMI reflects the impact of the social crisis approximately one month in advance of the Imacec, as is also the case with the sanitary crisis. This short-term predictability condition of the IS-News couples with its real-time availability, which ultimately makes it possible to anticipate by roughly two months what the activity indicators will show (figure 11).
In short, the news index developed for the period 2015-2020 shows a high correlation with indicators of citizens’ confidence regarding their personal and the country’s situation. Its best fit is achieved with expectations indicators referring to consumption intentions, which depend on greater or lower confidence in the conditions of the economy and their own income, and with their perceived economic stability level that allow sustaining said intended consumption over time (figure 12).

This result is consistent with international research findings, which show that the news contributes to the formation of individuals’ expectations, which in turn generate information flows that feed the perceptions surveys. Thus IS-News is capable of recording clearly and in advance the confidence shocks that will be later shown in confidence surveys.

**Figure 12:** IS-NewsRM vs activity and confidence indicators: 2015-2020 (confidence indices in bases; activity indices in 12-month percent variations; Pearson’s R)

8. Conclusions

This research uses the latest methodologies to perform sentiment analysis on voluminous news databases. By computationally reading the country’s leading economic and opinion newspapers, we have generated a time series that shows a high correlation with confidence-in-the-economy indicators, and somewhat lower correlations with indicators of economic activity.

The goodness of the results obtained is grounded on the quality of the database constructed, and on the use of a Spanish dictionary, generated from the same news database, extensive in the number of tagged words, complex in its variety of grammatical forms, tested on a sample of manually tagged stories and robust to statistical testing. This Spanish dictionary is an important contribution to text mining.
research in Spanish, not only because of its size, but also because it allows structuring in various ways the information coming from the written press.

The application of the PMI methodology succeeds in bringing significant improvements to the time series, showing the relevance of incorporating contexts into the process of formulating a manually tagged dictionary. In this sense, it corrects the weaknesses of sentiment analysis methods based solely on dictionaries, allowing the use of a dynamic lexicon, which feeds back as the news base keeps expanding.

In any case, it is important to emphasize that this research made use of the complete database to calculate the PMI, meaning that the valuation of the tagged dictionary words was adjusted according to the usage detected in the entire corpus of news articles. During the period considered, there were two major negative shocks, namely the social outbreak and the pandemic, which may have generated some bias in the application of the PMIs. Therefore, it is proposed here as part of some complementary research, to calculate the optimal and moving PMIs for a time series that uses tagged dictionaries.

The high correlations yielded by the results confirm the international evidence, in the sense that it is possible to computationally capture in real time the level of optimism or pessimism present in the news. There is also evidence that this way of capturing information makes it possible to anticipate confidence shocks, such as those that occurred in 2019 and 2020 in the Chilean economy.

The advantages of having a real-time press indicator are diverse. It provides information that is independent of other sources, has low implementation costs once the methodology has been installed, and can serve as a warning signal of an event of internal or external shock. Moreover, it enables additional applications such as topic analysis, to interpret the unfolding phenomena that accompany the business cycle, as well as the use of bags of words, to obtain intensity measurements.

This notwithstanding, the high correlations between the IS-News and the confidence indicators do not allow inferring that these variables are mutually affecting themselves, meaning that the consumption decisions are affected by the news and vice versa. This would be part of a complementary investigation to this paper.

Finally, the production of the IS-News can serve as a basis for the construction of a model that estimates the way in which sudden confidence-disrupting events detected through the press alter investment and consumption behavior. The early identification of individuals’ optimism or pessimism through the IS-News is an intermediate stage between the release of information by the media and the construction of a predictive model about the behavior of economic variables. Modeling this information also represents a way of extending this research and adding value to the fact of having structured information collected from the press.
Using the press as a real-time economic confidence indicator

References


USING THE PRESS AS A REAL-TIME ECONOMIC CONFIDENCE INDICATOR

AUGUST 26, 2022

RESEARCH TEAM:
María del Pilar Cruz N.
Hugo Peralta V.
Juan Pablo Cova M.

(*) The views and conclusions presented in this paper are exclusively those of the author(s) and do not necessarily reflect the position of the Central Bank of Chile or of the Board members.
ELABORATION OF A NEWS SENTIMENT INDEX

OBJECTIVE:
• Additional source complementing economic activity analysis.

ADVANTAGES:
• Real time data.
• High predictability of the economic cycle.
• Low costs compared to surveys.
• Larger population coverage.
• High effectiveness in changing conditions.
TRANSFORMING DATA INTO (DENSE) INFORMATION USEFUL FOR ANALYSIS

- News database includes main Chilean economic and financial newspaper companies.
- About 935K news articles from 2015 to 2020 were obtained, which were reduced to 417K after the text cleaning process.
- NLP functions from the spaCy library in Python were used to facilitate the construction of the dictionary and calculation of the sentiment indicator IS-News.
NEWS BY MEDIA AND SECTIONS: 2015-2020

NEWS BY MEDIA: 2015-2020
(filtered base; N° of news articles)

- **Estrategia**: 22,025
- **Pulso**: 38,329
- **La Segunda**: 52,759
- **Diario Financiero**: 53,590
- **La Tercera**: 93,811
- **El Mercurio**: 157,212

NEWS BY SECTIONS:
(filtered base; % of total news articles)

- **Economy and business**: 9% (2015), 10% (2016), 17% (2017), 12% (2018), 8% (2019), 8% (2020)
- **National**: 17% (2015), 17% (2016), 20% (2017), 22% (2018), 23% (2019), 26% (2020)
- **Other sections**: 43% (2015), 41% (2016), 35% (2017), 33% (2018), 34% (2019), 42% (2020)
A DICTIONARY TO READ NEWS ARTICLES
AN EXTENSIVE NEWS DICTIONARY WAS CREATED

- **7,617 unique words were identified from the data** (2,832 verbs, 4,157 adjectives and 628 adverbs)
- **374 words were labeled with 1, -1, 0.5 or 1.5**, equivalent to 5,419 words in their various inflected forms.
- **A random set of news items was manually classified**: testing with automatic labeling; similar results were obtained (Shapiro et. al., 2020).

**SIZE OF MAIN LABELED DICTIONARIES**

<table>
<thead>
<tr>
<th>Dictionary</th>
<th>Domain</th>
<th>Language</th>
<th>Labeled words</th>
</tr>
</thead>
<tbody>
<tr>
<td>Hu-Lui (HL)</td>
<td>Movies reviews</td>
<td>English</td>
<td>6,789</td>
</tr>
<tr>
<td>Is News</td>
<td>Chilean news articles</td>
<td>Spanish</td>
<td>5,419</td>
</tr>
<tr>
<td>Harvard General Inquirer (GI)</td>
<td>General English</td>
<td>English</td>
<td>4,206</td>
</tr>
<tr>
<td>Loughran-McDonald (LM)</td>
<td>Companies Financial Statements Report</td>
<td>English</td>
<td>2,683</td>
</tr>
<tr>
<td>Business Perceptions Report (BPR) Central Bank of Chile (CBCh)</td>
<td>BPR BCCh</td>
<td>Spanish</td>
<td>744</td>
</tr>
<tr>
<td>Financial Stability Report (FSR) Central Bank of Spain (CBS)</td>
<td>FSR CBS</td>
<td>Spanish</td>
<td>565</td>
</tr>
<tr>
<td>FSR CBCh</td>
<td>FSR CBCh</td>
<td>Spanish</td>
<td>361</td>
</tr>
</tbody>
</table>

**IS-NEWS HISTOGRAM: COMPARING MANUAL V/S AUTOMATIC LABELING**

![Histogram comparing manual vs automatic labeling](image)
ELABORATING THE IS-NEWS: 
ONE INDICATOR - TWO CALCULATION METHODOLOGIES

IS-NEWS WITH DICTIONARY:

• WORD = LABELED BY DICTIONARY

• NEWS SENTIMENT INDEX (ISN):

\[ ISN = \frac{\#\text{positive} + \#\text{negative} \cdot \#\text{modifiers}}{\#\text{total words within one piece of news}} \]

• DAILY NEWS SENTIMENT INDEX (ISD):

\[ ISD = \frac{\sum ISN}{\#\text{news per day}} \]

WITH POINTWISE MUTUAL INFORMATION (PMI)

• WORD = PMI

\[ PMI(w_i, c) = \log \left( \frac{p(w_i, c)}{p(w_i) \cdot p(c)} \right) \]

• NEWS SENTIMENT INDEX:

\[ ISN^{PMI} = S(w) \]

\[ S(w) = PMI(w_i, POS) - PMI(w_i, NEG) \]

• DAILY NEWS SENTIMENT INDEX:

\[ ISD^{PMI} = \frac{\sum ISN^{PMI}}{\#\text{news per day}} \]
IS-NEWS
RESULTS
IS-NEWS HIGHLY CORRELATED WITH CONSUMER SENTIMENT INDEX
IT IMPROVES USING PMI (Pearson’s R = 83%)

IS-NEWS\(^1\) v/s ECONOMIC PERCEPTION INDEX (IPEC)\(^2\)

(1) Index 2015-2020 average = 100.
(2) Value above (below) 50 indicates optimism (pessimism).
## IS-NEWS CORRELATES WITH CONFIDENCE AND ACTIVITY INDICES

### PEARSON CORRELATION COEFFICIENT: IS-NEWS<sup>PMI</sup> vs CONFIDENCE AND ACTIVITY INDICES

(Confidence indices in bases; activity indices in 12-month percent variations; %)

<table>
<thead>
<tr>
<th>CONFIDENCE INDICATORS</th>
<th>ACTIVITY INDICATORS</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Contemporaneous</td>
</tr>
<tr>
<td>Situation to buy household items (IPEC, sub-index)</td>
<td>92%</td>
</tr>
<tr>
<td>Economic perception index (IPEC)</td>
<td>83%</td>
</tr>
<tr>
<td>Construction, costs expectations (Monthly business confidence index - IMCE-, sub-index)</td>
<td>81%</td>
</tr>
<tr>
<td>Construction, actual general situation of companies (IMCE, sub-index)</td>
<td>75%</td>
</tr>
<tr>
<td>Index of business confidence (ICE)</td>
<td>74%</td>
</tr>
<tr>
<td>Service sector IMACEC (monthly indicator of economic activity)</td>
<td>83%</td>
</tr>
<tr>
<td>Non-mining IMACEC</td>
<td>79%</td>
</tr>
<tr>
<td>IMACEC</td>
<td>78%</td>
</tr>
</tbody>
</table>
**FINAL REMARKS**

1. **IS-News**: high correlations with confidence and activity indices, which grow with the application of the PMI.

2. **Effectiveness**: based on the construction of a dictionary of purpose, and a comparable size to those of greater use in the English language.

3. **Predictability**: anticipates economic shocks in the Chilean economy in a period of around 3 to 4 weeks.

4. **High availability and low cost**: the implementation of a news sentiment index in real time and independent from other data sources.

5. **Text mining in news articles**: many others research applications, such as topic analysis and bag of words to measure intensity.
Measuring macroprudential policy credibility using machine learning

Muhammad Abdul Jabbar, Nursidik Heru Praptono, Okiriza Wibisono and Alvin Andhika Zulen, Bank Indonesia

---

This presentation was prepared for the conference. The views expressed are those of the authors and do not necessarily reflect the views of the BIS, the IFC or the central banks and other institutions represented at the event.
Measuring Macroprudential Policy Credibility Using Machine Learning

Muhammad Abdul Jabbar\textsuperscript{1}, Okiriza Wibisono\textsuperscript{2}, Nursidik Heru Praptono\textsuperscript{3}, Alvin Andhika Zulun\textsuperscript{4}

Abstract

Macroprudential policies and their instruments can potentially be more effective when a central bank has credible track records. Credibility of Bank Indonesia’s macroprudential policy is used to be measured by using surveys to selected stakeholder. However, machine learning and text mining are recently proven to be able to provide less biased and timelier indicator of credibility. In our previous research, we have developed machine learning-based methodology for measuring central bank monetary policy credibility index. In this paper, we extend such methodology using news data in application to the measurement of central bank macroprudential policy credibility index.

Keywords: central bank credibility, macroprudential policy, central bank communication, machine learning, text mining.

JEL classification: E58, C88

\textsuperscript{1} Department of Statistics, Bank Indonesia, email: muhammad_abdul@bi.go.id
\textsuperscript{2} Department of Statistics, Bank Indonesia, email: okiriza_w@bi.go.id
\textsuperscript{3} Department of Statistics, Bank Indonesia, email: nursidik_hp@bi.go.id
\textsuperscript{4} Department of Statistics, Bank Indonesia, email: alvin_az@bi.go.id
Contents

1. Background ........................................................................................................................................... 3

2. Literature Review ................................................................................................................................. 5
   2.1 Bank Indonesia’s Policy Credibility Survey ................................................................................. 5
   2.2 Macroprudential Policy Credibility and Communication ......................................................... 5
   2.3 Text Mining of Economic News Data ......................................................................................... 6

3. Methodology .......................................................................................................................................... 7
   3.1 Data Collection ............................................................................................................................... 7
   3.2 Data Annotation ............................................................................................................................. 7
   3.3 Data Pre-processing ....................................................................................................................... 8
   3.4 Model Construction ....................................................................................................................... 9
   3.5 Index Calculation .......................................................................................................................... 9
   3.6 Model Validation ........................................................................................................................... 11

4. Result and Discussion .......................................................................................................................... 12
   4.1 Machine Learning Model Evaluation ......................................................................................... 12
   4.2 Index Results ............................................................................................................................... 13

5. Conclusion & Future Works ............................................................................................................... 14
   5.1 Conclusion .................................................................................................................................... 14
   5.2 Future Works ............................................................................................................................... 14

References .................................................................................................................................................. 16

Appendix A: Macroprudential Keyword ............................................................................................... 17

Appendix B: Sample of Annotated Sentences ...................................................................................... 17
1. Background

Bank Indonesia is the macroprudential authority as mandated by Act No. 21/2011 concerning Financial Services Authority. Since 2013, the effort to maintain financial system stability especially since the financial crisis of 1997/1998, has been focused on macroprudential policy. Global financial authorities also realize the growing importance of Macroprudential policy in financial stability especially since the global financial crisis of 2007/2008. Bank Indonesia considers macroprudential policy as one of the three main policy framework in their mandate as central bank.

In upholding its macroprudential policy mandates, Bank Indonesia relies on a set of macroprudential policy instruments. Currently there are 5 macroprudential policy instruments:

1. Countercyclical Capital Buffer (CCB)

   The Countercyclical Capital Buffer (CCB) functions as an additional buffer to anticipate losses caused by excessive credit growth with the potential to disrupt financial system stability. The risks are associated with procyclical lending in the banking industry, where banks tend to increase lending during an expansionary economic boom period and restrict lending during a contractionary economic bust period. Procyclicality in Indonesia necessitated CCB implementation, as evidenced by the direct correlation between credit growth and economic growth.

2. Loan-to-value / Financing-to-Value (LTV/FTV)

   The Loan-to-Value or Financing-to-Value (LTV/FTV) Ratio is the ratio of the value of the loan/financing disbursed by a Conventional or Islamic Commercial Bank against the value of collateral in the form of property when the loan is originated based on the latest evaluation. One goal of LTV/FTV policy is to maintain financial system stability and mitigate systemic risk stemming from higher property prices.

3. Macroprudential Intermediation Ratio (MIR)

   The Macroprudential Intermediation Ratio (MIR) and Sharia Macroprudential Intermediation Ratio (Sharia MIR) are macroprudential instruments to ensure the bank intermediation function is managed in line with economic capacity and target growth, while maintaining prudential principles.

4. Macroprudential Liquidity Buffer (MPLB)

   The Macroprudential Liquidity Buffer (MPLB) and Sharia Macroprudential Liquidity Buffer (Sharia MPLB) are minimum liquidity reserves denominated in rupiah that must be maintained by conventional commercial banks and Islamic banks in the form of rupiah securities that can be used for monetary operations, the level of which is set by Bank Indonesia as a percentage of rupiah deposits.

5. Short-Term Liquidity Assistance (PLJP)

   Short-term liquidity assistance (PLJP) is provided by Bank Indonesia to the banking industry in order to overcome short-term liquidity difficulties. Meanwhile, sharia short-term liquidity assistance (PLJPS) is sharia-compliant financing provided by Bank Indonesia to Islamic banks experiencing short-term liquidity difficulties.

These macroprudential policy instruments are the main tools that are constantly being monitored to ensure countercyclicality with Indonesia current economic conditions. Their application and effect into the economy and financial system will
usually triggers responses from stakeholders. These responses are usually made public and captured by the press as news about responses of macroprudential policies. The press also publish news regarding stakeholder analysis about current financial system stability condition that can be provide insights to the public.

Central Bank credibility is defined as a commitment to follow well articulated and transparent rules and policy goals. Since the 2008 financial crisis, central bank macroprudential policy framework has been proven to help maintain macroeconomic and financial stability. The history of central bank credibility is tied up with the history of policy regimes (Bordo & Siklos, 2015).

Empirically, credibility is a qualitative concept, which is not easy to measure. There are several approaches have been used in measuring credibility, including using survey and constructing composite index from several indicators. For Indonesia’s case Bank Indonesia regularly conducts survey to external stakeholders to measure the policy credibility. The policy credibility survey is prepared based on 6 (six) aspects of credibility, i.e. formulation, independence, communication, accountability, coordination, and effectiveness. But, in practice, the survey method has several weaknesses for measuring policy credibility (Zulen et al., 2020). Bank Indonesia has been using macroprudential policy instruments since 2011 to maintain overall financial stability in Indonesia. Therefore macroprudential credibility should also be taken into account for the overall Bank Indonesia credibility.

Central bank has been using machine learning for several of use cases. For example, they have used natural language processing to produce economic or policy uncertainty indices from textual data (Baker et al., 2016) or to gauge sentiment in response to monetary policy announcements, including those for unconventional policy measures (Hansen & McMahon, 2016). Combining natural language processing / text mining and financial stability part of Bank Indonesia mandates, this paper explores how the utilization of news data to measure its macroprudential credibility mandates.

Credibility survey that previously used are expensive, untimely and can contain biases from the stakeholders. News data that we have are highly available and can be collected in a daily basis. Using big data and machine learning we can utilize this data to capture unbiased public opinion regarding macroprudential policy much more timely. Previously we have developed methodology of Bank Indonesia monetary credibility using news data and text mining in (Zulen et al., 2020). In this paper, we use similar methodology to measure the macroprudential perspective of central bank credibility.

The paper is organized as follows. In section 2, we provide literature reviews on central bank usage of big data and machine learning, Bank Indonesia’s policy credibility survey, macroprudential policy credibility and communication, and text mining of economic and financial news. In section 3, we discuss the data and methodology. In section 4, we provide a summary of the results. Lastly in section 5, we conclude the paper and offer some thoughts for future works.
2. Literature Review

2.1 Bank Indonesia’s Policy Credibility Survey

From 2015 to 2018, Bank Indonesia conducted semiannual surveys to external stakeholders to measure macroprudential policy credibility. The survey is expected to provide a measure for policy credibility that is objective, accurate, reflecting broad view of stakeholders (including general public), and available timely in many aspects of policy credibility, which then can be used as feedback for formulating future policy and its communication strategies. The survey consists of a series of questions regarding macroprudential policy ranging from general aspects such as:

1. **Formulation**: Bank Indonesia’s policy is formulated carefully according to its purpose;
2. **Independence**: Bank Indonesia formulates its policy independently, without intervention from any party;
3. **Communication**: Bank Indonesia’s policy has been well communicated to the public;
4. **Accountability**: Bank Indonesia’s policies are well accounted for;
5. **Coordination**: Bank Indonesia and the Government always coordinate well; and
6. **Effectiveness**: Bank Indonesia policies are effective in achieving its objectives.

And also specific questions such as:

1. The difference between the responsibility of Bank Indonesia and Financial Services Authority
2. The stakeholder understanding of Bank Indonesia macroprudential policy
3. If the stakeholders are confident that Bank Indonesia is able to maintain financial stability in Indonesia.

The macroprudential policy survey doesn’t have consistent questions during the 3 years it’s being conducted since there are different aspects and specific questions that are set as the survey objectives every semester. Therefore, we are unable to use this survey as a benchmark to compare the result. This paper methodology calculates credibility within 4 aspects of credibility from the survey that can be captured from the news sentiment, which are formulation, effectiveness, coordination, and communication.

2.2 Macroprudential Policy Credibility and Communication

Bank Indonesia is a central bank which have a policy mix consisting of 3 policies and its macroprudential policy should be part of its credibility. Wairjiyo argues that there are three key reasons why central banks should assume macroprudential policy. First, the performance of their monetary policy functions provides central banks with macroeconomic focus and an understanding of financial markets, institutions and infrastructures needed for the exercise of macroprudential policy. Second, financial instability can be caused by and affect macroeconomic performances, with substantial
consequences for economic activity, price stability and monetary policy transmission. And third, central banks are the ultimate source of liquidity for the economy, through its monetary policy and lender of the last resort functions, and appropriate liquidity provision is crucial for financial system stability (Warjiyo, 2016).

There are several reasons why central bank involvement in macroprudential policy is beneficial. Combining financial supervision with monetary policy tasks can lead to synergies, e.g. through information gains, thereby possibly leading to a more effective conduct of monetary policy and/or to more effective crisis prevention and management (Borio, 2011). Central banks should monitor and regulate systemic risk not only because a financial stability objective is related to the objectives of monetary policy, but also because it is likely to require a lender of last resort function Blinder (2010).

Macroprudential policy communication reflects the credibility of the central bank. A credible central bank policy should be able to affect the financial market and be properly reacted by the general public and stakeholders. In his paper, Born et al. (2010) states some example; if the central bank expresses a rather pessimistic view about the prospects for financial stability, and this view gets heard in financial markets, we would expect that stock prices for the financial sector decline. In that sense, these communications “create news”. The other motive, to “reduce noise”, should then be reflected in market volatility, in the sense that a communication by the central bank should contribute to reducing uncertainty in financial markets, thereby reducing volatility.

2.3 Text Mining of Economic News Data

Text mining is a methodology of automatically extracting high quality information from text data. Text data has proven to provide insightful information for many use cases, including for policy and analysis done by central bank.

Bholat et al. (2015) states the methodologies of text mining that central bank has used. One of them is supervised machine learning. To quote Bholat, “perhaps the most fruitful application of supervised learning techniques in economics is when the researcher has well-motivated text classes”. Supervised machine learning algorithms that start with a researcher manually classifying training data with predefined classes, as in dictionary-based methods. In order to avoid the issue of over-fitting, the algorithm is then validated on another set of documents termed test data before being applied to the rest of the corpus.

This paper develops further from Zulen et al. (2020) methodology in developing a measurement of policy credibility using news data and text mining. In the paper, the focus is monetary policy while in this paper the focus is macroprudential policy. The methodology differs in the keyword that specifically tailored toward macroprudential policy by using different set of keywords while having mostly the same text mining methodology and credibility aspects that are measured compared to the previous research.
3. Methodology

We utilize machine learning to develop sentiment prediction model of news sentences to measure the credibility of Bank Indonesia macroprudential policy in 4 aspects of credibility as stated in chapter 2.1. The workflow is as describe in figure 1.

### 3.1. Data Collection

The news data is obtained from Bank Indonesia Institute’s Cyber Library that collected economy and financial news from print and online press release. The news are filtered using macroprudential keywords that are acquired by consulting with Bank Indonesia Macroprudential policy department. The news data that we use are in Bahasa Indonesia. The keyword and its variants that we use to filters consists of word regarding policy mix, macroprudential policy, financial system stability, its instruments and also its effect in the financial system.

The news are filtered using the keywords above and then processed into sentences. From there we obtain 9060 sentences from January 2013 to June 2021.

### 3.2. Data Annotation

From the 9.060 sentence we randomly pick 5.030 sentences to be annotated with labels to provide each sentence the labels of relevancy and sentiment to each aspects of credibility. First phase annotation is done by 2 people and second phase is done by a 3rd person annotating if the result of annotation is not uniform for a sentence in the first phase. The annotation is done by the authors, and subject matter experts from the Macropdurndential Policy Department of Bank Indonesia using an annotation guidebook that previously formulated by the annotators. The guidebook contains cases of news for each of the credibility aspect that can help the annotator to give the correct label to each news sentence that is relevant to macroprudential policy. Table 2 contains the possible labels for each credibility aspects:
Possible Label for Each Credibility Aspect

<table>
<thead>
<tr>
<th>Credibility Aspects</th>
<th>Possible Credibility Labels</th>
</tr>
</thead>
<tbody>
<tr>
<td>Formulation</td>
<td>Positive: 1</td>
</tr>
<tr>
<td></td>
<td>Negative: -1</td>
</tr>
<tr>
<td></td>
<td>Not Relevant: -</td>
</tr>
<tr>
<td>Effectiveness</td>
<td>Positive: 1</td>
</tr>
<tr>
<td></td>
<td>Negative: -1</td>
</tr>
<tr>
<td></td>
<td>Not Relevant: -</td>
</tr>
<tr>
<td>Coordination</td>
<td>Positive: 1</td>
</tr>
<tr>
<td></td>
<td>Negative: -1</td>
</tr>
<tr>
<td></td>
<td>Not Relevant: 0</td>
</tr>
<tr>
<td>Communication/Expectation</td>
<td>Contractive: 1</td>
</tr>
<tr>
<td></td>
<td>Accommodative: -1</td>
</tr>
<tr>
<td></td>
<td>Neutral: 0</td>
</tr>
<tr>
<td></td>
<td>Not Relevant: -</td>
</tr>
</tbody>
</table>

Table 3 provides the distribution of information labels annotated to the sentences:

<table>
<thead>
<tr>
<th></th>
<th>Positive</th>
<th>Negative</th>
<th>Neutral</th>
<th>Irrelevant</th>
</tr>
</thead>
<tbody>
<tr>
<td>Formulation</td>
<td>642 (12,8%)</td>
<td>70 (1,4%)</td>
<td>Not Applicable</td>
<td>4,318 (85,8%)</td>
</tr>
<tr>
<td>Effectiveness</td>
<td>387 (7,7%)</td>
<td>76 (1,5%)</td>
<td>Not Applicable</td>
<td>4,567 (90,8%)</td>
</tr>
<tr>
<td>Coordination</td>
<td>433 (8,6%)</td>
<td>30 (0,6%)</td>
<td>Not Applicable</td>
<td>4,567 (90,8%)</td>
</tr>
<tr>
<td>Communication/Expectation</td>
<td>72 (1,4%)</td>
<td>585 (11,6%)</td>
<td>25 (0,5%)</td>
<td>4,348 (86,4%)</td>
</tr>
</tbody>
</table>

1 Contractive for communication index. 2 Accommodative for communication index.

3.3. Data Pre-processing

There are 5 steps of pre-processing in order to ensure a high quality text data for the machine learning models.

1. **Sentence Cleansing**
   This step converts synonyms into one representative word, replace numbers, and replace common names into normative representative of each person.

2. **Tokenization**
   This step converts the news text dataset into tokens of words representation that will be vectorised into vector representation that is necessary for machine learning text model.
3. **N-gram and TF-IDF Feature Extraction**

This step converts the tokens dataset into n-grams of the tokens and their weighted frequency in TF-IDF (term frequency-inverse document frequency) vector representation for each of the sentences in the dataset. This TF-IDF will be the main representation of the feature that is used in machine learning model.

4. **Removal of Rarely Occurring Terms**

This step removes the rarely occurring terms from the feature.

5. **SMOTE**

From the distribution of annotated sentences, we can see a potential problem of imbalanced data of minority negative sentiment sentences in the dataset. To deal with imbalanced dataset we use SMOTE (Synthetic Minority Oversampling Technique) to oversample and undersample the dataset. SMOTE uses synthetic data of the minority classes in the dataset to improve the accuracy of the model.

3.4. **Model Construction**

As mentioned before, there is a problem of imbalanced data from the distribution of the annotated sentences. To solve that issue, we model the sentences using 2 phases of modelling before we construct the credibility index of each aspects:

1. 1st phase: Aspect relevancy model that separates sentences relevant to each credibility aspect, and;
2. 2nd phase: Sentiment model that predict the sentiment of each sentences that considered relevant from the 1st phase model.

The aspect and sentiment model of each credibility aspects are trained using 7 machine learning algorithms:

1. Logistic Regression;
2. K-Nearest Neighbor;
3. Support Vector Machine;
4. Naive Bayes;
5. Decision Tree;
6. Random Forest; and
7. XGBoost.

The model of each aspect is then evaluated out-of-sample test data with k-folds cross validation with 80:20 training and test data split using F1 score to get the best aspect and sentiment model. The best models then used to predict periodical sentences relevant aspects and sentiments.

3.5. **Index Calculation**

After we have the best model for each aspects, we can predict the rest of sentences in the news dataset from January 2013 until June 2021. Using the predicted sentiment
of each sentences we can calculate periodical index of each aspect of credibility. The index are calculated using these formulas:

1. **Formulation, Effectiveness, and Coordination credibility aspects**

   The indices are calculated using net balances of the number of positive and negative sentences for aspect $a$ in a period $t$.
   
   $$\text{Index}_{a,t} = \frac{\#\text{positive}_{a,t} - \#\text{negative}_{a,t}}{\#\text{positive}_{a,t} + \#\text{negative}_{a,t}}$$

2. **Communication Credibility Aspect**

   The communication index are calculated by calculating the difference between macroprudential policy forward guidance (Contractive, Neutral, or Accommodative) as stated in Bank Indonesia Board of Governor Meeting press release with the relative expectation of macroprudential policy in the news (expectation index). The forward guidance is available since January 2016 when Bank Indonesia communicate macroprudential policy communication is done more intensively to the general public. We calculate the expectation index using the number of contractive, accommodative and neutral predicted sentences from the period between previous month board of governors meeting to the current month board of governors meeting. The quarterly and semester index are calculated using average of monthly board of governors meeting index.
   
   $$\text{Index}_{\text{Communications},t} = 1 - |\text{FwdGuidance} - \text{Index}_{\text{Expectation},t}|$$

   $$\text{Index}_{\text{Expectation},t} = \frac{\#\text{Contractive} - \#\text{Accommodative}_t}{\#\text{Contractive} + \#\text{Neutral}_t + \#\text{Accommodative}_t}$$

3. **Overall Credibility Index**

   Lastly, the overall credibility index is calculated using simple average of each of credibility aspects.

   $$\text{Credibility Index}_t = \frac{1}{4}(\text{Index}_{\text{Communication},t} + \text{Index}_{\text{Effectiveness},t} + \text{Index}_{\text{Coordination},t} + \text{Index}_{\text{Communication},t})$$

The characteristics of the indices are as follows:

1. **Range of index**: [-100%, 100%].
2. The index will be close to 100% if there are more news with positive sentiment on the policy credibility aspect. For communication index, the index will be close to 100% if public expectation of macroprudential policy is in line with Bank Indonesia macroprudential policy forward guidance.

   The index will be close to -100% if there are more news with negative sentiment on the policy credibility aspect. For communication index, the index will be close to -100% if public expectation of macroprudential policy is not in line with Bank Indonesia macroprudential policy forward guidance.

3. Positive index means more news with with positive sentiment on the policy credibility aspect compared to the negative ones. For communication index,
positive index means public expectation of macroprudential policy is in line with Bank Indonesia macroprudential policy forward guidance.

Zero index means equal number of news with positive sentiment and negative sentiment on the policy credibility aspect.

Negative index means more news with negative sentiment on the policy credibility aspect compared to the positive ones. For communication index, positive index means public expectation of macroprudential policy is in line with Bank Indonesia macroprudential policy forward guidance.

4. If \( \text{index}_1 > \text{index}_2 \) then the proportion of news with positive sentiment on the policy credibility aspect is greater in \( t_1 \) than in \( t_2 \).

For communication index if \( \text{index}_1 > \text{index}_2 \) then public expectation on macroprudential policy is more in line with Bank Indonesia’s forward guidance in \( t_1 \) than in \( t_2 \).

If \( \text{index}_1 < \text{index}_2 \) then the proportion of news with negative sentiment on the policy credibility aspect is greater in \( t_1 \) than in \( t_2 \).

For communication index, if \( \text{index}_1 < \text{index}_2 \) then public expectation on monetary policy is more not in line with Bank Indonesia’s forward guidance in \( t_1 \) than in \( t_2 \).

3.6. Model Validation

For a machine learning to be used in production, it has to be evaluated to make sure that it can predict the aspect and sentiment of the sentences well. We use F1 score for our evaluation metrics since we want a model with balanced model precision and recall. Using F1 score we can also pick the model with the best predictive power from the 7 machine learning algorithms that we use. As explained in section 3.3, we use k-fold shuffled cross validation with 5 k to get a more robust validation. Lastly we use 80:20 split for each of the cross validation split.

To calculate F1 score we need to have the number of true and false prediction against the test data. F1 score is calculated from calculating the harmonic mean of precision and recall. Precision is the percentage measure of accuracy for sentences identified as credibility and recall is the percentage measure of identified credibility sentences.
The formulas to calculate precision, recall and f1 score are as follows:

\[
\text{Precision} = \frac{TP}{(TP + FP)}
\]

\[
\text{Recall} = \frac{TP}{(TP + FN)}
\]

\[
F1 \text{ Score} = \frac{2(\text{Precision} \times \text{Recall})}{(\text{Precision} + \text{Recall})}
\]

4. Result and Discussion

4.1 Machine Learning Model Evaluation

After performing a horse race of 7 machine learning algorithms, we obtained the best results for the macroprudential credibility news dataset for each credibility aspects. The model is evaluated using F1 score as explained before. The result is as follows:

<table>
<thead>
<tr>
<th>Macroprudential Credibility Evaluation Index</th>
<th>Table 3</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Best Model Combination</strong></td>
<td><strong>Aspect Model F1 Score</strong></td>
</tr>
<tr>
<td>Formulation</td>
<td>XGBoost &amp; Decision Tree</td>
</tr>
<tr>
<td>Effectiveness</td>
<td>Logistics Regression &amp; XGBoost</td>
</tr>
<tr>
<td>Coordination</td>
<td>Logistics Regression &amp; SVM</td>
</tr>
<tr>
<td>Communication/Expectation</td>
<td>Logistics Regression &amp; Logistics Regression</td>
</tr>
<tr>
<td><strong>Overall Average</strong></td>
<td>70.19%</td>
</tr>
</tbody>
</table>

- Logistic regression algorithm is the most accurate in majority of cases (4 out of 8 models). This suggests that the relationship between sentence features and credibility labels is mostly linear, or that more (annotated) data is needed to be able to extract more accuracy from using nonlinear algorithms.
The model individually is able to separate between the 4 aspects well and then predict positive and negative sentiment quite well. But due to the high number of non-aspects and non-sentiment news sentences in our dataset, the F1 score quite suffers from misclassifying non-relevant sentence as relevant.

Since the model is a joint combination of 2 models the errors are also multiplied (50.81% average). The models need to have higher precision and F1 score for each of the classification task. The resulting error of the models is still an issue to be dealt with in the future development of the models.

4.2 Index Results

Our previous paper for monetary policy credibility by Zulen et al. (2020) compared the credibility index result with the monetary policy credibility survey. Since the macroprudential policy survey doesn’t have consistent and similar aspect compared with the big data credibility index, the two indexes do not have a good correlation. The semiannual index compared to the available credibility survey only have -23% correlation, while the monetary policy index compared to its survey counterpart have a high value of 79.7% correlation since its aspect is more similar and consistent.

Quarterly Macroprudential Credibility Index (2013 – 2021)

Semiannual Macroprudential Credibility Index (2013 – 2021)
We can also perform event analysis in the earlier 2013 to 2016 where the index result is more volatile and explain the drops of index results. Since we have granular news sentence prediction, we can also explain any spike or drop in the index with this methodology.

As seen from the figure 3 and 4, the red communication credibility index line is only available since 2016 where Bank Indonesia started to communicate macroprudential policy more intensively to the general public and leveraged macroprudential policy forward guidance in its board of governors meeting press release. Since then the overall index that has been calculated is more stable and relatively high with overall average of 82% during 2016 to the 1st quarter of 2021. We can say that communication of macroprudential policy to the general public affect the sentiment or tone of news regarding Bank Indonesia macroprudential policy positively, as captured in this macroprudential policy credibility index.

5. Conclusion & Future Works

5.1. Conclusion

Based on our previous study of measuring monetary policy credibility, we develop a measurement of macroprudential policy credibility using text mining and machine learning. Using macroprudential policy keyword, we filter news data from our Bank Indonesia institute Cyber Library that collects economic and financial news data from online and printed press. Then we develop machine learning models to predict whether a news sentence is relevant to macroprudential policy and if it’s relevant, what its sentiment is, is it positively or negatively describe Bank Indonesia policy and credibility. Using the aspect and sentiment model predictions we develop an index that can be used to measure Bank Indonesia macroprudential policy credibility from the news.

The model is evaluated using F1 score, while still leaving a room for improvement regarding its accuracy in predicting aspect relevancy and sentiment since the usage of 2 models combined its errors. The aspect relevancy models have an average F1 score of 70.19% while the sentiment models have an average F1 score of 72.40%.

The index result is compared to its survey counterpart but the resulting correlation is quite low since the survey mostly ask various specific questions regarding macroprudential policy while not consistently ask questions regarding the 4 aspects that are measured in the credibility index. The correlation between the credibility survey and the credibility index from machine learning is -23%. Bank Indonesia starts to communicate macroprudential policy forward guidance more intensively starting in 2016, and the result is captured in the index with the 4 aspects of the index become more stable and positive starting in 2016.

5.2. Future Works

Some possible improvement and possible further research include:

- **Data collection and annotation of more recent data**: As mentioned before, the accuracy of the model is not that robust (50.81% average F1 Score). In the future with more available data we can update the model with news from 2nd semester of 2021 and also 2022. Hopefully with using and annotating more data
we can collect more relevant and recent news data so that the developed model can be more robust and produce better F1 score.

- **Expansion of macroprudential and policy mix**: There may be keywords that relevant to Bank Indonesia macroprudential policy or policy mix that contains positive and negative sentiment in the news during the covid-19 pandemic and recovery in the 2nd semester of 2021 and 2022. There are news regarding inclusive financing and SME definition expansion from 2022 that can be considered relevant to macroprudential policy. Updating the keyword list will help the model to be relevant to novel macroprudential policy formulated by Bank Indonesia in specific or unprecedented times.

- **Econometric Analysis**: The index that developed here measures Bank Indonesia macroprudential policy credibility using sentiment in the news. Following Wibisono et al. (2022) paper, we can perform econometric analysis to analyze whether the index can capture metrics of financial stability such as non-performing loan rate and credit growth and also whether the index capture the effect of any changes on macroprudential policy.
References


Appendix A: Macroprudential Keyword

<table>
<thead>
<tr>
<th>Macroprudential Policy Keywords</th>
<th>Table A.1</th>
</tr>
</thead>
<tbody>
<tr>
<td>Policy mix</td>
<td>DP / Down Payment /Property credit down payment/ Vehicle Down Payment/ Mortgage</td>
</tr>
<tr>
<td>Macroprudential / Macroprudential + Microprudential</td>
<td>Green Financing</td>
</tr>
<tr>
<td>Financial System Stability</td>
<td>Macroprudential Intermediation Ratio</td>
</tr>
<tr>
<td>Statutory Reserve Requirement / Statutory Reserve Requirement Incentive</td>
<td>Macroprudential Liquidity Buffer</td>
</tr>
<tr>
<td>Country Cyclical Buffer (CCB)</td>
<td>Short Term Liquidity Assistance</td>
</tr>
<tr>
<td>Loan to Value (LTV)</td>
<td>Small and Micro Enterprise Credit Ratio</td>
</tr>
<tr>
<td>Financing to Value (FTV)</td>
<td>Macroprudential Inclusive Financing Ratio</td>
</tr>
</tbody>
</table>

Appendix B: Sample of Annotated Sentences

<table>
<thead>
<tr>
<th>Sample of annotated sentences</th>
<th>Table B.1</th>
</tr>
</thead>
<tbody>
<tr>
<td>Sentences</td>
<td>Credibility Aspect</td>
</tr>
<tr>
<td>Bank Indonesia has taken various anticipative measures to maintain macroeconomy and financial system stability and support sustainability of growth through monetary and macroprudential policy.</td>
<td>Formulation</td>
</tr>
<tr>
<td>Bank Indonesia policy to tighten rules of mortgage receive criticism from property developer.</td>
<td>Formulation</td>
</tr>
<tr>
<td>With the rising consumption, relaxation of Loan-to-value ratio that will be applied on 1 August 2018, will help the growth of credit to its highest level to 12% this year.</td>
<td>Effectiveness</td>
</tr>
<tr>
<td>According to Paramadina University Rector, Prof Firmanzah Ph.D, Bank Indonesia macroprudential policy is not effective yet to support financial system stability.</td>
<td>Effectiveness</td>
</tr>
<tr>
<td>According to hendar, BI is preparing MOU with OJK. The purpose is to prevent data supply interference after the separation of macroprudential and microprudential function.</td>
<td>Coordination</td>
</tr>
<tr>
<td>Real Estate Company Union assess that there needs to be synchronization of loan-to-value ratio relaxation with the taxation system.</td>
<td>Coordination</td>
</tr>
</tbody>
</table>
Indonesia Credit Default Swap (CDS) can rise and Rupiah exchange rate will face turbulency and volatility that trigger vulnerability in the market. This will force BI to face it with monetary and macroprudential tightening.

<table>
<thead>
<tr>
<th>Communication/Expectation</th>
<th>Contractive</th>
</tr>
</thead>
</table>

Bank Indonesia will hold its loan to value policy on the property sector because it’s considered an effective measures to reduce speculation and control credit risk.

<table>
<thead>
<tr>
<th>Communication/Expectation</th>
<th>Neutral</th>
</tr>
</thead>
</table>

BI also predicted will loosen its macroprudential activity in the near future.

<table>
<thead>
<tr>
<th>Communication/Expectation</th>
<th>Accommodative</th>
</tr>
</thead>
</table>

*The news samples above are translated by the authors from Bahasa Indonesia to English
Measuring Macroprudential Policy Credibility Using Machine Learning

Muhammad Abdul Jabbar, Okiriza Wibisono, Nursidik Heru Praptono, Alvin Andhika Zulen
IFC Biennial Conference on 25-26 August 2022

*The views expressed here are those of the authors and do not necessarily reflect the views of Bank Indonesia*
Outline

1. Background and Goals
2. Framework
3. Methodology
4. Results and Analysis
5. Conclusion and Future Works
1. **Background & Goals**

**Background:**

1. Since the Global Financial Crisis in 2007/2008, macroprudential policy has been proven to help maintain macroeconomic and financial stability.

2. Central Bank credibility is defined as a commitment to follow well articulated and transparent rules and policy goals (Bordo & Siklos, 2015). **Macroprudential policy credibility should also be taken into account for the overall Bank Indonesia’s credibility.**

3. Bank Indonesia used **semiannual survey to stakeholders** for measuring policy credibility,

4. Previous research has shown evidence that **news data and machine learning** can be used to measure Bank Indonesia’s monetary policy credibility (Zulen et al., 2020).

**Goals:**

Utilizing **news data** and machine learning to measure unbiased and timelier public perception on Bank Indonesia’s macroprudential policy in 4 aspects: Formulation, Effectiveness, Coordination, and Communication.

**Literature Study:**


2. Framework & Scope

**Aspect Scope**

**Formulation:** Perception regarding macroprudential policy formulation

**Effectiveness:** Perception regarding the effectiveness of macroprudential policy

**Coordination:** Perception regarding Bank Indonesia’s coordination with other authorities

**Communication:** Perception regarding Bank Indonesia’s communication on its’ policy stance

---

**Framework**

1. **Filtering**
2. **Annotation**
3. **Pre-processing**
4. **Formulation Aspect Model**
5. **Effectiveness Aspect Model**
6. **Coordination Aspect Model**
7. **Communication Aspect Model**

**Data Source**

- **Source:** Cyber Library (internal repository of curated economic and financial news)
- **~30 domestic news (in Bahasa Indonesia)**
- **~850 articles daily**
- **Whole corpus:** since Jan 1999
- **Training data:** Jan 2013 – Jun 2021

**Index Calculation**

- **Credibility Index:**
  1. Formulation
  2. Effectiveness
  3. Coordination
  4. Communication

---

**Sentiment Models**

- **Formulation Sentiment Model:**
  - Positive
  - Negative

- **Effectiveness Sentiment Model:**
  - Positive
  - Negative

- **Coordination Sentiment Model:**
  - Positive
  - Negative

- **Communication Sentiment Model:**
  - Ekspsivne
  - Neutral
  - Contractive
Macroprudential keywords filter:

1. Generic keywords e.g: Policy mix, macroprudential, financial system stability
2. Macroprudential policy instruments e.g: Countercyclical Buffer / CCB, Loan to value (LTV), Financing to value (FTV), Macroprudential Intermediation Ratio (RIM), etc
3. Credit and macroprudential indicators e.g: Down payment, vehicle credit, property credit, green financing, prime lending rate, etc
4. Other financial authorities: Financial Services Authority (OJK), Indonesia Deposit Insurance Corporation (LPS), Financial System Stability Committee (KSSK).
5. Others e.g: Integrated Reporting, Government bond primary market.

### Methodology: Annotation

- Sample of filtered sentences are annotated with positive (+1), negative (-1), neutral (0), or not relevant (-) labels for each credibility aspect. One sentence can contain more than 1 aspect.
- Annotation are done by authors and domain experts within BI, using guidelines provided (incl. examples).
- Every sentence is annotated by 2/3 persons. The label are decided by the majority labels (or decided in annotator forum).

<table>
<thead>
<tr>
<th>Sentences</th>
<th>Credibility Aspect</th>
<th>Label</th>
</tr>
</thead>
<tbody>
<tr>
<td>Bank Indonesia has taken various anticipative measures to maintain</td>
<td>Formulation</td>
<td>Positive</td>
</tr>
<tr>
<td>macroeconomy and financial system stability and support sustainability of</td>
<td></td>
<td></td>
</tr>
<tr>
<td>growth through monetary and macroprudential policy.</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Bank Indonesia policy to tighten rules of mortgage receive criticism</td>
<td>Formulation</td>
<td>Negative</td>
</tr>
<tr>
<td>from property developer.</td>
<td></td>
<td></td>
</tr>
<tr>
<td>With the rising consumption, relaxation of Loan-to-value ratio that will</td>
<td>Effectiveness</td>
<td>Positive</td>
</tr>
<tr>
<td>be applied on 1 August 2018, will help the growth of credit to its highest</td>
<td></td>
<td></td>
</tr>
<tr>
<td>level to 12% this year.</td>
<td></td>
<td></td>
</tr>
<tr>
<td>According to Paramadina University Rector, Prof Firmanzah Ph.D, Bank</td>
<td>Effectiveness</td>
<td>Negative</td>
</tr>
<tr>
<td>Indonesia macroprudential policy is not effective yet to support</td>
<td></td>
<td></td>
</tr>
<tr>
<td>financial system stability.</td>
<td></td>
<td></td>
</tr>
<tr>
<td>According to hendar, BI is preparing MOU with OJK. The purpose is to</td>
<td>Coordination</td>
<td>Positive</td>
</tr>
<tr>
<td>prevent data supply interference after the separation of macropurudential</td>
<td></td>
<td></td>
</tr>
<tr>
<td>and microprudential function.</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Real Estate Company Union assess that there needs to be synchronization</td>
<td>Coordination</td>
<td>Negative</td>
</tr>
<tr>
<td>of loan-to-value ratio relaxation with the taxation system.</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Indonesia Credit Default Swap (CDS) can rise and Rupiah exchange rate</td>
<td>Communication/</td>
<td>Contractive</td>
</tr>
<tr>
<td>will face turbulency and volatility that trigger vulnerability in the</td>
<td>Expectation</td>
<td></td>
</tr>
<tr>
<td>market. This will force BI to face it with monetary and macroprudential</td>
<td></td>
<td></td>
</tr>
<tr>
<td>tightening.</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Bank Indonesia will hold its loan to value policy on the property sector</td>
<td>Communication/</td>
<td>Neutral</td>
</tr>
<tr>
<td>because it’s considered an effective measures to reduce speculation and</td>
<td>Expectation</td>
<td></td>
</tr>
<tr>
<td>control credit risk.</td>
<td></td>
<td></td>
</tr>
<tr>
<td>BI also predicted will loosen its macroprudential activity in the near</td>
<td>Communication/</td>
<td>Accommodative</td>
</tr>
<tr>
<td>future.</td>
<td>Expectation</td>
<td></td>
</tr>
</tbody>
</table>

| Total Annotated Sentences: 5,030                                        |

<table>
<thead>
<tr>
<th>Aspect</th>
<th>Positive</th>
<th>Negative</th>
<th>Neutral</th>
<th>Irrelevant</th>
</tr>
</thead>
<tbody>
<tr>
<td>Formulation</td>
<td>642</td>
<td>70</td>
<td>4.318</td>
<td>(12,8%)</td>
</tr>
<tr>
<td>Effective</td>
<td>387</td>
<td>76</td>
<td>4.567</td>
<td>(7,7%)</td>
</tr>
<tr>
<td>Coordination</td>
<td>433</td>
<td>30</td>
<td>4.567</td>
<td>(8,6%)</td>
</tr>
<tr>
<td>Communication</td>
<td>72</td>
<td>585</td>
<td>4.348</td>
<td>(1,4%)</td>
</tr>
</tbody>
</table>
All the sentences are pre-processed to change the sentence from text (unstructured) form to structured TF-IDF Vector that can be further processed using machine learning.

### Pre-processing

1. **Sentence cleansing**
   - Replace synonyms
   - Replace numbers
   - Replace common names

2. **Tokenization**

3. **\( n \)-gram vectorization**

4. **Remove rarely occurring terms**

### Machine Learning Flow

- Annotated data
- Preprocessing & Feature Extraction
  - TF-IDF Bag of Words vectorisation
- Machine Learning

### Machine Learning Algorithm:

- Logistic Regression
- K-nearest Neighbor
- Support Vector Machine
- Naïve Bayes
- Decision Tree
- Random Forest
- XGBoost

### K-Fold Cross Validation

80:20 Training and Testing Data Split

### Tuning

- Oversampling & Undersampling on learning process using SMOTE
After performing a horse race of 7 machine learning algorithms, we obtained the best model combination for each credibility aspects (using F1-score as evaluation metric).

### Model Evaluation

<table>
<thead>
<tr>
<th>Credibility Aspect</th>
<th>Best Model Combination</th>
<th>F1 Score Details</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>Phase 1 Aspect Relevancy</td>
</tr>
<tr>
<td>Formulation</td>
<td>XGBoost &amp; Decision Tree</td>
<td>66.15%</td>
</tr>
<tr>
<td>Effectiveness</td>
<td>Logistics Reg. &amp; XGBoost</td>
<td>66.87%</td>
</tr>
<tr>
<td>Coordination</td>
<td>Logistics Reg. &amp; SVM</td>
<td>70.64%</td>
</tr>
<tr>
<td>Communication</td>
<td>Logistics Reg. &amp; Logistics Reg.</td>
<td>77.13%</td>
</tr>
</tbody>
</table>

### Average F1 Score:

- Phase 1: 70.19%
- Phase 2: 72.40%

### Index Calculation

#### Communication aspect formula:

\[ \text{Index}_{\text{Communication}, t} = 1 - |P_t - \text{Tone}_t| \]

#### Formulation, Effectiveness, and Coordination aspects formula:

\[ \text{Index}_{\text{aspect}, k,t} = \frac{\#\text{positive}_{k,t} - \#\text{negative}_{k,t}}{\#\text{positive}_{k,t} + \#\text{negative}_{k,t}} \]

#### Credibility Index:

\[ \text{Credibility Index}_t = \frac{1}{4} (\text{Index}_{\text{formulation}, t} + \text{Index}_{\text{effectiveness}, t} + \text{Index}_{\text{coordination}, t} + \text{Index}_{\text{communication}, t}) \]
3. Result: Analysis

- Earlier index from 2013 to 2016 has fewer news therefore the indexes are more volatile.
- Since 2016, Bank Indonesia started to communicate macroprudential policy more intensively in its Board of Governors meeting press release. Thus, the overall index is more stable and positive with overall average of 82%. Communication of macroprudential policy to the general public affect the sentiment or tone of news positively, as captured in this macroprudential policy credibility index.

Quarterly Credibility Index per Aspect

- Coordination with Financial System Stability Committee needed for exchange rate
- Increase of Non-performing loan due to LTV Policy
- LTV regulation has minimum effect
- Mortgage policy formulation is not optimal
- Coordination to handle LTV policy considered slow
5. Conclusion & Future Works

Conclusion

1. Based on our previous study of measuring monetary policy credibility, we develop a measurement of macroprudential policy credibility using news data and machine learning.
2. The aspect relevancy models have an average F1 score of 70.19% while the sentiment models have an average F1 score of 72.40%.
3. Bank Indonesia starts to communicate macroprudential policy forward guidance more intensively starting in 2016, and the result is captured in the index with the 4 aspects of the index become more stable and positive starting in 2016.

Future Works

1. Data collection and annotation of more recent data.
2. Expansion of macroprudential and policy mix keywords.
3. Econometric analysis (econometric effect of the indexes on macro indicators).
THANK YOU!
A probabilistic method for reconstructing the FDI network 
in search of ultimate hosting economies

Nadia Accoto, Valerio Astuti and Costanza Catalano, 
Bank of Italy

---

1 This presentation was prepared for the conference. The views expressed are those of the authors and do not necessarily reflect the views of the BIS, the IFC or the central banks and other institutions represented at the event.
A probabilistic method for reconstructing the FDI network in search of ultimate hosting economies

Nadia Accoto, Valerio Astuti, Costanza Catalano
Department of Economics, Statistics and Research, Bank of Italy

Abstract

Bilateral Foreign Direct Investment (FDI) statistics struggle at identifying the ultimate hosting economies of a given country, due to the non-negligible presence of conduit jurisdictions. At the same time, determining the ultimate destinations of FDI investments is crucial for understanding the real paths followed by FDI investments among increasingly interdependent economies. In this paper, starting from the Coordinated Direct Investment Survey (CDIS) data collected and published by the International Monetary Fund (IMF), first we reconstruct the global FDI network through clustering techniques. Then we provide a method for computing an (approximate) distribution of the ultimate hosting economies of a country by using a probabilistic approach on this network, based on Markov chains. In particular, we analyse the Italian case.¹

Keywords: foreign direct investments, FDI network, clustering, ultimate host economies, absorbing Markov chains

JEL classification: C51, C60, F23, G15

¹ We thank Marta Crispino, Valerio Della Corte, Andrea Del Monaco, Silvia Fabiani, Stefano Federico, Fadi Hassan, Giacomo Oddo, Giacomo Romanini, Alfonso Rosolia, Silvia Sabatini and Simonetta Zappa for useful discussions and suggestions. The views expressed in this paper are those of the authors and do not involve the responsibility of the Bank of Italy and/or the Eurosystem.
1. Introduction

Foreign Direct Investment (FDI) is a category of financial cross-border investment in which an investor (direct investor) of one economy makes an investment in an enterprise (direct investment enterprise) of another economy that allows having control or a significant degree of influence on the management of that enterprise. FDI statistics, which provide information on investments by immediate counterparts between two different countries, are key indicators of countries’ participation in the global economy. However, in a world that is more and more interconnected, such statistics are not sufficient to reconstruct the investment chains, due to the increasing presence of multinational enterprises and countries that act as tax havens or investment hubs. Indeed, investments can pass through the so-called Special Purpose Entities (SPEs), enterprises commonly created and registered in tax havens and investments hubs that allow tax optimization by channelling investments through economies, before arriving to the final investment recipient country. Therefore, a large portion of FDI transits in and out of some countries before reaching their final destination, producing no real economic value in the crossing country. This led international organizations and national compilers to consider the development of extended experimental statistics, such as inward FDI by Ultimate Investing Economies (UIE) and outward FDI by Ultimate Host Economies (UHE). Compiling FDI statistics by UIE and UHE, other than by immediate counterpart country, in fact, would make the FDI statistics more complete and useful from a macroeconomic point of view, highlighting who ultimately controls the investments, the ultimate destination, and the financial connections between economies. The fourth edition of the OECD’s Benchmark Definition of Foreign Direct Investment [1] recommends to compile inward investment positions according to the UIE; as of today, only few countries (Italy included) provide such additional information as experimental statistics. FDI statistics by UHE belong to the research agenda of international organizations such as the IMF ([3], §1.43) and the OECD ([1], §672). A specific guidance and methodology concerning the UHE concept is being developed in the scope of the international statistical manuals’ revision process ([4]), which is still ongoing. Recently, experimental methods have been developed by the FDI statistical community [5], [6], although no countries have yet published comprehensive statistics on that.

In this framework, we propose a model to estimate, for each country, the distribution of its outward FDI by UHE. The mathematical model is based on absorbing Markov chain, and has as input the reconstructed global outward FDI network.

---

2 By definition, the direct investor owns 10 percent or more of the voting power in the direct investment enterprise [1].
3 Countries whose economies are (almost) entirely dedicated to the provision of offshore services. In this paper we consider as tax havens the countries listed in the Balance of Payments vademecum [2].
4 Jurisdictions that facilitate transit of investments due to favourable tax and investment conditions.
5 Economies where the first investments are originated.
6 Final recipient economies of the investment chains.
7 Information on the residency of ultimate investor is collected directly from reporting firms in the annual FDI survey and then cross-validated with commercial databases (Orbis).
The data on outward FDI are taken from the Coordinated Direct Investment Survey (CDIS) database, where the IMF collects yearly, on a voluntary basis, information about FDI stocks from all the world’s countries. Even though the process of compiling the survey can have its own characteristics in each respondent country, our model should not be heavily affected by these differences, given that we exploit only proportions of outward investments from each country and not the absolute declared values. Since the reporting countries in the CDIS database are only 89 over 246 world’s countries, and some of the disseminated data are kept confidential by the reporting country itself, some imputation methods are necessary to reconstruct the full network. In particular, we make use of mirroring and clustering techniques. The Markov chain model (inspired by [7]) is then set up on this reconstructed network, providing, for each country, an estimate of the percentage of its outward FDI towards each other country as final recipient.

The paper is structured as follows: Section 2 contains the description of how the outward FDI network is fully reconstructed, addressing the problem of missing data, while Section 3 describes the mathematical model used to estimate the FDI distribution by UHE for every country. Finally, Section 4 reports the results of such model in the Italian case and Section 5 presents some conclusions and future work.

2. The outward FDI network

We consider the bilateral data on outward stocks of FDI by counterpart country on year 2019, taken from the CDIS database.

The data in the CDIS database are presented according to the Extended Directional Principle ([3], §6.42); a direct investment is shown as either an investment abroad (outward investment) or an investment in the reporting economy (inward investment). Direct investments abroad cover assets and liabilities between resident direct investors and their non-resident direct investment enterprises. Direct investment in the reporting economy includes all liabilities and assets between resident direct investment enterprises and their non-resident direct investors. Assets and liabilities between resident and non-resident fellow enterprises are classified as outward (inward) investment if the ultimate controlling parent is resident (non-resident). The data are broken down by financial instruments (equity and debt instruments). The CDIS reporting economies are primarily requested to provide data on inward FDI, but they can complement them also providing data on outward FDI. Our aim is to reconstruct the (weighted, directed) network of bilateral outward FDI between the world’s countries and to use it to gain information about the ultimate hosting economies of each country (in particular, of Italy). More specifically, we want to reconstruct the network having as nodes all the CDIS countries, and as (directed) links of weight \(w\) from country \(i\) to country \(j\) the total FDI stocks of \(w\) dollars from country \(i\) towards country \(j\). This network will be the starting point of the subsequent Markov chain model that will provide approximated FDI statistics by UHE. The main difficulties in reconstructing the network using CDIS data can be summarized as follows:

1. Not all countries provide data on their outward FDI;

---

8 Debt instruments refer to debt positions between affiliated enterprises, see also §6.26 [3].
2. due to the extended directional principle presentation, some reported imports are negative,\(^9\) while we need positive weights for the Markov chain model;

3. reporting countries can decide to flag some of their data as confidential, thus masking the value of the investment.\(^{10}\)

In particular, the reporting countries are only 89 over the 246 possible ones and the confidential data are 24.7% of the available data, spread over 37 reporting countries. The percentage of confidential data varies between countries: it goes from a maximum of 91.8% of confidential data for Hong Kong to a non-null minimum of 0.8% of confidential data for Guatemala, and no confidential data for 52 countries. See also Figure 1. We handled item 2. by considering the absolute value of such negative amounts.\(^{11}\) In the next sections, we describe the solutions adopted in order to handle items 1. and 3.

![Figure 1: Percentage of confidential data in the outward FDI database by country.](image)

\(^9\) FDI data presented according to the extended directional principle can be negative; this happens when the debt positions from the subsidiary to the parent company exceed the value of the investment of the parent in the subsidiary.

\(^{10}\) This expedient is often used to avoid that individual information are indirectly disclosed by deduction from reported data. In this case, we know that a link between two countries exists but we do not know its weight.

\(^{11}\) Setting these negative values to zero would have been in contrast with the information that such link exists.
2.1 Missing data: mirroring the inward FDI

The CDIS database provides both the bilateral inward and outward FDI statistics supplied by the reporting countries. The inward and outward database, in theory, should be symmetric, as for every countries A and B it should hold:

\[
\text{Outward FDI from A to B} = \text{Inward FDI of B from A} \quad (1)
\]

In real data, these two values might not coincide as the one on the left hand side of (1) is reported by country A, while the one on the right hand side is reported by country B. Property (1) can nonetheless be of help in (partially) solving items 1. and 3. listed above. Indeed, in the inward database the reporting countries result to be 119, thus providing many additional data with which we can complement the network. We complement the outward database with data taken form the inward database and not vice versa, despite the latter having more reporting countries than the former, because that is the direction of the investment we want to exploit in the model when looking for the ultimate recipient countries. In particular, we use the inward database to:

- impute the confidential data. If the outward FDI of country A to country B is confidential, we check the inward FDI reported by country B from A: if it exists and it is different from zero, we substitute the confidential data with this amount;
- add missing links. If the outward FDI of country A to country B is equal to zero or it is not reported, we check the inward FDI reported by country B from A: if it exists and it is different from zero, we add a link from A to B with this amount as weight.\(^{13}\)

After all these operations, we again consider the absolute values of the negative amounts that might have been substituted.

This procedure let us impute 667 values over the 1918 starting confidential cells, and it let us add 7771 links (921 links that were reported as zeros in the outward FDI database plus 5915 extra links referring to non-reporting countries in the outward FDI database). The final network has 246 nodes (the total number of CDIS countries) and 14607 links, 3248 of which are still confidential (22.2% of the total data different from zero).\(^{14}\) The network results to be strongly connected, i.e. there is a path of outward FDI from any country to any other country, and quite sparse (sparsity score=0.759, i.e. only 24.1% of links over all the possible ones are present). The next section describes how we impute the values of the remaining confidential data.

2.2 Missing data: clustering techniques

After the initial imputation described in the previous section, we still have a high number of confidential data, which could potentially distort the results of the analysis.

\(^{12}\) All the reporting countries of the outward database are also reporting countries in the inward database but not vice versa.

\(^{13}\) Note: in this case, we also keep the inward FDI data that are confidential, as they signal the presence of a link that was not reported in the outward FDI database.

\(^{14}\) The percentage of confidential data is increased because we have complemented the outward FDI database also with the confidential data from the inward FDI database.
We propose in this section a second step of imputation, based on the proximity of countries in the "outward investment space". More specifically, each country $i$ defines a real-valued vector $\hat{v}_i$ in a 246-dimensional space, where each $j$-th component $v_i^j$ of the vector denotes the amount invested by country $i$ in country $j$. All of these components have non-negative values, such that the vectors span only a very small wedge of the embedding space. In addition, countries are obviously very different in terms of the total amount invested - which is linked to the magnitude of these vectors - but useful information on the investment pattern is contained in the ratios between the components. For this reason we define two different algorithms to study the similarity of the investment vectors: first, we consider the scalar product of the normalized investment vectors for two given countries $i$ and $j$:

$$
\sigma_{ij} = \hat{v}_i \cdot \hat{v}_j = \sum_{k=1}^{246} \frac{v_i^k v_j^k}{\|\hat{v}_i\| \|\hat{v}_j\|}
$$

(2)

where for any vector $\hat{v}$, we define its norm $\|\hat{v}\| = \left(\sum_k v^k \cdot v^k\right)^{1/2}$. As an alternative, we consider the scalar product between the normalized vectors after having subtracted their means:

$$
\rho_{ij} = \tilde{v}_i \cdot \tilde{v}_j = \sum_{k=1}^{246} \frac{v_i^k v_j^k}{\|\tilde{v}_i\| \|\tilde{v}_j\|}
$$

(3)

where $\tilde{v}_i^k = \frac{(\hat{v}_i^k - \bar{v}_i)}{\|\hat{v}_i - \bar{v}_i\|}$ and $\tilde{v}_j^k = \frac{(\hat{v}_j^k - \bar{v}_j)}{\|\hat{v}_j - \bar{v}_j\|}$ are the standardized vectors and $\hat{v}_i$ and $\hat{v}_j$ are the mean vectors.

While in principle leading to different results, the two proximity scores are in practice very similar for most of the countries; the reason is that many investment vectors have a large number of zero components, and few very large ones. This implies that subtracting the mean changes very little in terms of the components that most heavily influence the proximity scores. The underlying assumption of our imputation strategy is that countries that are close in this space have a similar investment pattern, such that, if country $i$ has a confidential value for investments towards a given country $k$, the average proportion of investment towards country $k$ from a set of countries which are close to $i$ will function as a good estimator for the missing value.

Clearly, the existence of missing values will have an influence also on the evaluation of the proximity score, in that we ignore the portion of the investment pattern which is hidden in confidential data. If for a given country the number of confidential entries is too large, we will probably have a distorted representation of the behaviour of the country. In the evaluation of the proximity score, neglecting the confidential values altogether would be equivalent to assuming that no investment is present in these cases. This assumption seems however unrealistic, in that one would think there are more reasons to have confidentiality on something existing, than on an absent investment. For this reason, to represent confidential data we choose a small, positive number - for definiteness fixed at 1.15 Whenever we encounter a confidential value for the investment from country $i$ to country $j$, we select the $N_{close}$

15 The exact value of this placeholder does not play any role in the clustering, and is exploited only for bookkeeping reasons.
countries with the highest values of the proximity scores with \( i \) having a non-confidential value for the investments towards \( j \) (such that the list of \( N_{\text{close}} \) countries can be different for any imputed confidential value). In addition we put a minimum threshold on the proximity score for the country to be considered as "close" to the one with the confidential value.\(^{16}\) From this sample of neighbour countries, we compute the average of the proportion of the investments towards country \( j \) (i.e. the investment towards country \( j \) divided by the total outward investments). If the sample of countries selected is empty (for example, if every country with the proximity score above the chosen threshold has also a missing value for country \( j \)), we impute the confidential value with a null value. For each country, we know the total outward FDI stocks toward the rest of the world. From this and the average proportions, it is easy to derive the absolute value of the imputed outward links. In particular, we impose for the sum of the outward FDI (both previously known and imputed) to be equal to the known total outward FDI of each country.

To test the performance of the algorithm, we consider two out-of-sample tests: first, we select a pair of countries and hide the value of the investment from the first to the second. We impute the hidden values, and evaluate the difference between the imputed value and the real ones. As a second test, we hide 10 outward investments of a single country and impute them. In this way, we can evaluate the correlation between the series of true values and the series of imputed ones. With \( N_{\text{close}} = 3 \), the percentage of zero values after the imputation procedure is close to 5%. For comparison, 75% of non-confidential values in the data-set are null. We report some summary statistics of the error distribution of the imputation process in Table 1, excluding the cases in which the imputed value is equal to zero. The absolute error is simply the difference between the imputed proportion of investment toward a given country, and the true value. To have a comparative measure of these values, the average proportion, corresponding to the inverse of the number of countries in the network, is \( 246^{-1} \approx 0.0041 \). For the correlation between the series of imputed and true values, we exclude the cases in which the series of true values is composed of null values only.

\(^{16}\) The proximity score \( \rho_{ij} \) can assume negative values, so when using it to establish the proximity between countries we impose the constraint \( \rho_{ij} > 0 \). In Table 1 we report summary statistics of the proximity score between each country and its selected neighbours.

<table>
<thead>
<tr>
<th>Proximity</th>
<th>Mean</th>
<th>25-th percentile</th>
<th>Median</th>
<th>75-th percentile</th>
</tr>
</thead>
<tbody>
<tr>
<td>Absolute error (10^{-3})</td>
<td>-0.39</td>
<td>-0.49</td>
<td>0.00</td>
<td>1.20</td>
</tr>
<tr>
<td>Correlation</td>
<td>0.53</td>
<td>0</td>
<td>0.69</td>
<td>0.99</td>
</tr>
</tbody>
</table>

Table 1: Summary statistics of the imputation process results

3. In search of ultimate host economies

Once we have reconstructed the network, we can proceed to build the model to estimate the distribution of FDI by ultimate host economies. We here provide a description on how such model works.
Conduit FDI arises when a multinational enterprise investing from home country A in host country B establishes an intermediate step through a third country C. This intermediate step is merely financial, as in country C no real "productive" investment takes place, and it is generally qualified as conduit FDI. In the model we allow for a conduit component in each country, representing the percentage of FDI received passing through the country. We now want to simulate the investment process on the FDI network following the investment from the investor to the final recipient. Starting from country A (the origin of the investment), we know from bilateral data the distribution of the investments of A towards its recipient countries (the out-neighbours of A in the network). If \( w_{A,B} \) represents the magnitude of the outward FDI of country A towards B, and \( s_A = \sum_B w_{A,B} \) is the total outward FDI coming from country A, we can say that an investment from A has a probability of \( \pi_{A,B} = w_{A,B}/s_A \) to be invested in country B. Suppose now that the investment has been indeed made towards country B: it either stays there (country B behaves as a non-conduit jurisdiction) and the investments chain stops making B the UHE of A, or country B behaves as a conduit and makes the investment pass through towards another country. The investment will pass from country B to a country C in the network according the probability \( \pi_{B,C} \) computed from country B’s FDI bilateral data.\(^{17}\) Again, either the investments stops there, making C the UHE, or country C behaves as a conduit and let the investment pass through towards another country. Eventually, the investment chain will stop in the final recipient economy, that is the UHE.

What we have just described is the behaviour of a random walk on (a modified version of) the FDI network. In the next sub-section we provide the rigorous mathematical framework of this model.

3.1 The Markov chain model

Consider the FDI network and suppose for each country/node A to be split into its conduit component \( A_c \) and its non-conduit component \( A_{nc} \), thus producing a total of \( 246 \cdot 2 = 492 \) nodes. To each country B we associate a probability \( q_c(B) \) to act as a conduit: how we estimate this parameter for every country will be addressed in the next section. Here we underline that we are making the strong assumption that the percentage of passing-through investments of a country B does not depend on the country from which it receives the investment: this assumption comes from the scarcity of the available data, as will be explained in the next section.\(^{18}\)

We now proceed to define a (discrete-time) Markov chain process\(^{19}\) on such augmented network by specifying the transition probabilities between any pair of nodes. The nodes of the network, according to standard notation, will be also referred to as the states of the Markov chain. We remind that \( \pi_{A,B} \) represents the percentage of outward FDI from country A to country B. For each pair of nodes \( (A_i, B_j) \), where

---

\(^{17}\) It is not guaranteed that conduit investments will follow the same distribution \( \pi_{A,c} \) as non-conduit ones; being unable to estimate the particular distribution of conduit investments, we make an uninformative assumption and conflate the two (in principle different) distributions.

\(^{18}\) A generalized model where such probabilities depend also on the investing country will be object of future work.

\(^{19}\) For formal definitions see [8].
\(i, j \in \{c, nc\}\), we define the probability \(p_{A_i B_j}\) that an investment passes from node \(A_i\) to node \(B_j\) in the following way:

- If \(i = nc\), then \(p_{A_{nc} B_j} = 1\) if \(B_j = A_{nc}\), otherwise \(p_{A_{nc} B_{nc}} = 0\).
  
  This represents the fact that country A acts as a non-conduit, so the investment remains in such country and does not pass through.

- If \(i = c = j\), then \(p_{A_{c} B_{c}} = \pi_{A,B} \cdot q_c(B)\).
  
  This represents the fact that A is investing on B and B will let the investment pass through (acts as a conduit).

- If \(i = c \text{ and } j = nc\), then \(p_{A_{c} B_{nc}} = \pi_{A,B} \cdot (1 - q_c(B))\).
  
  This represents the fact that A is investing on B and B does not let the investment pass through (acts as a non-conduit), i.e. it is the final recipient.

We call the non-conduit states \(A_{nc}\) an absorbing states, because once we enter that state we cannot leave. All the other states \(A_c\) are called transient. We can then arrange all the above probabilities in a matrix \(P = [p_{A_i B_j}]\) of dimension 492x492 ordering first by the absorbing states and then by the transient ones. The matrix \(P\) will then have the following structure:

\[
P = \begin{bmatrix}
\text{abs} & \text{trans} \\
I & 0 \\
R & Q
\end{bmatrix}
\]  

Where \(I\) is the 246x246 identity matrix, \(R\) represents the transition probabilities \(p_{A_{nc} B_{nc}}\) from transient states to absorbing states, and \(Q\) represents the transition probabilities \(p_{A_{c} B_{c}}\) between transient states. This defines a so-called absorbing Markov chain. The initial state of the Markov chain is any stochastic vector \(\eta\) of length 492. In particular, if we want the model to start from a given country A, the vector \(\eta\) will have all zero entries but \(\eta_{A_{nc}} = 1\). This initial condition on transient states is a mere modeling expedient to ensure that the investment moves towards another country in the first step (as otherwise it would not be a foreign direct investment).

In an absorbing Markov chain, a random walk starting from any of the states will eventually end up in one of the absorbing states (and remain there forever). The following theorem provides the long-run distribution of transition probabilities on an absorbing Markov chain (see also [9]).

**Theorem.**

The limiting distribution of an absorbing Markov Chain with transition matrix as in (4) is given by:

\[
\hat{P} = \lim_{n \to \infty} P^n = \begin{bmatrix}
\text{abs} & \text{trans} \\
(I - Q)^{-1}R & 0
\end{bmatrix}
\]  

\(20\) An absorbing state is a state that is impossible to leave (no outgoing links). An absorbing Markov chain is a Markov chain with at least one absorbing state that is reachable from every state.

\(21\) Entry-wise nonnegative vector whose entries sum up to 1.
The interpretation of the limiting matrix $P^*$ is the following: for any country that acts as source of the investment (the rows of the matrix), the process will end up after a sufficiently large number of investment steps, with probability 1, in one of the UHE, modeled as absorbing states (the columns of the matrix). In particular, if we want to retrieve the UHE distribution of a country A, we just need to take the row indexed by $A_c$ of the fundamental matrix $(I - Q)^{-1}R$: this is a distribution by UHE over the non-conduit version of the world’s countries. In other words, the share of FDI from country A that ends up in country B as final recipient is $P^*_{A_cB_{nc}} = ((I - Q)^{-1}R)_{A_cB_{nc}}$, i.e. the entry of $P^*$ on row $A_c$ and column $B_{nc}$; this is equivalent to perform the vector-matrix multiplication $\eta^T P^*$, with $\eta_{A_c} = 1$ and all its other entries equal to zero.

3.2 Estimation of the conduit parameters of the model

We here describe, for each country A, how we estimate the parameter $q_c(A)$ i.e. the probability that an investment arriving in A would pass through it heading to another country.

As we have already mentioned in the introduction, most conduit FDI in the world take place through a limited set of jurisdictions that act as global FDI hubs. Such hubs can be divided into two groups:

a) the tax havens, that are small jurisdictions whose economy is entirely, or almost entirely, dedicated to the provision of offshore financial services;

b) other investment hubs, that are countries that have a substantial real economic activity but also act as conduit jurisdictions due to their favourable tax and investment regime.

We consider as tax havens 38 countries, listed by the European Commission in the BoP Vademecum [2], plus 4 other countries identified by Casella in [7]. Note that we consider Hong Kong and Singapore belonging to group (b) instead of group (a), despite appearing in the BoP Vademecum, due to their relevant size in term of population, comparable to other investments hubs such as Luxembourg. Since tax havens act fully as conduit jurisdictions, we associate to them a probability $q_c = 1$.

Regarding the investments hubs belonging to group (b), there are some countries that report their yearly outward FDI investments made through Special Purpose Entities (SPEs). These data can then be used to compute their conduit probability as the ratio between the inward FDI through SPEs and total inward investment:

$$q_c(A) = \frac{\text{Inward FDI of } A \text{ in resident SPEs}}{\text{Total Inward FDI of } A}.$$  

22 Andorra, Anguilla, Antigua and Barbuda, Aruba, Bahamas, Bahrain, Barbados, Belize, Bermuda, Cayman Islands, Cook Islands, Curaçao, Dominica, Gibraltar, Grenada, Guernsey, Isle of Man, Jersey, Lebanon, Liberia, Liechtenstein, Marshall Islands, Mauritius, Montserrat, Nauru, Niue, Panama, Philippines, Saint Lucia, Saint Vincent and the Grenadines, Samoa, Seychelles, Sint Maarten, St Kitts and Nevis, Turks and Caicos Islands, Vanuatu, British Virgin Islands, U.S. Virgin Islands.

23 Malta, Monaco, Netherlands Antilles, San Marino.

In particular, as conduit probability we take the average of the above ratio on the years 2017-2019.\textsuperscript{25} The reason why we use in the computation of the conduit probability \((6)\) for country A its inward FDI is that we are trying to estimate the percentage of investments made to country A that will pass through it towards other countries, relative to the total investments received. The computed conduit probabilities of such self-reporting countries are reported in Table 2. To them we added four countries, namely United Kingdom, Ireland, Hong Kong and Singapore, where we consider as conduit probability the values estimated by Casella \cite{7}, through a regression method based on GDP\textsuperscript{26} (end of Table 2). The countries listed in Table 2 are the countries that we consider belonging to group (b).

To all the countries that belong neither to group (a) nor to group (b) we assign a conduit probability of $0\%$, thus implying that they always act as final recipient.

\begin{table}[h]
\centering
\begin{tabular}{|c|c|}
\hline
Country & \(q_c\) (\%) \\
\hline
Luxembourg & 94.94 \\
Netherlands & 67.16 \\
Hungary & 57.28 \\
Switzerland & 20.39 \\
Denmark & 18.90 \\
Austria & 18.81 \\
Sweden & 7.30 \\
Spain & 6.41 \\
Belgium & 6.10 \\
Portugal & 5.83 \\
Norway & 4.95 \\
Iceland & 4.17 \\
Lithuania & 3.78 \\
Finland & 3.19 \\
Estonia & 2.77 \\
Chile & 0.91 \\
South Korea & 0.41 \\
Poland & 0.31 \\
Hong Kong & 78.94 \\
Ireland & 60.31 \\
Singapore & 25.08 \\
United Kingdom & 20.16 \\
\hline
\end{tabular}
\caption{Conduit probability of self-reporting SPEs and estimated hubs.}
\end{table}

4. The Italian case: results

The weighted FDI network reconstructed in Section 2 and the conduit probabilities computed in Section 3.2 are all the ingredients we need to run our Markov chain

\textsuperscript{25} If some annual data are missing, we consider the average on the years with available data.

\textsuperscript{26} Since Casella in \cite{7} uses the outward FDI to compute such probabilities, for each country A we corrected the values by the coefficient \((\text{Total Outward FDI of A}/\text{Total Inward FDI of A})\). Singapore does not report its outward FDIs, so we used the sum of the imputed values.
model (Section 3.1). In particular, we focus on the Italian case; Figure 2 and 3 compare the Italian bilateral data on outward FDI from the CDIS database (presented as percentage over the total outward FDI) with the results we obtain from the model in terms of UHE (percentage of outward FDI towards final recipient countries).

We can notice that countries such as The Netherlands and Luxembourg, which are ranking respectively first and fifth according to Italian bilateral data, often let these investments pass through to other destinations: indeed, in the Italian FDI distribution by UHE, their ranks drop respectively to the sixth and 22th position. This result was somehow expected, as both countries are characterized by favorable tax regimes and by the presence of a large number of SPEs in their territory. These countries are also the ones that show the biggest (absolute) percentage difference between bilateral data and the model output by UHE (see Figure 4).

Contrarily, countries such as the United States, Germany and France show the opposite behaviour: the volume of Italian investments that they receive as ultimate recipient is larger than what is reported in bilateral data. This means that some investments originated in Italy have been channeled through investments hubs and/or tax havens before ending up to such countries. Figure 3 shows that the main Italian partners in terms of final recipient of investment chains are the United States, Germany, Spain and France (first four positions), while ranking respectively 3rd, 4th, 2nd and 6th when considering bilateral data.

It is also interesting to observe that Italy itself appears as an Italian UHE: the results indeed show that a small percentage of Italian FDI (around 0.5%) returns to Italy, highlighting the presence of round-tripping phenomena.27 Finally, the results obtained at the country level, in particular for Luxembourg, The Netherlands and USA, are reflected in the aggregation across main areas (see Figure 5).

![Bar chart](image.png)

**Figure 2:** Comparison between the Italian bilateral FDI from CDIS and the estimated distribution of Italian FDI by ultimate host economies, sorted in decreasing order by bilateral data. Only values over 1% are displayed.

---

27 Round-tripping refers to capital that leaves the economy and then goes back to it, see also §6.46 [3]
Figure 3: Comparison between the Italian bilateral FDI from CDIS and the estimated distribution of Italian FDI by ultimate host economies, sorted in decreasing order by ultimate host economies. Only values over 1% are displayed.

Figure 4: Difference between the estimated distribution by ultimate host economies and the Italian bilateral FDI. Only differences over 0.5 in absolute value are displayed.
5. Conclusions and future work

In this paper we first presented a methodology on how to reconstruct the full outward FDI network starting from the (incomplete and with confidentiality issues) CDIS database. Second, we proposed a model to estimate the FDI distribution by ultimate host economies for any given country. The results of the model in the Italian case show that some of the main Italian partners in terms of bilateral FDI receive much smaller volumes of investments as final recipients.

Future work would involve testing the robustness of the model for the different assumptions that were made, as well as possibly validating it by comparing the results with other experimental FDI statistics by UHE. Moreover, we plan to consider only the equity part of the FDI data to build (and reconstruct) the network, which should let us avoid negative weights and then maybe obtain a more realistic network. It would be of interest also to refine the model by considering conduit probabilities that depend also on the country making the investment, and not only on the receiving country. Also the search of newer techniques to compute the conduit probabilities would make it possible to extend the list of investments hubs and provide more reliable estimates. Finally, we plan to perform a deeper analysis on the FDI network in terms of connectivity, resistance to shock propagation and centrality measures [10].
References


A probabilistic method for reconstructing the FDI network in search of ultimate host economies

Nadia Accoto, Valerio Astuti, Costanza Catalano
26/08/2022 11th IFC Biennial Conference, Basel

Department of Economics, Statistics and Research - Bank of Italy
Foreign Direct Investments (FDI) and CDIS database

**FDI**: When an investor of one economy makes an investment in an enterprise of another economy, that allows to have control or a significant degree of influence. [Source: IMF]

The Coordinated Direct Investment Survey (CDIS)

- Promoted by IMF, annual and voluntary
- Inward FDI (investments received) and Outward FDI (investments made) by counterpart economy

Investments hubs and tax havens: facilitate transit of investments due to favourable tax regimes and off-shore services.

**Goal of MultiNational Enterprises**: Tax optimization by channeling investments through several economies → **Goal of MultiNational Enterprises**: Tax optimization by channeling investments through several
Ultimate investing economy and Ultimate host economy

Difficult to interpret FDI statistics by immediate partner economy as it does not show the ultimate sources and destinations of FDI.

**Ultimate investing economy (UIE):** where the investment originated

**Ultimate host economy (UHE):** final recipient of the investment

- who ultimately controls the investments/ultimate destination
- reveals the financial connections between economies
- info on businesses using offshore centers

Difficult to collect/provide data on UHE (no data available).
The Outward FDI Network

246 countries represented by their capitals. Reporting countries: 89
Orange: Tax havens, White: investment hubs, Blue: all the others

Red links: Confidential data. The import of the investment is censored (∼24.7% of the reported data)
Reconstructing the network

Imputation of confidencials and adding missing links:

- **Mirroring the Inward FDI database:**
  Outward of A by counterpart B = Inward of B by counterpart A

Imputation of remaining confidencials:

- **Clustering techniques:** "countries that are similar have the same outward ratios"

1. We cluster the countries according to how similar their outward FDI ratio are;
2. If the FDI of country A towards B is confidential, we impute it by averaging the FDI towards B of the first $N_{close}$ countries of A that report such data.
In search of UHE: the Markov chain model

We simulate the investment process on the FDI network following the investment from the investor to the final recipient.¹

We allow for a conduit component \( q_c(j) \) in each country \( j = \% \) of FDI received passing through it:

\[
q_c(j) = \frac{SPE_{out}(j)}{FDI_{out}(j)}
\]

[OECD statistics, average over 3 years.]

**Tax haven:** \( q_c(j) = 1. \)
All the investments pass through, no real investment is made in such countries.

**Otherwise** \( q_c(j) = 0: \) no investment passes through.

<table>
<thead>
<tr>
<th>Country</th>
<th>( q_c ) (perc)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Luxembourg</td>
<td>94.42%</td>
</tr>
<tr>
<td>Hungary</td>
<td>80.92%</td>
</tr>
<tr>
<td>Netherlands</td>
<td>59.83%</td>
</tr>
<tr>
<td>Austria</td>
<td>17.33%</td>
</tr>
<tr>
<td>Lithuania</td>
<td>12.21%</td>
</tr>
<tr>
<td>Denmark</td>
<td>11.66%</td>
</tr>
<tr>
<td>Switzerland</td>
<td>10.68%</td>
</tr>
<tr>
<td>Portugal</td>
<td>9.08%</td>
</tr>
<tr>
<td>Belgium</td>
<td>7.72%</td>
</tr>
<tr>
<td>Spain</td>
<td>7.48%</td>
</tr>
<tr>
<td>Estonia</td>
<td>7.33%</td>
</tr>
<tr>
<td>Iceland</td>
<td>6.73%</td>
</tr>
<tr>
<td>Norway</td>
<td>3.45%</td>
</tr>
<tr>
<td>Sweden</td>
<td>3.37%</td>
</tr>
<tr>
<td>Poland</td>
<td>2.42%</td>
</tr>
<tr>
<td>Chile</td>
<td>2.11%</td>
</tr>
<tr>
<td>Finland</td>
<td>0.14%</td>
</tr>
<tr>
<td>South Korea</td>
<td>0.01%</td>
</tr>
<tr>
<td>Hong Kong</td>
<td>78.00%</td>
</tr>
<tr>
<td>Ireland</td>
<td>65.00%</td>
</tr>
<tr>
<td>Singapore</td>
<td>57.00%</td>
</tr>
<tr>
<td>United Kingdom</td>
<td>21.00%</td>
</tr>
</tbody>
</table>

¹see also B. Casella, *Looking through conduit FDI in search of ultimate investors*, UNCTAD
In search of UHE: the Markov chain model

We simulate the investment process on the FDI network following the investment from the investor to the final recipient.

1. Starts from country $i$: the investment has a probability of ($\%$FDI from $i$ to $j$) to be invested in country $j$;
2. it either stays in $j$ ($j$ is the UHE of $i$) with prob $(1 - q_c(j))$ or it passes through to another country with prob $q_c(j)$ ($j$ behave as a conduit);
3. Reiterate steps 1. and 2. from $j$.

$(\%$FDI of country $i$ towards country $j) \times$ (prob country $j$ acts as a conduit)

$(\%$FDI of country $i$ towards country $j) \times$ (prob country $j$ does not act as a conduit)
In search of UHE: the Markov chain model

Eventually the investment chain (random walk) will end up in one non-conduit node, which will be the final recipient economy (UHE).

⇒ Absorbing Markov chain.
We can retrieve the final FDI distribution by UHE for each country.

Theorem

\[
P = \begin{bmatrix}
\text{abs} & \text{trans} \\
I & 0 \\
R & Q
\end{bmatrix}
\xrightarrow{\text{as } n \to \infty} P^* = \begin{bmatrix}
\text{abs} & \text{trans} \\
I & 0 \\
0 & 0
\end{bmatrix}
\]

\[
(I - Q)^{-1}R
\]

UHE distribution
Comparison between Italian FDI bilateral data (blue) and the estimated distribution of Italian FDI by UHE (green), sorted in decreasing order by bilateral data. Only values over 1% are displayed.
Thank you for your attention!

Questions?
Banks' real estate exposures:
Risk-based approach to measurement of exposures and concentrations

Patrick Slovik and Farah Azman

1 This presentation was prepared for the conference. The views expressed are those of the authors and do not necessarily reflect the views of the BIS, the IFC or the central banks and other institutions represented at the event.

2 Lead author
Banks' Real Estate Exposures: Risk-Based Approach to Measurement of Exposures and Concentrations

Patrick Slovik and Farah Azman

Abstract
The study develops a holistic risk-based approach to measuring banks' real-estate exposures and real-estate sector concentrations. As banks' real-estate exposures increasingly extend beyond traditional lending, adequate statistics shall cover broader types of on-balance sheet exposures and off-balance sheet exposures. The study describes a novel approach to measuring banks' exposures and concentrations utilising a risk-based approach aligned with the more granular post-crisis reforms of the Basel standards. It outlines a method of integrating the revised Basel standards with statistics on banks' sectoral exposures and concentrations and proposes refined metrics based on credit risk-weighted assets.

Keywords: Basel Accords; Bank credit; Commercial real estate; Concentration risk; Credit risk; Data gaps; Financial crisis; Real estate; Residential real estate; Sector concentration

JEL Classification: G01, G21, G28, G31, G32
Banks' Credit Risk and Sector Concentrations

Historical experience shows that the concentration of credit risk in banks' portfolios has been one of the main causes of bank distress (BCBS, 2006). Sector concentrations in banks' portfolios arise from excessive exposure to a sector or several highly correlated sectors. The Basel Accords' Pillar 1 approach has been portfolio invariant, i.e. appraising the risk of a single exposure without consideration for the portfolio's structure or concentrations. The business sector concentration risk thus warrants a focus on appropriate methodologies for measuring sector concentrations.

Banks' credit risk concentrations, particularly real estate exposures, exert a material impact on the soundness of the financial system. Distress in the real estate sector propelled numerous financial crises globally.1,2 Such financial crises either originated directly in real estate sector distress or real estate sector distress aggravated the severity and length of financial crises. Real-estate-linked financial crises were not limited to a particular category of real estate, as their causes stem from exposures to both residential real estate and commercial real estate (relatively equitably).1,3

While banks' real estate financing has been traditionally dominated by loans, other types of on-balance sheet and off-balance sheet exposures have increased in prominence. The share of conventional lending to total banks' assets has declined over time (Slovik, 2012). This is attributable to broadening financial innovation and deepening market-based finance. In response to the impact of financial innovation and market-based finance, authorities should adjust measurements to accommodate changes in financial structures of financial institutions (Lumpkin, 2010).

Extending Measurement Scope of Sector Exposures

Adequate measurement and analysis of banks' real estate exposures require extending their scope to include all types of on-balance sheet and off-balance sheet exposures to the real estate sector. The measurement and disclosure of broadly defined sector exposure categories and their inclusion in sector concentration aggregates remains scant, with most sector-specific research, analysis, and dashboards still predominantly based on loans. Consolidated measurement of sector exposures covering all on-balance sheet and off-balance sheets exposures can be applied to any sector.4

<table>
<thead>
<tr>
<th>Consolidated Sector Exposures</th>
<th>( \sum_{i=1}^{n} On_{SE_i} + \sum_{j=1}^{m} Off_{SE_j} \times CCF_j )</th>
</tr>
</thead>
<tbody>
<tr>
<td>Legend: CSE – Consolidated Sector Exposures; On_SE – On-Balance Sheet Sector Exposure; Off_SE – Off-Balance Sheet Sector Exposure; and CCF – Credit Conversion Factor.</td>
<td></td>
</tr>
</tbody>
</table>

1 Based on a review conducted by the authors of 20 different financial crises globally between 1970 and 2020, covering countries in North America, Europe, and Asia.
2 Similarly, the Basel Committee for Banking Supervision's review of bank failures concluded: "Credit concentration risk, usually in real estate, was cited in 9 out of 13 episodes." (BCBS, 2004).
3 The European Systemic Risk Board offers similar conclusions: "Adverse real estate market developments in some member states, both in residential real estate and commercial real estate, resulted in large losses in the past and negatively impacted the real economy." (ESRB, 2017).
4 Sectors represent business sectors, although the methodology can also be applied to other segments exhibiting default dependencies, such as geographical regions, several highly correlated sectors, or sub-segments of specific business sectors (for instance, residential real estate or commercial real estate).
Common on-balance sheet sector exposures cover lending and investments but may include other relevant types of on-balance sheet credit exposures. In the case of real estate, common categories include residential real estate loans, commercial real estate loans, loans to real estate companies, investment in real estate securities, direct investment in real estate or property acquired in settlement of debt, and may also include other types of material exposure categories (while the scope commonly excludes infrastructure projects or socially beneficial real estate projects).

Measurement and disclosure of off-balance sheet sector exposures are commonly missing even in relatively advanced contemporary frameworks and dashboards. This omission represents a material data gap. Off-balance sheet exposures to a particular sector (such as real estate) could be sizable, often exceeding more observable types of on-balance sheet sector exposure categories. Off-balance sheet items can be converted into credit-exposure equivalents through the use of harmonised credit conversion factors that were developed as part of the BCBS’s Basel Accords.

**Enhancing Risk Sensitivity of Sector Exposures**

While the consolidated sector approach defines an all-inclusive scope, it might not be sufficiently risk-sensitive. Risk sensitivity recognises that same-sized credit exposures do not necessarily have the same credit risk. The credit risk of real estate exposures relates to factors such as loan-to-value (LTV) ratios, counterparties’ external credit ratings, whether the property is income-producing or self-occupied, or exposures’ asset quality. Credit risk weights derived from the harmonised BCBS’s Basel Accords can be utilised to obtain credit risk-weighted consolidated sector exposures.

<table>
<thead>
<tr>
<th>Credit Risk-Weighted Consolidated Sector Exposures</th>
<th>Equation (2)</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\text{CRWA(CSE)} = \sum_{i=1}^{n} \text{CRW}<em>i \times \text{On}</em>\text{SE}<em>i + \sum</em>{j=1}^{m} \text{CRW}<em>j \times \text{Off}</em>\text{SE}_j \times \text{CCF}_j$</td>
<td></td>
</tr>
</tbody>
</table>

Legend: CRWA(CSE) – Credit Risk-Weighted Consolidated Sector Exposures; CRW – Credit Risk Weight; On_SE – On-Balance Sheet Sector Exposure; Off_SE – Off-Balance Sheet Sector Exposure; CCF – Credit Conversion Factor.

Note: The equation has been simplified to provide a more intuitive expression of the use of risk weights and credit conversion factors for the measurement of risk-sensitive sector exposures. A more detailed calculation might also recognise the impact of credit risk mitigants as defined in the BCBS’s Basel Accords, the role of specific provisions, and other relevant factors.

**Revised Risk Sensitivity of Real Estate Exposures**

Exposures to residential or commercial real estate with lower LTV ratios have different risk characteristics than exposures with higher LTV ratios. This is based on the equity-maximisation theory that links borrowers’ defaults with rational comparisons between costs and returns given contemporaneous LTV ratios. In contrast, the ability-to-pay theory relates borrowers’ defaults primarily to cash-flow constraints. The evidence from several studies supports the equity-maximisation model in lieu of the ability-to-pay model (Jackson et al., 1980; Wong et al., 2004).

---

5 Real estate companies can be referred to as corporate entities with a majority share of total turnover derived from real estate activities or financing of real estate activities.

6 Off-balance sheet exposures commonly include guarantees, letters of credit, acceptances, commitments, and other types of off-balance sheet exposures of financial institutions.
The revised standardised approach for credit risk (part of finalising post-crisis Basel III reforms) enhances risk-weight granularity subject to varying LTV ratios of residential and commercial properties. It also provides risk-weight granularity subject to risk weights of counterparties, source of repayment cash flows, or issue-specific ratings of covered bonds. The revision enables a greater integration between the measurement of sector exposures and the harmonised Basel Accords' risk sensitivity (along with the use of harmonised credit conversion factors or credit risk mitigants).  

### Risk Weights for Residential Real Estate Exposures

Table (1)

<table>
<thead>
<tr>
<th>Repayment is not materially dependent on the cash flow generated by the property</th>
<th>LTV ratio</th>
<th>≤50%</th>
<th>50% to 60%</th>
<th>60% to 80%</th>
<th>80% to 90%</th>
<th>90% to 100%</th>
<th>&gt;100%</th>
</tr>
</thead>
<tbody>
<tr>
<td>Risk weight</td>
<td>20%</td>
<td>25%</td>
<td>30%</td>
<td>40%</td>
<td>50%</td>
<td>70%</td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Repayment is materially dependent on the cash flow generated by the property</th>
<th>LTV ratio</th>
<th>≤50%</th>
<th>50% to 60%</th>
<th>60% to 80%</th>
<th>80% to 90%</th>
<th>90% to 100%</th>
<th>&gt;100%</th>
</tr>
</thead>
<tbody>
<tr>
<td>Risk weight</td>
<td>30%</td>
<td>35%</td>
<td>45%</td>
<td>60%</td>
<td>75%</td>
<td>105%</td>
<td></td>
</tr>
</tbody>
</table>

Source: BCBS (2017)

### Risk Weights for Commercial Real Estate Exposures

Table (2)

<table>
<thead>
<tr>
<th>Repayment is not materially dependent on the cash flow generated by the property</th>
<th>LTV ratio</th>
<th>≤60%</th>
<th>&gt;60%</th>
</tr>
</thead>
<tbody>
<tr>
<td>Risk weight</td>
<td>Min (60%, Risk weight of counterparty)</td>
<td>Risk weight of counterparty</td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Repayment is materially dependent on the cash flow generated by the property</th>
<th>LTV ratio</th>
<th>≤60%</th>
<th>60% to 80%</th>
<th>&gt;80%</th>
</tr>
</thead>
<tbody>
<tr>
<td>Risk weight</td>
<td>70%</td>
<td>90%</td>
<td>110%</td>
<td></td>
</tr>
</tbody>
</table>

Source: BCBS (2017)

### Risk Weights for Covered Bond Exposures

Table (3)

<table>
<thead>
<tr>
<th>Rated covered bond exposures</th>
<th>Issue-specific rating</th>
<th>AAA to AA-</th>
<th>A+ to A-</th>
<th>BBB+ to BBB-</th>
<th>BB+ to B-</th>
<th>Below B-</th>
</tr>
</thead>
<tbody>
<tr>
<td>Risk weight</td>
<td>10%</td>
<td>20%</td>
<td>20%</td>
<td>50%</td>
<td>100%</td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Unrated covered bond exposures</th>
<th>Risk weight of issuer</th>
<th>20%</th>
<th>30%</th>
<th>40%</th>
<th>50%</th>
<th>75%</th>
<th>100%</th>
<th>150%</th>
</tr>
</thead>
<tbody>
<tr>
<td>Risk weight</td>
<td>10%</td>
<td>15%</td>
<td>20%</td>
<td>25%</td>
<td>35%</td>
<td>50%</td>
<td>100%</td>
<td></td>
</tr>
</tbody>
</table>

Source: BCBS (2017)

---

7 The Basel Accords' internal ratings-based (IRB) approach relies upon banks' internal assessments of credit risk weights with a primary objective of further enhancing risk sensitivity (BCBS, 2001). Similar to the revised standardised approach, risk weights modelled based on the IRB approach decline with lower LTV ratios (PRA, 2017). Both revised standardised and IRB approaches can be utilised within risk-sensitive exposure frameworks subject to the compilers' objectives. A post-crisis Basel III reform of the IRB approach introduced an output floor of 72.5% of the total risk-weighted assets, calculated based on the standardised approach (BCBS, 2017), designed to limit inconsistencies in RWAs between banks.
Risk-Weighted Sector Concentration Ratio

Sector concentrations in banks’ portfolios represent a key driver of credit risk (Bonti, 2006). Sector concentration risk has been one of the main causes of banks’ failures (BCBS, 2004). Based on empirical studies, sector concentrations increase banks’ credit risk and, in turn, an optimal level of economic capital by 20% to 40% (BCBS, 2006). Therefore, adequate measurement of sector concentrations remains crucial for risk management and banking supervision. There has been a remarkable diversity in the way different banks and supervisors approach sector concentrations.

Similar-sized sector exposures of different banks or banking systems may have significantly different risk characteristics, rendering concentration risk statistics that are not risk-sensitive inadequate. Varying scopes and lack of risk sensitivity restrict cross-bank or cross-country comparability of sector concentration ratios and their interpretability by stakeholders. A sector concentration ratio defined through credit risk-weighted assets benefits from the Basel Accord’s harmonised methods for risk sensitivity (and also, credit conversion factors, risk mitigants, etc.).

<table>
<thead>
<tr>
<th>Sector Concentration Ratio</th>
<th>Equation (3)</th>
</tr>
</thead>
<tbody>
<tr>
<td>[ \frac{\text{CRWA(CSE)}}{\text{CRWA}} ]</td>
<td></td>
</tr>
</tbody>
</table>

Legend: CRWA(CSE) – Credit Risk-Weighted Consolidated Sector Exposures; CRWA – Total Credit Risk-Weighted Assets.

Conclusion

The concentrations of banks’ exposures in sectors or segments can substantially increase the credit risk of banks’ portfolios. This warrants a greater focus on developing holistic and harmonised approaches for measurement, management, and supervision of sector exposures. Such comprehensive measurement shall include all types of on-balance sheet and off-balance sheet exposures to a sector. Off-balance sheet exposures can be converted into credit-exposures equivalents using harmonised credit conversion factors developed as part of the Basel Accords.

In addition to a holistic and harmonised scope, sector exposure measures need to be adequately risk-sensitive. The BCBS’s Basel Accords provide the only globally harmonised methodology for risk sensitivity. The study outlines a method for integrating sector exposure statistics with the revised BCBS’s Basel Accords. Banks’ sector exposures and concentrations can thus be expressed in terms of credit risk-weighted assets. Defining exposures in terms of credit risk-weighted assets also provides a more comparable and intuitive approach related to credit risk and capital adequacy.
Reference:


Banks' Real Estate Exposures: Risk-Based Approach to Measurement of Exposures and Concentrations

Patrick Slovik and Farah Azman

Abstract

The study develops a holistic risk-based approach to measuring banks' real-estate exposures and real-estate sector concentrations. As banks' real-estate exposures increasingly extend beyond traditional lending, adequate statistics shall cover broader types of on-balance sheet exposures and off-balance sheet exposures. The study describes a novel approach to measuring banks' exposures and concentrations utilising a risk-based approach aligned with the more granular post-crisis reforms of the Basel standards. It outlines a method of integrating the revised Basel standards with statistics on banks' sectoral exposures and concentrations and proposes refined metrics based on credit risk-weighted assets.
Banks' Credit Risk and Sector Concentrations

Sector Exposures and Concentration Risk

- Historical experience shows that sector concentration risk in banks' portfolios has been one of the main causes of bank distress, which warrants a focus on appropriate statistical methodologies.

- Sector concentrations in banks' portfolios arise from excessive exposures to a sector, several highly correlated sectors, and also apply to other exposures exhibiting high default dependencies.

Sector Exposures and Financial Distress

- Banks' credit risk concentrations, particularly real estate sector exposures, exert material impact on the soundness of the financial system and contributed to numerous financial or bank crises globally.

- Real-estate-linked financial crises were not limited to a particular real estate category, as the causes stem from exposures to both residential real estate and commercial real estate rather equitably.

Sector Exposures and Financial Innovation

- While banks' real estate financing has been traditionally dominated by loans, other types of on-balance sheet and off-balance sheet exposures to the real estate sector increased in prominence.

- This is attributable to broadening financial innovation and deepening market-based finance. In response, authorities should adjust statistics to accommodate changes in financial structures.
Adequate measurement and analysis of banks' real estate exposures require extending their scope to include all types of on-balance sheet and off-balance sheet exposures to the real estate sector.

Measurement and disclosure of off-balance sheet sector exposures are commonly missing even in relatively advanced contemporary frameworks and dashboards, resulting in material data gaps.

### Consolidated Sector Exposures

\[ CSE = \sum_{i=1}^{n} On_{SE_i} + \sum_{j=1}^{m} Off_{SE_j} \times CCF_j \]

**Legend:**
- **CSE** – Consolidated Sector Exposures
- **On SE** – On-Balance Sheet Sector Exposure
- **Off SE** – Off-Balance Sheet Sector Exposure
- **CCF** – Credit Conversion Factor

Off-balance sheet items can be converted into credit-exposure equivalents through the use of harmonised credit conversion factors that were developed as part of the BCBS’s Basel Accords.
While the consolidated sector approach defines an all-inclusive scope, it might not be sufficiently risk-sensitive. Risk sensitivity recognises that credit risk of same-sized credit exposures differs.

The credit risk of real estate exposures relates to factors such as loan-to-value ratios, counterparties' external credit ratings, source of repayment cash flows, exposures' asset quality, or others.

Credit Risk-Weighted Consolidated Sector Exposures

\[
CRWA(CSE) = \sum_{i=1}^{n} CRW_i \times On\_SE_i + \sum_{j=1}^{m} CRW_j \times Off\_SE_j \times CCF_j
\]

Legend: CRWA(CSE) – Credit Risk-Weighted Consolidated Sector Exposures; CRW – Credit Risk Weight; On_SE – On-Balance Sheet Sector Exposure; Off_SE – Off-Balance Sheet Sector Exposure; CCF – Credit Conversion Factor.

Note: The equation has been simplified to provide a more intuitive expression of the use of risk weights and credit conversion factors for the measurement of risk-sensitive sector exposures. A more detailed calculation might also recognise the impact of credit risk mitigants as defined in the BCBS's Basel Accords, the role of specific provisions, and other relevant factors.

Credit risk weights derived from the harmonised BCBS's Basel Accords can be utilised to obtain credit risk-weighted consolidated sector exposures.
The revised standardised approach for credit risk (part of finalising post-crisis Basel III reforms) enhances risk-weight granularity of bank real estate exposures to residential and commercial properties.

Revised risk-weight granularity varies based on loan-to-value ratios, risk weights of counterparties, sources of repayment cash flows, issue-specific ratings of covered bonds, and other relevant factors.

### Risk Weights for Residential Real Estate Exposures

<table>
<thead>
<tr>
<th>Repayment is not materially dependent on the cash flow generated by the property</th>
<th>Table (1)</th>
</tr>
</thead>
<tbody>
<tr>
<td>LTV ratio</td>
<td>≤50%</td>
</tr>
<tr>
<td>Risk weight</td>
<td>20%</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Repayment is materially dependent on the cash flow generated by the property</th>
</tr>
</thead>
<tbody>
<tr>
<td>LTV ratio</td>
</tr>
<tr>
<td>Risk weight</td>
</tr>
</tbody>
</table>

Source: BCBS (2017)

The revised standardised approach enables a greater integration between the measurement of sector exposures and the harmonised Basel Accords’ risk sensitivity (also applicable in the IRB approach).
**Revised Risk Sensitivity of Real Estate Exposures (continued)**

### Risk Weights for Commercial Real Estate Exposures

<table>
<thead>
<tr>
<th>Repayment is not materially dependent on the cash flow generated by the property</th>
<th>Table (2)</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>LTV ratio</strong></td>
<td><strong>Risk weight</strong></td>
</tr>
<tr>
<td>≤60%</td>
<td>Min (60%, Risk weight of counterparty)</td>
</tr>
<tr>
<td>&gt;60%</td>
<td>Risk weight of counterparty</td>
</tr>
</tbody>
</table>

| Repayment is materially dependent on the cash flow generated by the property |
|---|---|
| **LTV ratio** | **Risk weight** |
| ≤60% | 70% |
| 60% to 80% | 90% |
| >80% | 110% |

Source: BCBS (2017)

### Risk Weights for Covered Bond Exposures

<table>
<thead>
<tr>
<th>Rated covered bond exposures</th>
<th>Table (3)</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Issue-specific rating</strong></td>
<td><strong>Risk weight</strong></td>
</tr>
<tr>
<td>AAA to AA-</td>
<td>10%</td>
</tr>
<tr>
<td>A+ to A-</td>
<td>20%</td>
</tr>
<tr>
<td>BBB+ to BBB-</td>
<td>20%</td>
</tr>
<tr>
<td>BB+ to B-</td>
<td>50%</td>
</tr>
<tr>
<td>Below B-</td>
<td>100%</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Unrated covered bond exposures</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Risk weight of issuer</strong></td>
</tr>
<tr>
<td>30%</td>
</tr>
<tr>
<td>40%</td>
</tr>
<tr>
<td>50%</td>
</tr>
<tr>
<td>75%</td>
</tr>
<tr>
<td>100%</td>
</tr>
<tr>
<td>150%</td>
</tr>
</tbody>
</table>

| **Risk weight** |
| 10% |
| 15% |
| 20% |
| 25% |
| 35% |
| 50% |
| 100% |

Source: BCBS (2017)
Risk-Weighted Sector Concentration Ratio

- Sector concentrations in banks' portfolios represent a key driver of credit risk. For this reason, adequate measurement of concentrations remains crucial for risk management and banking supervision.

- Similar-sized sector exposures of different banks or banking systems may have significantly different credit risk characteristics, rendering concentration risk statistics that are not risk-sensitive inadequate.

<table>
<thead>
<tr>
<th>Sector Concentration Ratio</th>
<th>Equation (3)</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\text{Sector Concentration Ratio} = \frac{\text{CRWA(CSE)}}{\text{CRWA}}$</td>
<td></td>
</tr>
</tbody>
</table>

**Legend:** CRWA(CSE) – Credit Risk-Weighted Consolidated Sector Exposures; CRWA – Total Credit Risk-Weighted Assets.

A sector concentration ratio defined through credit risk-weighted assets benefits from the Basel Accords' harmonised methods for risk sensitivity (and also, credit conversion factors, risk mitigants, etc.).
Extending Measurement Scope of Sector Exposures

- Holistic and harmonised approaches for measurement of sector exposures shall include all types of on-balance sheet and off-balance sheet exposures to a sector.

- Off-balance sheet exposures can be converted into credit-exposures equivalents using harmonised credit conversion factors developed as part of the Basel Accords.

Risk-Based Measurement of Sector Exposures

- Sector exposure and concentration measures need to be adequately risk-sensitive. The BCBS's Basel Accords provide the only globally harmonised methodology for risk sensitivity.

- Sector exposures and concentrations expressed in terms of credit risk-weighted assets provide a more comparable and intuitive approach related to credit risk and capital adequacy.
Examining concentration and similarity in institutional investors’ holdings

Ariel Mantzura and Michael Bell,
Bank of Israel

---

1 This presentation was prepared for the conference. The views expressed are those of the authors and do not necessarily reflect the views of the BIS, the IFC or the central banks and other institutions represented at the event.
Examining Concentration and Similarity in Institutional Investor's Holdings

An application Using Network Analysis

Ariel Mantzura, Bank of Israel

Examining Concentration and Similarity in Institutional Investor's Holdings

The stock market is considered to be a good example of a complex system, one that consists of many entities and agents, interacting in such a way that their collective market behaviour is beyond a simple combination of their individual behaviours. Such systems are also characterized by various dimensions of interactions, that can be measured and modelled using data that is becoming increasingly more and more available. A key challenge in modern finance is finding efficient ways of summarizing and visualizing the stock market data to obtain useful information about its behaviour. Over the past two decades methodologies have been introduced to address such challenges, including machine learning techniques, and network analysis methods that have become very popular. In this study, we test how interdependence both between the agents and their assets within the network of portfolio holdings can be a source of systemic vulnerability. We study a real-world, institutional investors holdings network based on a granular dataset of all common assets (stocks and bonds) holdings in Israel from 2010-2021 and compare it with various alternative scenarios from randomization and rebalancing of the original investments. The scenarios generation relies on algorithms that satisfy the global constraints imposed by the numbers of outstanding shares in the market and the regulatory constraints. We extensively analyse the interplay between portfolio diversification and differentiation and examine how the outreach of exogenous shocks depends on these factors as well as on the type of shock and the size of the network with respect to the market. We find that real portfolios are diversified but highly similar, that portfolio similarity correlates with systemic fragility and that rebalancing can come with an increased similarity depending on the initial network configuration.

Keywords: Network analysis, cosine similarity, concentration

JEL classification: D85, G01, G18
Contents

Introduction ......................................................................................................................... 1
Data .................................................................................................................................... 3
Measures of Concentration and Similarity ........................................................................ 4
  Herfindahl Hirschman Index .............................................................................................. 4
  Entropy .............................................................................................................................. 5
  Cosine similarity ............................................................................................................... 8
Network analysis of institutional investor portfolios ....................................................... 10
1. Introduction

Finding efficient ways of summarizing and visualizing the stock market data to extract useful information about its behavior as well as finding interconnectedness patterns among market participants is crucial for identification of systemic vulnerabilities in the financial system and particularly, vulnerabilities arising from commonalities in asset holdings (overlapping portfolios). Using Network analysis for the Modelling of financial systems has become popular in recent years (Poledna et.al, 2020).

The aim of this study is to empirically analyse systemic vulnerabilities due to commonalities in asset holdings in the Israeli financial system across different types of financial institutions: Insurance companies, Mutual funds, Provident funds and Pension funds. The Institutional investors in Israel are required to publish on a quarterly basis their detailed asset holdings in accordance with the Capital Market Authority, Insurance and Savings reporting directives. Similarly, Mutual fund managers are required to publish on a monthly basis their detailed asset holdings in accordance with the Israeli Securities Authorities reporting directives. Using such required disclosures can allow for a study of commonalities across such investors.

For this study we use Praedicta’s Granular Data base of Management company asset holdings (equity and bonds) on a quarterly basis. Specifically, we test how interdependence both between the agents and their assets within the networks of portfolio holdings can be a source of systemic vulnerability. We study a real-world, institutional investors holdings network based on a granular dataset of all common assets (equity and bonds) holdings in Israel from 2010-2021.

Systemic vulnerabilities in financial markets may lead to the realization of systemic risks where exogenous shocks, failure or distress in one or several institutions may affect other financial institutions through some contagion channel. There are several contagion channels referred to in the literature. Poledna (2020) distinguishes between two types of contagion channels: direct and indirect contagion. Direct contagion (see Upper and Worms (2002)) is due to bilateral exposures of two institutions, i.e., through borrowing and lending and is quantified in the institutions balance sheet liabilities. Alternatively, indirect contagion can arise when financial institutions invest in the same assets where this type of contagion is referred to as overlapping portfolios (Levy-Carciente at al., 2015). In this case contagion has the following mechanism: As a result of some exogenous shock to some financial institution or to one or some of its assets it is forced to liquidate some of its securities at lower prices due to binding constraints i.e., contractual or regulatory constraints. Liquidation of securities at discounted prices referred to as fire sales propagates through the financial system hence causing losses to other financial institutions holdings those assets. Girardi et al. (2018) studied empirically this notion of indirect contagion in the insurance industry using the cosine similarity measure. Huang et al. (2013), Caccioli (2015) model overlapping portfolios as a bipartite network (banks and assets) of the banking system to describe the risk propagation process during crises and find that

1 http://praedicta.com
their model can be useful for systemic risk stress testing for the banking system. Elsinger et al. (2006) examined systemic risk in the Austrian banking system and found that correlations in banks asset portfolios dominates contagion as the main source of systemic risk. Barucca et al. (2021) examined interconnections and systemic vulnerabilities related to price-mediated contagion (overlapping portfolios) across multiple types of institutions.

The paper is organized as follows: In section 2 we describe to more detail the data we use in our investigation and present in section 3 some descriptive statistics and various measures of concentration and similarity in asset holdings. In section 4 we model the financial system as a network and extract important measures.

2. Data

We use in this work Praedicta’s granular Data base. The data base includes the total asset holdings of management companies (including institutional investors) from 2010Q1 up to the present\(^2\). The management companies are one of four: Insurance companies, pension funds, provident funds and mutual funds. The data is quarterly and detailed up to a particular asset and investment channel (funds). For each company and fund the total holding in assets is given. In Table 1 is an example of typical records of the data base.

Table 1: Example of Typical Records in Database

<table>
<thead>
<tr>
<th>Channel Name</th>
<th>Asset type</th>
<th>Asset name</th>
<th>Market Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Fund A investment channel for age 60+</td>
<td>Corporate bond</td>
<td>Poalim bond 29</td>
<td>279,855</td>
</tr>
<tr>
<td>Fund A general investment channel</td>
<td>Equity TA 35</td>
<td>Elbit Systems</td>
<td>25,955,137</td>
</tr>
<tr>
<td>fund B</td>
<td>Corporate bond</td>
<td>Poalim bond 29</td>
<td>32,811</td>
</tr>
<tr>
<td>fund B</td>
<td>Equity TA 35</td>
<td>Elbit Systems</td>
<td>856,926</td>
</tr>
<tr>
<td>fund C</td>
<td>Corporate bond</td>
<td>Poalim bond 29</td>
<td>8,552,713</td>
</tr>
<tr>
<td>fund C</td>
<td>Equity TA 35</td>
<td>Elbit Systems</td>
<td>292,002,936</td>
</tr>
</tbody>
</table>

\(^2\) The analysis in this work used data till 2020.Q1.
3. Measures of Concentration and Similarity

Concentration Measures

We examine concentration segmented by Management company type, namely, Insurance companies, Pension funds, Mutual funds and Provident funds. We measure concentration in 3 different dimensions:

1. Concentration of Management companies
2. Concentration of Funds
3. Concentration of Assets

We aim to answer the following question: What is the effective number of companies, funds, and assets? The effective number differs from simple counting in the sense that effective number accounts also for size. To this end, we use the well-known indices: The Herfindhal- Hirschman Index (HHI) (Herfindhal, 1959; Hirschman 1945) and Entropy. These measures are calculated for each quarter and therefore we can examine their dynamics as a quarterly time series.

Herfindhal-Hirschman (HHI)

Assume for example that the aggregate portfolio of the management companies is invested in \( k \) assets with asset allocation represented by the following vector:

\[
v = (v_1, \ldots, v_k)
\]

Where \( v_i > 0 \) is the total holding in asset \( i \). Let,

\[
p = (p_1, \ldots, p_k)
\]

be the vector of proportions of asset holdings where \( p_i = \frac{v_i}{\sum_{j=1}^{k} v_j} \). The HHI is defined by the squared norm of these proportions:

\[
HHI = ||p||^2 = \sum_{j=1}^{k} p_i^2
\]

Entropy

The second measure we use is based on Entropy and is defined as follows:

\[
EFF = 2^{H(p)}
\]

Where,

\[
H(p) = - \sum_{j=1}^{k} p_i \log_2(p_i)
\]
The Entropy is a common measure of diversity that gets values between 0 which points at full concentration and between $\log_2(k)$ which reflects full diversity in the case where $p_1 = p_2 = \cdots = p_k = 1/k$.

These 2 measures are calculated in the same manner for companies, assets, funds and branches. For instance, when examining concentration of management companies $k$ will be the number of management companies and $v_1, \ldots, v_k$ will be the vector of total assets that each company holds. In figures 1-4 we display concentration measures calculated in Insurance Companies.

**Figure 1 - Insurance Companies, Total vs Effective Numbers**

From figure 1 we can observe that the effective number of managers, assets and funds is significantly smaller (and smoother) than the total number. This indicates of higher concentration i.e. high percent of capital is concentrated in fewer managers, assets and funds, than the ones indicated by the total number of Managers, assets and funds, respectively. We can also observe a slight improvement across the years in diversity of the portfolio as indicated in the effective number of assets whereas there is no change in the effective number of managers. Moreover, there is a small effective number of Management companies which indicates of high concentration in the insurance industry.
In the pension funds as can be observed in figure 2 there is also a slight improvement in the diversification of the asset portfolio. Also, in the pension field there is a small effective number of companies indicating high concentration. A structural break in 2017 is apparent in the effective number of funds. This is due to the change in reporting directives which required to report on specific investment channels rather than general management companies. This is apparent in the HHI index which shows the divergence between managers and funds beginning from 2017.
From figure 3 we can also observe a small number of managers and although there is an indication of worsening in the total number of companies the effective number indicates of stability in the number of managers. Also, the improvement in the number of assets is slighter than the one indicated by the total of assets.

Figure 4 - Mutual Funds, Total vs Effective Numbers
Similarity Measures

We examine similarity between portfolios of different management companies using Cosine similarity. To this end, we constructed for each management company the total holdings in each asset aggregated over the different funds. We then calculated between every 2 companies the cosine similarity measure. Cosine similarity assumes values between -1, which reflects absolute lack of similarity and 1 which reflects full similarity. In each quarter we represent these measures in a matrix that in cell i,j is the cosine similarity between company j and company i. These matrices can be constructed as heat maps where similar companies are close in the matrix. Following is a more formal definition of cosine similarity.

Cosine Similarity

Assume 2 management companies with the following vectors of total asset holdings in assets 1,...,k.

\[ w = (w_1, \ldots, w_k) \]
\[ v = (v_1, \ldots, v_k) \]

Where \( w_i, u_i > 0 \) the total asset holdings of each is company in asset i. The cosine similarity measure is defined between these two vectors is defined as:

\[ \cos(v, w) = \frac{\sum_{j=1}^{k} w_i v_i}{\sqrt{\sum_{j=1}^{k} v_i^2} \sqrt{\sum_{j=1}^{k} w_i^2}} \]

Theoretically this measure can assume negative values but in our case, we only have non-negative asset holdings so this measures assumes only non-negative values.

Figure illustrates a heat map of portfolio similarity calculated with the cosine similarity measure for December 2014 in pension funds

Figure 5 – Heat Map of Cosine similarity between effective managers of Insurance Companies, 2019.Q1
Figure 6 – Heat Map of Cosine similarity between effective managers of Provident funds, 2019.Q1

Figure 7- Heat Map of Cosine similarity between effective managers of Pension Funds, 2019.Q1
The matrix of cosine similarity describes the portfolio similarity between each pair of companies.

4. Network analysis of institutional investor portfolios

A fully connected undirected weighted graph \( G = (V, E) \) where \( E = V^2 \) can be induced from this matrix. In this graph each node represents a management company and edges between every pair of nodes is the cosine similarity between any two companies. We would like to filter this graph in a way that preserves the most relevant part of the data interdependency patterns measured by cosine similarity. We use for this purpose a method called Maximal Planar Graphs.

The Maximal planar graph is a filtering method of a dense graph or network. The main idea under these methods (see Massara et. al (2015)) is to filter a dense matrix of weights (correlations or cosine similarities) by retaining the largest and most significant sub graph while imposing constraints on the structure of the resulting graph. The minimal (or maximal) spanning tree retains the edges with the largest weights (correlations) while constraining the sub graph to be a spanning tree. The Maximal planar graph does the same while constraining the sub graph to be a planar graph (no edge crossing).
We can observe from figure 10 that there is high similarity between management companies of the same type. The network exhibits clusters among provident funds and Mutual funds and also among pension funds. Insurance companies do not seem to cluster.

The density of a network is a measure of a graph is the ratio between the number of edges present in the graph and the maximum number of edges that the graph can contain. The density may reflect high aggregate similarity between management companies.

**Figure 11- Density graphs by Management Company type**

We can observe from figure 11 the structural break in the density around 2016 due to the inclusion of ETF’s in addition to ETN’s and hence weakening the density since 2016. There is an increase in the density among provident funds since 2017 due to the change in reporting directives as mentioned earlier.
References


Quantification of systemic risk from overlapping portfolios in the financial system
S Poledna, S Martínez-Jaramillo, F Caccioli, S Thurner Journal of Financial Stability 52, 100808

Examining Concentration and Similarity in Institutional Investor's Holdings

Ariel Mantzura, Michael Bell
Bank of Israel
Data

• We use in this work Praedicta’s granular Data base.

• The data base includes the total asset holdings of management companies (including institutional investors) from 2010Q1 up to the present.

• The management companies are one of four: Insurance companies, pension funds, provident funds and mutual funds.

• The data is quarterly and detailed up to a particular asset and investment channel (funds).

• For each company and fund the total holding in assets is given.
Concentration: Questions?

• How many Management companies?

• How many effectively?

• How similar are the investment portfolios of the companies?

• How many assets? How many effectively?
Concentration Measures

- Assume that the aggregate portfolio is invested in k assets with asset allocation represented by the following vector:
  \[ v = (v_1, \ldots, v_k) \]
- Where \( v_i > 0 \) is the total holding in asset i.
- Let,
  \[ p = (p_1, \ldots, p_k) \]
  be the vector of proportions of asset holdings where \( p_i = \frac{v_i}{\sum_{j=1}^{k} v_i} \).
- The HHI is defined by the squared norm of these proportions:
  \[ HHI = ||p||^2 = \sum_{\{j=1\}}^{k} p_i^2 \]
- Entropy
- The second measure we use is based on Entropy and is defined as follows:
  \[ EFF = 2^{H(p)} \]
- Where,
  \[ H(p) = -\sum_{\{j=1\}}^{k} p_i \log_2(p_i) \]
Pension Funds, Total vs Effective Numbers
Portfolio Similarity – Cosine Similarity

• Assume 2 management companies with the following vectors of total asset holdings in assets 1,…,k.

\[ w = (w_1, \ldots, w_k) \]
\[ v = (v_1, \ldots, v_k) \]

• Where \( w_i, u_i > 0 \) the total asset holdings of each is company in asset i.

• The cosine similarity measure is defined between these two vectors is defined as:

\[
\cos(v, w) = \frac{\sum_{j=1}^{k} w_i v_i}{\sqrt{\sum_{j=1}^{k} v_i} \sqrt{\sum_{j=1}^{k} w_i}}
\]
Heat Map of Cosine similarity between effective managers of Insurance Companies, 2019.Q1
Filtering with maximal planar

- The Maximal planar graph is a method of filtering dense matrix of weights (correlations or cosine similarities) by retaining the largest and most significant subgraph while imposing constraints on the structure of the resulting graph.

- The Maximal planar graph does the same while constraining the subgraph to be a planar graph (no edge crossing).
Triangulated Maximally Filtered Graph - all types - 2014-12-31
To Conclude

- The Effective Number is significantly smaller than the total number.
- Significant portion of effective companies are highly similar.
- Similarity is often present in same company type.
Statistical matching for anomaly detection in insurance assets granular reporting

Vittoria La Serra and Emiliano Svezia, Bank of Italy

---

1 This presentation was prepared for the conference. The views expressed are those of the authors and do not necessarily reflect the views of the BIS, the IFC or the central banks and other institutions represented at the event.

2 Vittoria La Serra and Emiliano Svezia (Bank of Italy) received the IFC award for the best paper presented by a young statistician, "Statistical matching for anomaly detection in insurance assets granular reporting" from the IFC Chair Rashad Cassim.
Statistical matching for anomaly detection in insurance assets granular reporting

Vittoria La Serra, Emiliano Svezia

Abstract

Since 2016, insurance corporations report granular asset data in Solvency II templates on a quarterly basis. Assets are uniquely identified by codes that are required to be kept stable and consistent over time; nevertheless, due to reporting errors, unexpected changes in the codes may occur, causing inconsistencies when compiling insurance statistics. The paper addresses this issue as a statistical matching problem and a supervised classification approach is proposed to detect such anomalies. Test results show the potential benefits of machine learning techniques on data quality management processes and the efficiency gains arising from automation, especially during situations of constraints on human resources, as the ongoing pandemic.

Keywords: insurance data, data quality management, record linkage, statistical matching, machine learning.

JEL classification: C18, C81, G22

Contents

Statistical matching for anomaly detection in insurance assets granular reporting

Introduction and Motivation ............................................................................................................... 2

1. Data description .................................................................................................................................. 3

2. The proposed approach ...................................................................................................................... 4

2.1 A record linkage problem ................................................................................................................. 4

2.2 Model selection ................................................................................................................................. 6

2.3 The results ........................................................................................................................................ 8

3. Conclusions and further developments ............................................................................................. 10

References ............................................................................................................................................... 11

1 Bank of Italy, Statistical Data Collection and Processing Directorate. The views expressed herein are those of the authors and do not necessarily reflect those of the Bank of Italy.
Introduction and motivation

In the process of collecting, processing and disseminating statistics, an effective and efficient data quality management (DQM) is of paramount importance in order to ensure the high quality of data. The automation of DQM processes became crucial in presence of increasingly granular databases. Furthermore, since the beginning of the Covid-19 pandemic in 2020, it has become clear the importance of investing resources in making such processes as much automatic as possible, in order to enhance their resilience in presence of situations of human resource constraints.

In the statistical literature, machine learning models are emerging as important tools to approach the DQM on very granular data in an automated way, since they generally outperform traditional modelling approaches in prediction tasks (Chakraborty et al., 2017). Restricting the issue to central banks statistics, the Bank of Italy has already applied successfully several machine learning methods to specific DQM processes (see Buzzi et al., 2020, Cusano et al., 2021, Zambuto et al., 2021, Maddaloni et al., 2022) and further research in this field is ongoing.

This paper proposes a machine learning approach in a statistical matching framework to solve - in an accurate and efficient automated way - a DQM issue on insurance granular assets data, specifically to check for anomalies in identification codes (ID) reporting. More in detail, the assets' IDs are expected to remain unique and consistent over time, meaning that the IDs assigned by the insurance corporations (ICs) cannot be subject to changes throughout the reporting history of the assets. However, from quarter to quarter, unexpected changes in the code for the same asset can occur. This is either due to annual updates of the requirements, which imply a change for assets’ codes, or, more often, it is a consequence of reporting errors. Therefore, an insurance corporation might either consciously revise an asset’s code, following requirements from an updated version of the regulation, or erroneously change it to a new, different code than the one used in the previous reporting quarter. Either ways, such changes have important consequences on the work of supervisory authorities and central banks, since they signal a change in the reported assets that in practice has not occurred; this raises DQM issues when analyzing assets’ time series and compiling the IC statistics that are then disseminated.

The paper is organized as follows. Section 1 describes the data from which the dataset used in the analysis is derived, presenting its structure and details on the Italian case. In Section 2, a record linkage approach based on machine learning models for classification is proposed; different models are tested on an Italian dataset and a robust and high-performance random forest approach is chosen, for whom results are presented. Section 3 summarizes the main conclusions, showing the advantages of the proposed approach and opening new ground for future research.
1. Data description

Since 2016 European insurance corporations (ICs) report to their national supervisory
authority, according to the Implementing Technical Standards (ITS)\(^2\) drawn by EIOPA,
and to the national central banks, quarterly data on their individual balance sheets.
The data is organised in templates according to the Solvency II Directive\(^3\). They
provide very granular and highly valuable information especially with template
S.06.02 which contains asset-by-asset information on the single holdings of insurance
corporations, showing the investments in debt securities, equity and investment fund
shares, as well as loans, deposits and properties.

Template S.06.02 allows, on the one hand, supervisory authorities to perform a
comprehensive and detailed risk assessment upon insurance undertakings and, on
the other, central banks to compile statistics on insurance sector, useful to analyse its
interconnections within the financial system and to gather knowledge on households’
wealth and income from insurance policies. This template is used both for supervisory
and statistical purposes, enriched with specific for the latter.

More in details, template S.06.02 comprises quantitative information on each position
held, such as the market and nominal value, quantity and accrued interest of the asset,
along with qualitative features, which include – wherever applicable – the type of
insurance undertaking\(^4\), the type of asset\(^5\), the issuer and/or counterparty sector, the
issuer and/or counterparty area, the currency, the issue and maturity dates, the name
of the issuer, the description of the asset.

Each asset in the template is reported with an identification code. Asset codes are
standardized in most cases (e.g. ISIN codes for securities), although in some cases
insurance corporations can assign internal codes and report them (CAU, Code
Attributed by the Undertaking).

The dataset used in the paper consists of Italian data from the S.06.02 template and
it is also integrated with attributes from the European Central Bank’s Centralised
Securities Database (CSDB) - the European harmonised security registry. In detail, the
population of Italian ICs is composed of around 100 entities; overall, reported data
comprises almost 30 reporting quarters and around 70,000 assets at each period.

On average, in each quarter, there is a turnover of 8% in number and 4% in market
value share for the reported assets\(^6\). In line with the reporting instructions, asset ID
codes that are only reported in one of two adjacent quarters should consist in new

\(^2\) Commission Implementing Regulation (EU) 2015/2450 of 2 December 2015 and following amendments,
laying down implementing technical standards with regard to the templates for the submission of
information to the supervisory authorities, according to Directive 2009/138/EC of the European

up and pursuit of the business of Insurance and Reinsurance.

\(^4\) Life, non-life, composite and reinsurance.

\(^5\) As from now CIC (Category Identification Code).

\(^6\) Between two adjacent quarters, the turnover of the assets is the percentage of brand new reported codes
(not reported in the first quarter) over the total reported assets in the second quarter.
purchased or sold assets. However, these also include the cases of changes in the codes, whose exact percentage in the data is unknown. Therefore, 8% can be taken as the maximum expected percentage of cases of anomaly, which highlights how, even having a limited impact on the general data quality, errors in ID codes cannot be neglected.

2. The proposed approach

2.1 A record linkage problem

Statistical matching techniques, as described in D’Orazio et al. (2006), have the objective to draw information from two (or more) different datasets by linking them with respect to some common observed variables. Such techniques were originally proposed with the aim of data integration, i.e. to link two (or more) datasets coming from independent surveys and build a richer dataset containing information from both (Okner, 1972).

A specific case of statistical matching is record linkage, which is applied when the statistical units in two datasets are supposed to be at least partially overlapping (D’Orazio et al., 2006) and the objective of the analysis is to identify the list of common units between the two.

The topic was first introduced and formalized by Fellegi et al. (1969) and a classical approach was then proposed by Jaro (1989). Since mid-Nineties, many applications of record linkage have concerned the issue of linking historical census data, as in the works of Ferrie (1996), Rosenwaike et al. (1998) and Ruggle (2002). Different methodologies for performing record linkage have later been proposed, such as mixture models (Larsen et al., 2001) and Bayesian approaches (Fortini et al., 2001; Tancredi et al., 2011).

More recent contributions to this topic make use of machine learning techniques (Feigenbaum, 2016, Rijpma et al., 2020), which is the framework of the current work. In fact, the issue of unexpected changes of ID codes in insurance data introduced above can be approached as a record linkage problem.

As described in Section 1, each reported asset in a quarter is identified by a unique ID code and comes with a set of reported features. If a change in an asset’s ID code occurs, so that an insurance corporation $I$ reports asset "$a$" in quarter $Q_t$ and recodes it as "$b$" in quarter $Q_{t+1}$, it is expected for the reported features of the two apparently different assets "$a$" and "$b$" to be the same", since they actually refer to the same asset. Comparing the reported features of the two assets is therefore necessary to assess whether their difference in ID code is in fact an anomaly, stemming from an unexpected change that has taken place.

7 The features selected in the dataset are “structural”. For this reason, they should remain stable except for limited statistical reclassifications that can occur.
As in a record linkage framework, two datasets of assets, each referring to two adjacent reporting quarters \(Q_t, Q_{t+1}\), can be compared to assess whether there are common units between the two datasets, where each unit is an asset, identified by its ID code and its reporting IC.

Assets in the two quarters are compared with respect to the observed features and such comparisons are carried out using distance measures, one for each feature’s type, whether categorical (nominal or ordinal), numerical or textual.

Nominal variables, such as reporting/counterparty sector or issuer/counterparty area, are compared with an overlap measure (Boriah et al., 2008), taking 0 as a measure of minimum distance if the reported values are equal and 1 otherwise; ordinal variables, such as the categorized maturity date, are compared via the Manhattan distance; lastly, textual variables as the assets’ description are compared using a Levenshtein measure for strings. All distance measures are normalized to take values in the interval \([0,1]\).

Each pair of assets, one reported in \(Q_t\) and one reported in \(Q_{t+1}\), can either be a “match” if the two assets share the same ID code or a “non-match” if they do not.

A comparison matrix is built, as reported in Table 1. Each row in the matrix refers to a pair of assets from the two adjacent quarters and each column refers to an observed feature, either nominal, ordinal, numerical or textual, for which a distance measure \(d(.)\) is chosen. The \(d_f(a,b)\) distance measure for two assets \(a\) and \(b\) on a feature \(f\) is a value calculated in \([0,1]\), where the endpoints respectively indicate minimum and maximum distance between the two observed values for feature \(f\).

The features in the comparison matrix are used as input (covariates) to supervised statistical models together with the status of each pair that is a binary target variable with values “match” or “non-match” to be predicted.

**Input to supervised models: the comparison matrix and the target variable**

<table>
<thead>
<tr>
<th>Asset codes</th>
<th>TARGET VARIABLE</th>
<th>Nominal</th>
<th>Ordinal</th>
<th>Numerical</th>
<th>Textual</th>
</tr>
</thead>
<tbody>
<tr>
<td>(Q_t)</td>
<td>(Q_{t+1})</td>
<td>Status</td>
<td>(i \in {1 \ldots n_i})</td>
<td>(j \in {1 \ldots n_j})</td>
<td>(k \in {1 \ldots n_k})</td>
</tr>
<tr>
<td>a</td>
<td>a</td>
<td>Match</td>
<td>(d_i^1(a,a) \ldots d_i^{n_i}(a,a))</td>
<td>(d_j^1(a,a) \ldots d_j^{n_j}(a,a))</td>
<td>(d_k^1(a,a) \ldots d_k^{n_k}(a,a))</td>
</tr>
<tr>
<td>a</td>
<td>b</td>
<td>Non-match</td>
<td>(d_i^1(a,b) \ldots d_i^{n_i}(a,b))</td>
<td>(d_j^1(a,b) \ldots d_j^{n_j}(a,b))</td>
<td>(d_k^1(a,b) \ldots d_k^{n_k}(a,b))</td>
</tr>
<tr>
<td>b</td>
<td>b</td>
<td>Match</td>
<td>(d_i^1(b,b) \ldots d_i^{n_i}(b,b))</td>
<td>(d_j^1(b,b) \ldots d_j^{n_j}(b,b))</td>
<td>(d_k^1(b,b) \ldots d_k^{n_k}(b,b))</td>
</tr>
<tr>
<td>...</td>
<td>...</td>
<td>...</td>
<td>...</td>
<td>...</td>
<td>...</td>
</tr>
</tbody>
</table>
For the application, two subsets are selected from the whole Italian database, including all reported assets from the two subsequent quarters 2021-Q1 and 2021-Q2, and the observed arrays of features on pairs of reported assets are compared by building a comparison matrix; with the goal of detecting the changes in ID codes that ICs have reported between the two quarters, comparison is only made for pairs referring to the same ICs. Even with this constraint, the number of rows in the matrix, i.e. the number of compared pairs, approaches 150 million units. Given the size of the dataset, it would be impossible for data analysts to manually check all pairs of assets.

The comparison matrix is built and afterwards split into a “training set” and a “test set”, respectively including 80% and 20% of the data. Moreover, both datasets are proportionally stratified with respect to the “asset type” feature, in order to obtain a representative dataset.

The two datasets are respectively used to train and test supervised classification models, to predict the status variable in the comparison matrix from the computed distances between features.

### 2.2 Model selection

Four widely used supervised classification models are considered: logit, bagging, random forests and neural networks.

The logit model is adopted as a benchmark, being a high-performing yet easy-to-interpret classifier (Feigenbaum, 2016). Among machine learning classification models, bagging (Breiman, 1996), random forests (Breiman, 2001) and neural networks (Bishop, 1995) are considered and respective hyperparameters are tuned for each model: number of bootstrap samples for bagging, numbers of trees and variables randomly sampled as candidates at each split ($mtry$) for random forest, number of nodes in the hidden layer for neural networks.

In order to assess robustness, the models are trained and tested multiple times on differently unbalanced data with respect to the binary target variable. More in detail, the models are trained and tested on comparison matrices having $p\%$ cases of match and $(1-p)\%$ cases of non-match, with $p$ ranging from 1% (extreme unbalance) to 50% (perfect balance).

The test results are presented averaging over the results obtained on test sets unbalanced with different $p$.

Average Receiver Operating Characteristics (ROC) curves are shown in Figure 1, built through the computation of average false positive and true positive rates, varying with the probability threshold for classification.

---

8 Third digit of the CIC code. Ten classes are considered.
Panels A and B both show the average ROC curves for the four models; panel A shows the whole curves, while panel B focuses on a smaller range, to better spot the differences in the models.

Area Under Curve (AUC) indexes for the average ROC curves are presented in Table 2; this index is the chosen criterion to assess each model’s best hyperparameters combination.

The results reported in Table 2 show that all four models perform well on the tested data, with a minimum AUC measure of 95.66% observed for the logit model and a maximum AUC of 99.64% observed for the random forest. As expected (Breiman, 2001), among the random-trees-based models, random forest outperforms bagging.

The superiority of the random forest model against the others can be clearly observed in Figure 1 - panel B, where its ROC curve always shows larger true positive rates over false positive rates, for all probability thresholds for classification, with respect to the other models. For instance, with a 1% false positive rate, the random forest can

---

9 The best hyperparameters combination for the three machine learning models are 100 bootstrap samples for bagging, 200 trees and 7 mtry for random forest, 30 neurons in the hidden layer for neural networks.
achieve a 99% true positive rate, while the neural network, bagging and logit models respectively achieve 98%, 95% and around 91% rates.

2.3 The results

The random forest model is selected among the tested ones, in relation of its superiority in the robustness analysis held in Subsection 2.2.

More results for the selected model are shown in this Section for an unbalance proportion $p$ of 5%, chosen for illustrative purpose, being smaller than 8%, namely the maximum expected unbalance proportion in the dataset (see Section 1).

Performance measures for the model are presented in Table 3, varying with the probability threshold for classification. The presented indexes are accuracy, balanced accuracy, true positive rate (TPR), true negative rate (TNR), false discovery rate (FDR) and the difference between true positive rate and false discovery rate.

<table>
<thead>
<tr>
<th>Probability threshold for classification</th>
<th>0.1</th>
<th>0.2</th>
<th>0.3</th>
<th>0.4</th>
<th>0.5</th>
<th>0.6</th>
<th>0.7</th>
<th>0.8</th>
<th>0.9</th>
<th>Average</th>
</tr>
</thead>
<tbody>
<tr>
<td>Accuracy</td>
<td>99.05</td>
<td>99.41</td>
<td>99.51</td>
<td>99.55</td>
<td>99.56</td>
<td>99.54</td>
<td>99.46</td>
<td>99.35</td>
<td>99.16</td>
<td>99.4</td>
</tr>
<tr>
<td>Balanced accuracy</td>
<td>99.27</td>
<td>98.99</td>
<td>98.71</td>
<td>98.36</td>
<td>97.98</td>
<td>97.52</td>
<td>95.91</td>
<td>94.38</td>
<td>92.27</td>
<td>97.04</td>
</tr>
<tr>
<td>TPR</td>
<td>99.52</td>
<td>98.53</td>
<td>97.82</td>
<td>97.05</td>
<td>96.23</td>
<td>95.27</td>
<td>91.96</td>
<td>88.86</td>
<td>84.62</td>
<td>94.43</td>
</tr>
<tr>
<td>FDR</td>
<td>15.72</td>
<td>9.52</td>
<td>7.3</td>
<td>5.87</td>
<td>5.02</td>
<td>4.43</td>
<td>2.82</td>
<td>2.07</td>
<td>1.53</td>
<td>6.03</td>
</tr>
<tr>
<td>TPR-FDR</td>
<td>83.79</td>
<td>89.02</td>
<td>90.53</td>
<td>91.17</td>
<td>91.21</td>
<td>90.83</td>
<td>89.13</td>
<td>86.78</td>
<td>83.09</td>
<td>88.39</td>
</tr>
</tbody>
</table>

General model performance measures such as accuracy and balanced accuracy show that the model correctly identifies true cases of match and true cases of non-match with high frequencies, for all probability thresholds. Mean values for the two measures are respectively 99.40% and 97.04%.

True negative rate (TNR) shows very good results, remaining stably around a 99% value for all thresholds. Making 95% of the unbalanced test set, the cases of non-match are naturally easier to be detected by the classification model.

With the aim of choosing the best probability threshold to consider in the model, true positive rates (TPR), false discovery rates (FDR) and the difference between the two are also presented, both in Table 3 and in Figure 2.
Benjamini et al. (2001) define FDR as the “expected proportion of false discoveries among the discoveries”. Given a binary confusion matrix, FDR is the percentage of negative cases that the model incorrectly classifies as positive cases; in the current analysis, FDR coincides with percentage of non-matching assets that are erroneously classified as matches by the model and it is therefore an interesting cost measure to minimize.

As observable in Table 3 and Figure 2 - panel A, both the true positive rate and false discovery rate slowly decrease with the probability threshold.

For instance, taking the lowest threshold for classification, the selected model ensures a 99.52% rate of correctly classified cases of match, with the consequence of a 15.72% incorrectly classified cases of non-match; instead, taking the highest threshold, less than 2% false discoveries are made but only 84.62% cases of match are correctly identified.

Therefore, a trade-off between the two indexes must be found in order to choose an appropriate probability threshold for the model. However, the two rates do not have the same weight in the current analysis: although false discovery rate is a cost measure to minimize, maximizing the true positive rate is considered as a priority for the model effectiveness. Detecting most of the true cases of match is in fact the goal of the analysis and it is therefore desirable to select a lower probability threshold for classification which would ensure to reach the goal, even if that implies that some cases of non-match are erroneously classified as matches.

To assess for the best threshold, the difference between TPR and FDR is calculated and reported in Table 3 and Figure 2 - panel B. The maximum value for the index is reached on a 0.5 threshold; however, the maximum increase in the index is observed when switching from threshold 0.1 to 0.2. The latter ensures a 98.53% true positive rate, with the cost of a 9.52% false discovery rate. Therefore, in the current analysis the threshold value chosen for the model is 0.2.
In light of the presented results, with the goal of identifying the anomalous cases of changes in assets’ ID codes between two quarters, assuming that the percentage of such changes in a quarter is 5% of all reported assets, a random forest model results as the best choice. In fact, among the tested models, the random forest ensures large accuracy and balanced accuracy. Moreover, selecting a probability threshold for classification of 0.2, the best model provides a true positive rate around 99% and a cost of a 9% in terms of false discovery rate that is considered acceptable in the DQM process.

3. Conclusions and further developments

Annual updates of the requirements or errors in insurance reporting can cause unexpected and undesirable changes in the reported asset codes, from quarter to quarter. An automated method to detect such changes is necessary to improve the data quality of insurance statistics, which are published at international level, given the size of the available data, the level of granularity on the single assets and the unneglectable impact that such changes have on compiled statistics.

A record linkage approach is proposed to reach the goal, making use of supervised machine learning classification models.

Real Italian data from 2021 are used for the application; four models are considered, i.e. logit model as a benchmark, bagging and random forests as random-tree-based machine learning models and neural networks. Robust results are presented, testing the models on differently sampled data, stratified on varying percentages of cases of changes in the codes in two adjacent quarters, since the true proportion of such anomalies in the data is currently unknown with precision.

The tested models show good performance in terms of average AUC and results show the superiority of the random forest model to approach the problem with respect to the other tested classifiers.

Assuming a 5% proportion of anomalies in the data and taking a 0.2 probability threshold for classification, the selected random forest model shows good performance for all measures of interest, both in terms of effectiveness and efficiency, ensuring large accuracy and balanced accuracy, with around 99% rate of correctly identified cases of changes in ID codes (TPR), accepting the cost of a false discovery rate that approaches 9%.

The presented test results give a robust estimate of the improvement in data quality that would derive from running the selected model on production data, with the goal of successfully identifying cases of unexpected and unwanted changes in the ID codes. However, the actual performance in production of the proposed methodology must be validated through the cross-check with the insurance corporations of the estimated cases of changes in a quarter during a real data production round; in that occasion TPR and FDR could decrease and increase, respectively, from the test results presented in the paper.
In the future, the model training phase might be improved by considering all the available Italian data, not only focusing on two reporting quarters but analyzing all couples of subsequent reporting quarters since 2016, in order to gain a larger amount of information on the assets. This possibility has not been explored yet due to the computational effort needed to elaborate all historical data available.

Moreover, further analyses might be conducted to evaluate the performance of the classifiers on different “asset type”, since the results might vary depending on that feature and potentially might be improved by training different models for specific category of assets.

Finally, in the future, the presented approach might be extended to a symmetrical data issue on the ID codes, such as the reuse of the same code for two different assets in two subsequent quarters, which is again an unexpected behavior in the reporting of the assets’ ID codes that can attempt at the quality of the data.

References


Statistical matching for anomaly detection in insurance assets granular reporting

Vittoria La Serra, Emiliano Svezia
Bank of Italy
Statistical Data Collection and Processing Directorate
Basel, 26th August 2022
• **Data Quality Management** (DQM) in central banks: necessary to ensure high quality in disseminated statistics.

• **Automation** of DQM processes is crucial:
  • to manage the volume of increasingly granular databases
  • to ensure resilience in situations of human resources constraints (pandemic)

• **Machine Learning** (ML) models: emerging to solve DQM issues

• **Proposal**: a record linkage approach using ML models to deal with a DQM issue on insurance granular assets data.
- **European Insurance Corporations** (ICs) quarterly report to national supervisory authorities and central banks since 2016 (Solvency II Directive).

- **Asset-by-asset information** is provided in template S.06.02 and used for statistical purposes by central banks.

- **Each asset** of an IC is reported with:
  - An identification (ID) code → required to be kept stable and consistent over time
  - A set of qualitative and quantitative features
• The DQM issue: reporting errors in ID codes might occur.

• Consequences:
  • two assets from two subsequent quarters are perceived as different when in reality are the same;
  • decrease in quality of IC statistics to be compiled and disseminated.

• The goal: to build a model that is capable of identifying pairs of assets that do not share the same ID code but actually refer to the same asset.
A record linkage approach

- Select two datasets containing assets from two subsequent quarters $Q_t$ and $Q_{t+1}$.
- Same assets are similar on the observed features → build a comparison matrix to compare all pairs of assets reported by the same IC on observed features (qualitative/quantitative) via distance measures.
- Fit supervised classification models on the matrix, where the target variable to predict is the binary status of each pair: \{match, non-match\}.

<table>
<thead>
<tr>
<th>Asset codes</th>
<th>Target</th>
<th>Distance measures on the observed features</th>
</tr>
</thead>
<tbody>
<tr>
<td>Quarter $Q_t$</td>
<td>Quarter $Q_{t+1}$</td>
<td>Status</td>
</tr>
<tr>
<td>Code A</td>
<td>Code A</td>
<td>Match</td>
</tr>
<tr>
<td>Code A</td>
<td>Code B</td>
<td>Non-match</td>
</tr>
<tr>
<td>Code B</td>
<td>Code B</td>
<td>Match</td>
</tr>
<tr>
<td>...</td>
<td>...</td>
<td>...</td>
</tr>
</tbody>
</table>
• From the Italian database, assets from two subsequent quarters are selected and compared, building the comparison matrix.

• 70,000 assets reported on average at each quarter → billions of pairs of assets to compare.

• Different supervised classification models have been trained and tested:
  o Logit (benchmark), bagging, random forest, neural networks.
  o Fine tuned for different hyperparameters combinations (e.g. number of trees, number of hidden layers).
  o Repeatedly fitted on differently unbalanced datasets (w.r.t. the target) to ensure robust results.
TEST RESULTS
for the Random Forest

- Hypothesizing 5-95% unbalance in the target
- Selecting a probability threshold of 0.2

<table>
<thead>
<tr>
<th>Balanced accuracy</th>
<th>99%</th>
</tr>
</thead>
<tbody>
<tr>
<td>Correctly classified cases of match (True positive rate)</td>
<td>98.5%</td>
</tr>
<tr>
<td>Erroneously classified cases of non-match (False discovery rate)</td>
<td>9.5%</td>
</tr>
</tbody>
</table>

Model | AUC index
---|---
Logit | 95.66%
Bagging | 98.62%
Neural network | 99.52%
Random forest | 99.64%

*Different percentages of unbalance
Conclusions

• The proposed methodology returns promising results to reach the goal with high performance.

• An automated method to detect errors in reported ID codes is necessary to ensure high quality of insurance statistics, given: the need for resilience in DQM processes; the volume of IC assets; the impact that such errors have on compiled statistics.

Further developments

• Improvement in the model training phase: considering all the available Italian data since 2016, not only focusing on two subsequent quarters.

• Evaluation of model performance on different “asset types” (e.g. securities, deposits, loans).

• Monitoring of production results: cross-check with the insurance corporations during a real data production round.
Thank you for your attention.

Vittoria La Serra  vittoria.laserra@bancaditalia.it
Emiliano Svezia  emiliano.svezia@bancaditalia.it
Making omelettes without new eggs: 
a story about a self-sufficient way of producing new statistic\textsuperscript{1}

Miguel Fonseca and Sonia Mota, 
Banco de Portugal

\textsuperscript{1} This presentation was prepared for the conference. The views expressed are those of the authors and do not necessarily reflect the views of the BIS, the IFC or the central banks and other institutions represented at the event.
Making omelettes without new eggs: a story about a self-sufficient way of producing new statistics

Sónia Mota and Miguel Fonseca

In this paper, Banco de Portugal presents its strategy to respond to the new requirements for statistical information on other financial intermediaries, financial auxiliaries and captive financial institutions and money lenders (OFIFA). The approach was fully anchored in the data management principles adopted by Banco de Portugal in 2018, which privilege the rationalization of data collection processes and the maximization of the usefulness of the data already available internally.

In this context, using a full range of microdata already available at Banco de Portugal – loans, securities, external transactions and individual financial statements (for supervisory purposes) – a set of information for each subsector of the OFIFA is already produced and published. The production of individual subsector statistics based on this method will allow to respond to the European Central Bank’s new requirements in the scope of financial accounts statistics (breakdowns for each subsector of the OFIFA), in an autonomous way from the institutions concerned.

Keywords: microdata; other financial intermediaries, financial auxiliaries and captive financial institutions and money lenders; balance sheet; financial accounts

JEL classification: E51; G23.

1 The opinions expressed here are those of the authors and not necessarily those of Banco de Portugal or the Eurosystem. Any errors and omissions are the sole responsibility of the authors. The authors are thankful to the comments and suggestions provided by Luis Teles Dias and Diogo Guerreiro.
Contents

Making omelettes without new eggs: a story about a self-sufficient way of producing new statistics........................................................................................................................................ 1

1. Introduction................................................................................................................................................................................... 3

2. The sources of information for the production of the OFIFA subsectors......... 4
   2.1. Balance sheet data .................................................................................................................................................................. 4
   2.2. How do we calculate the data gap for our quarterly commitments and present additional details for the financial accounts?................................................. 5

3. The OFIFA sector in numbers......................................................................................................................................................... 6

4. Conclusion..................................................................................................................................................................................... 10

5. References...................................................................................................................................................................................... 11
1. Introduction

In 2018, Banco de Portugal started an integrated information management program, which promoted a rationalization of data collection processes, and a greater sharing of internal information throughout all structures. The “report once” principle was at the forefront of this program, as one of the main goals was to avoid double reporting by institutions.

In this sense, internal task forces were constituted to catalogue available information and to identify the potential flows of data across units and departments, in order to break remaining information silos. It was through this process that Banco de Portugal was capable of retrieving the necessary information to compile the balance sheet items and financial accounts of the OFIFA sector, without requiring additional reporting from these institutions, thus, “making omelettes without new eggs”.

Previously to the European System of National and Regional Accounts 2010 (ESA 2010), OFIFA data was compiled mostly with basis on quarterly supervision information, reported to Banco de Portugal. With the manual, part of holding corporations were reclassified to the captive financial institutions and money lenders subsector (CFIML). These entities, however, did not qualify for supervisory reporting. This presented an added challenge, as no direct source of information was apparently available to guarantee the compilation of the entire OFIFA sector accounts.

To address this challenge, an administrative database with the balance sheet of all corporations operating in Portugal was used as source data. Nevertheless, the periodicity of information in this database is annual, presenting yet another challenge: How could Banco de Portugal fill the data gaps arising from the periodicity mismatch between the administrative data source (annual data) and the reporting requirements (quarterly data)?

The answer lied on a production method anchored on several complementing sources of information, which made use of a full range of microdata already available internally on loans, securities and external transactions. The production of individual OFIFA subsector statistics using this method allowed Banco de Portugal to timely respond to the European Central Bank’s new requirements for a breakdown of the OFIFA sector in terms of the financial accounts statistics.

In the remainder of the paper, we will firstly describe the sources of information used in the compilation of these statistics, explaining the needs each one aims to fulfil, and how they complement each other in order to produce the most accurate end-product possible. Then, using produced data stemming from those information sources, we characterize the OFIFA sector in the scope of the Portuguese financial system. The case of Banco de Portugal illustrates an example of how the “report once” principle was respected, while still being successful in meeting new statistical requirements.
2. The sources of information for the production of the OFIFA subsectors

Currently, Banco de Portugal produces and publishes balance sheet items information for each OFIFA subsector. This is possible through the combination of seven different sources of information. Each one aims to fill a specific need, complementing each other, in order to construct the most accurate picture of the balance sheet possible.

2.1. Balance sheet data

In order to compile the quarterly statistical balance sheet for each subsector, we use two primary sources of information: IES – *Informação Empresarial Simplificada* (census information approach) and FINREP (Financial Reporting).

IES is a mandatory annual survey to all corporations operating in Portugal, containing information mostly of accounting nature, including financial statements. After ESA 2010, and with the inclusion of holding corporations that were previously part of the non-financial corporations sector, IES became the main source of information for the production of the balance sheet of the OFIFA sector, with 99% of entities being currently retrieved from this source.

FINREP contains information of supervisory nature and is reported by credit institutions and investment firms. In the FINREP data, it is possible to find balance sheet information, with a higher level of granularity and a breakdown by counterpart sector, which IES does not offer. This allows a one for one correspondence with the statistical instruments, unlike IES, which has a lower level of detail.

Most importantly, FINREP is also consistent with the required quarterly periodicity of the data, while IES is only reported on an annual basis. Still, only 1% of total OFIFA entities currently report FINREP, and thus, to fill in for the data gaps arising for the remainder of the entities IES information is complemented with several other internal sources. This process is conducted while guaranteeing compatibility between the annual IES data, and the complementary quarterly information sources.

For the specific case of financial vehicle corporations engaged in lending (FVC), the main constituent of the other financial intermediaries subsector, there is a specific ECB regulation that allows for good information coverage. Through it, securitization corporations report loans conceded and debt securities issued, which are mandatorily reported to the ECB. The remainder of the balance sheet is compiled using available public financial statements and IES. In the case of securitization funds, Banco de Portugal receives an indirect report, shared by the Portuguese Securities Market Commission (CMVM).
2.2. How do we calculate the data gap for our quarterly commitments and present additional details for the financial accounts?

To compile quarterly information for the entities that have basis IES information, we use a set of five additional information sources available at Banco de Portugal, shared in the scope of the integrated information management principles put in place. Together, these information sources complete the balance sheet “puzzle”, with compatibility with IES annual data also being ensured.

For deposits vis-à-vis non-residents, information is retrieved from balance of payments and international investment position data (BoP) and, for deposits with resident banks, data are obtained from the balance sheet items data (BSI). The current BSI regulation foresees the reporting of the institutional sector counterpart, with a breakdown between all the OFIFA subsectors.

In terms of securities, Banco de Portugal manages a micro-database with data on securities issues and portfolios, on a “security-by-security” and “investor-by-investor” basis, denominated Securities Statistics Integrated System (SSIS). With SSIS, it is possible to determine the portfolios of all OFIFA entities, in short-term and long-term debt securities, quoted shares and on investment fund shares/units, as well as all the short-term and long-term debt securities issued by them, along with issued capital.

Regarding loans, there are three sources of information used: the Central Credit Register (CCR), BoP and ITENF.

The CCR is another database administered by Banco de Portugal, and discloses “entity-by-entity” and “loan-by-loan” information, reported by banks and other financial intermediaries operating in the country. CCR loan data is used for the identification of the counterpart sector for loans granted by the other financial intermediaries, and to identify all the loans granted by banks to OFIFA entities. BoP data is used for external loans operations. From ITENF (Inquérito Trimestral às Empresas Não Financeiras), a quarterly survey, it is possible to obtain a picture of the remaining loans for the CFIML subsector, the only one covered by the survey.

For the remaining assets and liabilities, information is also sampled from ITENF.

Table 1 below summarizes, by showing how each quarterly information source is used in the balance sheet, in order to fill the gap created by the annual frequency of IES information:
Making omelettes without new eggs: a story about a self-sufficient way of producing new statistics

Table 1 | Balance sheet items and the respective quarterly information source

<table>
<thead>
<tr>
<th>Assets</th>
<th>BSI</th>
<th>BoP</th>
</tr>
</thead>
<tbody>
<tr>
<td>Currency/deposits</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Debt securities</td>
<td>SSIS</td>
<td></td>
</tr>
<tr>
<td>Loans</td>
<td>ITENF</td>
<td>CCR</td>
</tr>
<tr>
<td>Equity</td>
<td>SSIS</td>
<td></td>
</tr>
<tr>
<td>Other assets</td>
<td>ITENF</td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Liabilities</th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Debt securities</td>
<td>SSIS</td>
<td></td>
</tr>
<tr>
<td>Loans</td>
<td>ITENF</td>
<td>CCR</td>
</tr>
<tr>
<td>Equity</td>
<td>SSIS</td>
<td></td>
</tr>
<tr>
<td>Other liabilities</td>
<td>ITENF</td>
<td></td>
</tr>
</tbody>
</table>

By combining all these sources of information, it becomes possible to produce both monetary and financial statistics and financial accounts for each subsector of the OFIFA, while avoiding multiplicity of reporting by these institutions.

3. The OFIFA sector in numbers

At the end of 2021, the Other Monetary Financial Institutions (banks) and the Central Bank comprised 73% of the total assets of the Portuguese financial system and the non-monetary financial institutions represented the remaining 27%. A different reality is observed concerning the number of entities, where the non-banking financial sector included 99% of total entities (Graph 1). These entities, besides not being authorised to receive deposits or close substitutes for deposits, are relevant for the economy since they channel or help channelling funds from savers, to entities that need financing. Some of the most relevant subsectors of the non-banking financial sector are the OFIFA sector, which represent 91% of the total entities in the financial system and 14% of total assets.
Graph 1 | Total assets and entities belonging to the Financial Sector in 2021

Despite the magnitude of the OFIFA sector in Portugal, in the last decade, we have seen a clear downsizing of the sector (Graph 2). The downsizing has been led by the other financial intermediaries, which saw a reduction of more than 60%, in total assets, since 2011, due mostly to a decrease in securitization operations by financial vehicles corporations after the European sovereign debt crisis. CFIML subsector also suffered a balance sheet reduction of 29%, driven by a simplification of group structures, which led to a fall in holding corporation’s assets, the most predominant type of institution of the subsector.

Although there has been a tendency for a reduction in the activity of the other financial intermediaries, financial auxiliaries and captive financial institutions and money lenders, in 2021 total assets rose by 344 million euros, ascending to 130.1 billion. This growth was fuelled by the captive financial institutions subsector.

Graph 2 | Total assets by OFIFA subsector (End of period, millions of euros)
A particularity of the OFIFA sector is its heterogeneous decomposition between the three subsectors, with the captive financial institutions and money lenders encompassing 73% of total OFIFA assets, the other financial intermediaries 22%, and lastly, the financial auxiliaries with only 5%, at the end of 2021.

From available balance sheet items information, we can also obtain a picture of the activity developed by the institutions belonging to each of the subsectors.

**Graph 3 |** Breakdown of the balance sheet asset side of the captive financial institutions and money lenders subsector (End of period, millions of euros)

Currently, the most predominant instrument on the asset side of the captive financial institutions and money lenders (Graph 3) are equity securities held, followed by loans. This is a consequence of the activity of holding corporations, which currently account for 90% of the assets of the subsector. These entities own controlling levels of equity in subsidiary companies and often concede financing to them, without developing any type of management. Thus, as expected, securities and loans dominate the balance sheet of the subsector, being also noticeable that both instruments now hold a larger share of the balance sheet than a decade ago.

In what concerns the other financial intermediaries, the dominant instrument on the asset side are loans, which arise from the operations of financial vehicle corporations engaged in securitization transactions, which currently hold 68% of total assets of the sector. These institutions’ main activity consists in acquiring loans from other institutions, while selling debt securities to finance these purchases.
While the composition of the balance sheet of the subsector has remained somewhat constant across time, the same cannot be said for its size, which as stated has decreased drastically due to a reduction in securitization activity (Graph 5). This reduction has been led by the fact that the European Central Bank started accepting the underlying loans (often non-performing) of these operations as collateral for monetary policy operations, and thus, banks lost one of the biggest incentives to take part in them. Before, banks would sell the loans to securitization entities, while buying the debt securities that the entity would use to finance the operations, with these securities in turn being eligible for monetary policy operations.

For loans conceded by the other financial intermediaries subsector, there is currently a higher level of breakdown available in the Central Credit Register, such as the debtor’s counterpart sector (Graph 6). Currently, Monetary Financial Institutions are the main counterpart of the lending with 38% of total loans, mostly due to non-
Making omelettes without new eggs: a story about a self-sufficient way of producing new statistics

derecognized securitization operations. Then, households and non-financial corporations represent 30% and 24% respectively, due to derecognized securitization operations and also to the activity of financial corporations engaged in lending.

**Graph 6 | Loans granted by other financial intermediaries, by counterpart sector (End of 2021, percentages)**

4. Conclusion

A bigger shift towards microdata and a rationalization of existing information are important tools to allow for a higher flexibility in the response to evolving statistical standards, which have continuously demanded higher levels of disaggregation.

Even if the collection of information based on microdata does not have as its initial objective the compilation of statistics, it allows the producer to be well equipped to respond to new reporting requirements requested by international organizations without additional costs for reporting institutions, would these come to arise.

This strategy was applied at Banco de Portugal, making it possible to respond to the new requisites for the production of financial accounts, which, from June 2022 onwards, required a breakdown between the OFIFA subsectors. In this sense, by combining information from micro-databases, such as loans and securities, and through a broader sharing of information, it was possible to produce the required sets of information, even without a guideline that mandated a direct reporting of the balance sheet by these institutions.

The outputs of the production process allowed for a characterization of the OFIFA sector in the scope of the Portuguese financial sector, including individual subsector balance sheets, and even a disaggregation of loans by counterparty sector, in the case of the other financial intermediaries.
5. References


[7] Instruction No 31/2005 of Banco de Portugal of 17 of September on the regulation of the reporting of statistical information to Banco de Portugal, for the compilation of securities statistics.

[8] Instruction No 27/2012 of Banco de Portugal of 17 of September on the regulation of the statistical reporting to Banco de Portugal, with the main objective of compiling statistics on external transactions and positions recorded in the balance of payments and the international investment position of Portugal.

[9] Instruction No 17/2018 of Banco de Portugal of 27 of August on the regulation of the communication to Banco de Portugal of actual or potential liabilities arising from credit operations, in any form or modality.

[10] Instruction No 14/2021 of Banco de Portugal of 18 November on the regulation of the reporting of balance sheet and interest rate statistics of monetary financial institutions.
MAKING OMELETTES WITHOUT NEW EGGS: A STORY ABOUT A SELF-SUFFICIENT WAY OF PRODUCING NEW STATISTICS

AUGUST 2022 | BASEL

SÓNIA MOTA | HEAD OF MONETARY AND FINANCIAL STATISTICS PRODUCTION UNIT
Lot of entities but few impact on the total assets of the financial sector

We do not want to overwhelm with new reporting requirements

How to compile the OFIFA sector without new eggs?
MORE NUMBERS, NO MORE EGGS

Change of Banco de Portugal mindset

Greater knowledge of the available information throughout Banco de Portugal

Information more visible and useful

Catalog all available information

Identify potential flows of data across departments

Make information methodologically clear for internal users

This is the key to compile the OFIFA sector, without requiring additional reporting from these institutions
IF WE DO NOT RECEIVE INFORMATION DIRECTLY FROM ENTITIES, WHERE DOES THE EGGS COME FROM?

FINREP - FINancial REPoring Standards

- Lost importance with ESA2010
- Mandatory supervision information reported by credit institutions
- Contain quarterly balance sheet information
- Several important breakdowns

1% of total OFIFA entities currently report FINREP

Coverage: supervised institutions by Banco de Portugal
IF WE DO NOT RECEIVE INFORMATION DIRECTLY FROM ENTITIES, WHERE DOES THE EGGS COME FROM?

Simplified business information survey

- Mandatory survey to all corporations operating in Portugal
- Contain financial statements
- Annual information

99% of the information of the OFIFA sector comes from this survey

Main entities covered: holdings corporations

CHALLENGE

How to guarantee the consistency between annual and quarterly data?
IF WE DO NOT RECEIVE INFORMATION DIRECTLY FROM ENTITIES, WHERE DOES THE EGGS COME FROM?

THE QUARTERLY “EGGS”

<table>
<thead>
<tr>
<th>Assets</th>
<th>BSI</th>
<th>BoP</th>
<th>SSIS</th>
<th>ITENF</th>
<th>CCR</th>
<th>BoP</th>
</tr>
</thead>
<tbody>
<tr>
<td>Currency/deposits</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Debt securities</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Loans</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Equity</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Other assets</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Liabilities</th>
<th>SSIS</th>
<th>ITENF</th>
<th>CCR</th>
<th>BoP</th>
</tr>
</thead>
<tbody>
<tr>
<td>Debt securities</td>
<td>SSIS</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Loans</td>
<td>ITENF</td>
<td>CCR</td>
<td>BoP</td>
<td></td>
</tr>
<tr>
<td>Equity</td>
<td>SSIS</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Other liabilities</td>
<td>ITENF</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

- **BSI**: Monetary and financial institutions balance sheet items
- **BoP**: Balance of Payment and International investment position data
- **SSIS**: Securities Statistics Integrated System
- **ITENF**: Quarterly corporate survey
- **CCR**: Central Credit Register
THE OFIFA SECTOR IN NUMBERS

Total assets by OFIFA subsector (End of period, millions of euros)

Balance sheet reduction over the years

Heterogeneous decomposition between the three subsectors
THE OFIFA SECTOR IN NUMBERS

Captive financial institutions and money lenders subsector – Assets (End of period, millions of euros)

Equity securities are the predominant instrument on the asset side

Other financial intermediaries - Assets (End of period, millions of euros)

Loans are the predominant instrument on the asset side
Loans granted by other financial intermediaries, by counterpart sector (End of 2021)

- **Monetary Financial Institutions** are the main counterpart due to non-derecognized securitization operations.

- **Households and non-financial corporations** represent 30% and 24% respectively, due to derecognized securitization operations and due to the other lending activity.
CONCLUSIONS

Rationalization and sharing of all available information, breaking down silos

Enhance the use of micro databases to respond to new reporting requirements without new cost for reporting agents

Banco de Portugal combine information from micro databases to produce the required sets of information with all the details and breakdowns
11th Biennial IFC Conference on “Post-pandemic landscape for central bank statistics”
BIS Basel, 25-26 August 2022

The euro short-term rate (€STR) – the new role of central bank statistics in financial markets – a financial benchmark fully based on statistical microdata

Ludovica Amorese, Javier Huerga and Ronald Rühmkorf,
European Central Bank

---

1 This presentation was prepared for the conference. The views expressed are those of the authors and do not necessarily reflect the views of the BIS, the IFC or the central banks and other institutions represented at the event.
The euro short-term rate (€STR) – the new role of central bank statistics in financial markets

A financial benchmark fully based on statistical microdata

L. Amorese, J. Huerga, R. Rühmkorf

Abstract

The euro short-term rate (€STR), which has been published by the European Central Bank (ECB) since October 2019, is the overnight interest rate benchmark for the euro. The launch of the €STR was part of the global reform of financial benchmarks. The €STR is determined every morning by the ECB on the basis of money market statistical microdata. This is a new use of statistical microdata and a new task beyond the classic statistical functions of central banks, with particular requirements in terms of governance, methodology, determination process and audit. The article focuses on these challenges as well as on the solutions implemented by the ECB, illustrating the evolving role of statistics and the use of statistical microdata in this field.

Keywords: euro short-term rate, €STR, financial benchmarks, microdata, granular data, money market

JEL classification: E43
Contents

Introduction ............................................................................................................................................... 3
  Historical background on financial benchmarks ........................................................................... 3
The euro short-term rate (€STR) ....................................................................................................... 4
  From EONIA to the €STR ............................................................................................................. 6
The €STR determination process ..................................................................................................... 7
  The €STR determination process in the ECB statistical function ............................................ 7
  Advantages of basing the €STR on money market statistical data ........................................... 9
Challenges and solutions in producing the €STR ................................................................. 10
  Governance .................................................................................................................................. 11
  Methodology ............................................................................................................................... 11
  Determination process ............................................................................................................... 13
  Audit .............................................................................................................................................. 14
Conclusions .............................................................................................................................................. 13
References ................................................................................................................................................ 14
Introduction

In the era of evidence-based policymaking, official statistics set the foundations for policy and decision-making that can impact millions of citizens and influence the expenditure of billions of euros. ECB official statistics are subject to high standards that ensure the accuracy, consistency and timeliness. They are produced in line with international standards, without any outside interference. The independence of official statistics is the pillar for safeguarding their quality, avoiding political influence and conflicts of interest that could affect the provision of data.

Following the global financial crisis, granular data collections and micro-databases have become increasingly relevant to produce statistics. As a result, the scope of central bank statistics is moving beyond aggregates to micro - hence more detailed - data. The use of granular data allows the both the direct use of microdata and the compilation of aggregates (on a regular and on a flexible ad-hoc basis). The Money Market Statistical Reporting (MMSR) and subsequently the euro short-term rate (€STR) are two noteworthy examples of how the ECB uses microdata.

Historical background on financial benchmarks

Cases of manipulation of interest rate benchmarks, as occurred some years ago with the London Interbank Offered Rate (LIBOR), had a negative impact on financial stability and ultimately affected the real economy. In this respect, initiatives at both global and European level were set up to introduce a global reform to address the vulnerability of some benchmarks to possible manipulation when volumes declined in the markets they were supposed to represent, ensuring the accuracy and integrity of the indices used as benchmarks and promoting near risk-free rates.

The principles for the reform of financial benchmarks obtained a wide consensus. However, in some cases previously existing benchmarks had difficulties in implementing these principles. For this reason, several central banks around the world decided to publish new financial benchmarks, addressing the risk of a sudden discontinuation of existing benchmarks and a potential market failure. Against this background, in September 2017 the ECB decided to develop the euro short-term rate (€STR) based on data already available to the Eurosystem.

The ECB was already collecting granular, timely, daily statistical data on the money market activities of selected euro area banks across four market segments: unsecured money market, secured money market, foreign exchange swap market and overnight index swaps market. The data were readily available to the ECB, in order to fulfil its tasks, in particular the monetary policy implementation. The data were considered of sufficient quality and timeliness to serve the daily production of a reference rate.

The €STR has been published by the ECB since 2 October 2019. The ECB, as administrator of the benchmark, has overall responsibility for providing the rate. The determination process of the €STR follows precise steps and rules, from the sending of the data by reporting agents up to the publication of the rate by the ECB. The €STR determination process is designed to provide the maximum quality while delivering the rate daily and on time, and complying, where relevant and appropriate, with the International Organization of Securities Commissions (IOSCO) principles on financial benchmark, covered in section 4.
This paper describes the €STR and its determination, as well as the challenges faced by the ECB in its production, illustrating the evolving role of statistics in this field. The €STR represents a new use of statistical microdata and a new task beyond the classical functions of central banks, with a number of new requirements.

The main features of the €STR are summarised in the second section, the €STR determination process and the advantages in producing the €STR from statistical microdata are covered in the third section. The fourth section describes the challenges and solutions related to the €STR. The fifth section concludes.

The euro short-term rate (€STR)

The €STR is a rate which reflects the wholesale euro unsecured overnight borrowing costs of euro area banks. It is exclusively based on borrowing transactions in euro conducted with financial counterparties deemed to be executed at arm’s length. Out of the potential MMSR instrument categories, the €STR is calculated using overnight unsecured fixed-rate deposit transactions over € 1 million. The €STR is calculated as a volume-weighted trimmed mean of the relevant transactions, removing the top and bottom 25% rates in terms of volume before calculating the mean.

The €STR publication is accompanied by complementary information on the data used for its computation (volume and number of €STR-eligible transactions), the reporting agents (number of banks reporting €STR-eligible transactions, share in the total volume of the largest 5 contributors), indicators on the dispersion of the data (rate at 25th and 75th percentiles), the publication type (standard or republication) and the calculation method (normal or contingency), see charts 1 and 2.

Chart 1. Euro short-term rate and quartiles

Percentages
The euro short-term rate (€STR) – the new role of central bank statistics in financial markets

The rate is published for each TARGET2 business day based on transactions conducted and settled on the previous day (reporting date T) with a maturity date on T+1.

From EONIA to the €STR

Following a carefully planned transition, which involved a transition period during which both benchmarks existed and during which EONIA methodology had been modified to become €STR plus a fixed spread on 8.5 basis points, the €STR successfully replaced the Euro OverNight Index Average (EONIA), as the benchmark overnight rate for the euro. The transition took place over several years, guided by a private sector working group on euro risk-free rates (WG RFR) in line with the guidance of the Financial Stability Board (FSB). Users of EONIA managed to successfully switch to the new benchmark within the required deadlines.

With the discontinuation of EONIA on 3 January 2022, the €STR became the only overnight benchmark rate for the euro. The €STR, much like previously EONIA, is now mainly used in derivatives such as Overnight Index Swaps (OIS) contracts. In response to the recommendations of the FSB, the WG RFR is considering other uses, including in cash market and cross-currency products.

The €STR is also the fallback in Euro Interbank Offered Rate (EURIBOR) contracts, should that rate cease to exist in future. The ISDA introduced €STR-based fallback provisions in its standard documentation to cater for discontinuation of EUR London Interbank Offered Rate (LIBOR) and EURIBOR. The WG RFR issued recommendations for €STR-based fallback rates in cash market products linked to EURIBOR. Depending on the asset class, the recommendations suggest using either forward-looking €STR rates (subject to their future availability), or a compounded €STR rate in all other cases.
In response to market feedback, the ECB publishes also a compounded €STR average rates and a compounded index based on the €STR. The rates are backward-looking compounded averages of the €STR calculated over standardised tenors of one week, one month, three months, six months and twelve months, see chart 3. The compounded €STR index makes it possible to calculate a compounded €STR average rate over any other tenor of choice. The ECB started on 15 April 2021 with the publication of the compounded average rates and a compounded index based on the €STR.

Benchmark rates like the €STR are a useful reference for many financial contracts, as they are publicly accessible, published by an independent institution on a regular basis following a transparent methodology that reflects market developments fairly and objectively, and are an important component for the monetary policy transmission. Reliable benchmarks are also necessary for the smooth functioning of money markets, and therefore for financial stability.

The €STR determination process

The €STR determination process in the ECB statistical function

The process of collecting MMSR data starts at the reporting agents, i.e. the 47 banks that currently constitute the MMSR reporting population. The reporting banks compile the data and send them to the corresponding National Central Bank (NCB) or ECB collection platform. The collection platforms receive the data and automatically perform initial checks on their format and content. The data are then submitted to the ECB by 07:00 Central European Time (CET).
The ECB performs additional quality checks and applies filters to automatically select the subset of transactions traded in the unsecured market segment eligible for calculating the €STR. Targeted data quality checks are then carried out exclusively on this subset. In particular, the correctness of the selected transactions is checked with reporting agents before the calculation of the €STR.

The €STR and the accompanying information are then automatically calculated and published, after a final check, at 08:00 CET. If errors with an impact larger than two basis points are detected following the publication, the ECB will revise and re-publish the €STR at 09:00 CET, although such an event has never occurred at the time of writing. No changes are made to the €STR after that time. At 09:15 CET, the compounded €STR (C-€STR) average rates and index, fully based on the €STR are published. Chart 4 shows a graphical representation of the €STR determination process.

The performance of the €STR determination process as part of the statistical function provides benefits for the overall daily process and data management. In general terms, the pre-existence of the MMSR data quality process provides a strong basis for the €STR, supported by sets of automatic checks and daily contacts by the ECB and NCBs with the reporting agents. The overall €STR data checking process includes the following elements:

- **Formal checks.** The formal correction of the reported files is ensured through the so-called level one (L1) checks. These checks ensure that the files received are fully in line with the technical requirements as established in the MMSR reporting requirements.

- **Basic business checks.** Checks on the data reported are then performed, rejecting data that do not comply with the MMSR rules or providing warnings in dubious cases. These checks, the so-called level two (L2) checks, ensure that the data reported for each transaction comply with the respective minimum standards on data quality.
• Plausibility checks. Plausibility checks performed on targeted on specific €STR eligible transactions, so-called level three (L3) checks, also take advantage of the pre-existing contact points and expertise in reporting agents. Reporting agents confirm the correctness of these selected transactions before the calculation of €STR, excluding them in case errors are detected.

• Post-production checks. The so-called level four (L4) checks, performed after €STR publication, aim at detecting potential systematic issues regarding €STR eligible transactions and are carried out together with the daily performance of MMSR data quality process. The L4 checks include outlier detection as well as consistency checks with other MMSR data.

From an institutional perspective, the statistical function also provides several features and objectives, as established in the Public Commitment on European Statistics by the ECB, that serve as well the €STR determination process. In particular, the following features are underlined:

• Institutional features. The ECB’s statistical function is granted with professional independence, legal mandates for the data collection, dedicated resources, an explicit commitment to quality, statistical confidentiality rules in place and aims at impartiality and objectivity.

• Processes and output. The ECB’s statistical function also has a commitment to appropriate procedures, minimizing the reporting burden and being cost-effective. The output objectives include relevance, accuracy and reliability, timeliness, consistency and accessibility.

Advantages of basing the €STR on money market statistical data

The €STR is based on daily confidential statistical information on individual money market transactions collected by the ECB. The use of pre-existing statistical data to produce the €STR involves a series of relevant technical and operational advantages:

• Avoid double reporting. The use of the pre-existing statistical data permits to avoid the establishment of an additional reporting for the purpose of calculating the €STR. The transaction-by-transaction reporting provides the necessary flexibility to aggregate the data according to the €STR methodology, in addition to other aggregations for analytical purposes, without changing the underlying data collection.

• Avoid submissions for the purpose of compiling a benchmark. The use of the pre-existing statistical data permits to avoid the collection of data submitted solely for the purpose of compiling a benchmark, a practice that can create additional vulnerabilities such as conflicts of interest and incentives for manipulation, as indicated in the IOSCO principles. The production of benchmarks based on submissions may also imply risks on the continuity of the benchmark, as these submissions are typically voluntary, which may result in an insufficient number of contributors, low business volumes, high concentration and/or low representativeness.

• Avoid expert judgement. The €STR is automatically calculated using pre-existing statistical data related to actual transactions conducted by the MMSR reporting agents in financial markets. No extrapolations or
adjustments are made to the reported values, therefore avoiding expert judgement as defined in the IOSCO principles.

- **Use of international standards and harmonized data.** The MMSR transaction-by-transaction data include the necessary attributes for the checking and use of the data. Each transaction is reported with a proprietary identifier assigned by each reporting agent, which permits to detect duplicates and facilitates the exchange of information on specific transactions. The data collection also requires the reporting of the Legal Entity Identifier (LEI) of the reporter and, if available, of the transaction counterparty. This allows standardization of the processes, a high degree of harmonization in the economic sector classification of the counterparties and deeper quality checks. The ECB ensures that a single economic sector classification is consistently assigned to each relevant institution for all statistical purposes in line with the European System of Accounts 2010, which implements the System of National Accounts 2008 in the European Union. ISO international standards are used wherever appropriate, including ISO20022 which defines the format of the reported data. This results in a single standardized reporting framework applicable across the euro area. The ISO20022 standard is also used for example by the Bank of England for its Sterling Money Market (SMM) data collection or for the Securities Financing Transactions (SFT) Reporting of the European Securities and Markets Authority (ESMA). This allows reporting agents to implement similar format standards for different reporting requirements, reducing the reporting burden and minimizing the risk of errors in the reporting systems.

- **Compliance procedure on statistical data submission.** The data quality management process is complemented by the overall statistical non-compliance framework that provides the tools for the monitoring and enforcement of the reporting agents compliance with the minimum standards of transmission, accuracy, and revisions, as defined in the MMSR Regulation. In case non-compliance is observed, an infringement procedure can be initiated, and sanctions may be imposed in accordance with the applicable legal framework.

- **Allow the publication of the pre-€STR.** The pre-€STR was an indicator published by the ECB previously to the launch of the €STR, for information purposes. The pre-€STR was also used to calculate the fixed spread EONIA-€STR for the transition period in which both benchmarks co-existed. During this time the EONIA was calculated as €STR plus the fixed spread. The pre-€STR, which included more than two years of daily data, could be provided to the market participants in preparation of the launch of the €STR thanks to the pre-existence of the MMSR data.

---

2 The International Organization for Standardization (ISO) is the most relevant international institution in the field of standardization.
Challenges and solutions in producing €STR

In July 2013 the IOSCO issued a set of principles for financial benchmarks (the IOSCO Principles), following a series of attempted market manipulation of global reference rates coupled with a declining liquidity in the underlying markets. These developments reduced the confidence of financial markets in the reliability and robustness of existing interbank interest rate benchmarks which play a pivotal role in the global financial system because of their usage in a broad range of financial products and contracts.

The IOSCO Principles were intended to promote the reliability of benchmark determinations, and address benchmark governance, quality and accountability mechanisms. The Principles provide a framework of standards that administrators should implement according to the specificities of each benchmark and cover various areas including governance aspects, the methodology of benchmarks, the determination process as well as accountability aspects.

As a central bank, the ECB is not within scope of the IOSCO Principles, nor is it subject to the EU Benchmark Regulation, since central banks already meet the principles, standards and procedures which ensure that it performs its activities with integrity and in an independent manner. Nevertheless, it was considered proper, where relevant and appropriate, for the €STR framework to be in line with international best practice for the provision of financial benchmarks, in particular with the IOSCO Principles. The implementation of these requirements goes beyond classical statistics in some aspects, as set out below.

Governance

While the €STR is based on transactions that reporting agents submit in accordance with the MMSR Regulation a distinction is made between the governance of the MMSR and of the €STR itself.

Being a statistical reporting, the MMSR is part of the ECB statistical framework, which is governed by the Statute of the European System of Central Banks (ESCB) and of the ECB. Regarding the collection of statistical information, the Statute establishes that the ECB, assisted by NCBs, collects the statistical information and contributes to the rules and practices governing the collection, compilation and distribution of statistics. In addition, the European Union Council Regulation (EC) 2533/98 concerning the collection of statistical information by the ECB lays down the regulatory powers of the ECB in the field of statistics. It establishes, inter alia, the potential reporting population for statistical purposes, the areas on which the ECB can collect data and the overall confidentiality and enforcement (non-compliance) aspects of ECB statistical data collections.

The Regulation (EU) No 1333/2014 (ECB/2014/48) provides the legal foundation for the MMSR data collection allowing to calculate the €STR on previously existing data. It directly addresses the reporting agents, establishes their daily reporting obligations including the scope and timeliness of the data. The MMSR Regulation lists and defines the specific attributes to be reported in respect of each transaction in the secured, unsecured, foreign exchange swap and overnight index swap euro money market segments. The MMSR Regulation also defines the roles of the NCBs with respect to...
the collection of the data and establishes as set of minimum standards regarding the transmission, accuracy, conceptual compliance and revisions, as well as minimum standards for data integrity. In case of non-compliance with the reporting obligations, an infringement procedure can be initiated, and sanctions may be imposed on the reporting agents.

In turn the €STR is specifically governed by Guideline (EU) 2019/1256 on the euro short-term (ECB/2019/19) (the €STR Guideline). ECB guidelines are legal acts that are legally binding on the ECB itself and/or on the ESCB members whose Member State has adopted the single currency. The €STR Guideline establishes the overall governance of the €STR, including the ECB’s responsibility for the administration and oversight of the €STR and the tasks and responsibilities of the ECB and NCBs with respect to their contribution to the €STR determination process and related procedures. Internal Eurosystem rules complement the Guideline by further specifying the required high criticality of the IT systems as well as detailing the operational tasks to be performed in the €STR calculation. The €STR governance framework ensures consistency with the €STR methodology and policies and also guarantees that no expert judgement is involved in the €STR.

The €STR Guideline also establishes the ECB’s control framework to protect the integrity and independence of the rate and its determination process and to address any existing or potential conflicts of interest that might otherwise compromise its integrity and reliability. The control framework refers to the ECB’s and the Eurosystem’s common corporate ethical culture as embedded in the ECB Ethics framework (which applies to all ECB staff), the ethical standards for all central banks of the Eurosystem (Eurosystem Guideline), established by the Governing Council, and the Code of Conduct for high-level ECB officials. For example, staff members and high-level officials are expressly prohibited from taking advantage of inside information in any private financial transaction or to recommend or advise against such transactions.

The €STR Guideline further establishes a €STR Oversight Committee which reviews, challenges and reports on all aspects of the €STR determination process as established by the €STR Guideline. The Oversight Committee acts as an advisory body to the ECB’s Executive Board and the Governing Council.

Methodology

The method of calculation of the €STR is defined in the published €STR methodology and policies. The €STR methodology was developed by the ECB’s Directorate General Market Operations together with experts from other business areas. To align the definition and features of new rate with the needs of the prospective user base two public consultations were conducted during the design phase. While the first consultation focused on more general considerations such as scope, the second consultation was dedicated to more detailed methodological aspects regarding key operational and technical parameters. Market participants strongly backed the proposals put forward for consultation. In addition, the design and implementation of the new unsecured overnight rate aims at being consistent with international best practices as set out in the IOSCO Principles.

The definition of the €STR has two elements regarding (1) the statement of the underlying interest represented by the euro short-term rate and (2) the statement of
the methodology setting out which transactions are considered for the calculation and the details on the calculation process. The underlying interest of the €STR was defined as the wholesale euro unsecured overnight borrowing costs of euro area banks. In comparison to EONIA that was based on interbank lending only, the €STR includes short-term borrowing from a wider set of counterparties by covering borrowing activity beyond the interbank segment. The €STR is calculated using euro denominated overnight unsecured fixed-rate deposit transactions over €1 million received from financial counterparties. Unsecured deposits are standardised and are the most frequent means of conducting arm’s length transactions on the basis of a competitive procedure, thereby limiting idiosyncratic factors potentially influencing the volatility of the rate. The €STR is calculated for each TARGET2 day as a volume weighted trimmed mean of the eligible transactions. The trimming aims at protecting the rate from idiosyncratic volatility caused by transactions priced off the market, or from errors in the underlying statistical data. The volume-weighted trimmed mean is calculated by (1) ordering transactions from the lowest rate to the highest rate; (2) aggregating the transactions occurring at each rate level; (3) removing the top and bottom 25% in volume terms; and (4) calculating the mean of the remaining 50% of the volume-weighted distribution of rates.

The methodology also includes a contingency formula for calculating the rate in case of insufficient underlying data. This could include cases where (i) there is a lack of data; (ii) there is a possible concentration of inputs; or (iii) systems break down, preventing a sufficient data feed and thereby hindering the calculation of a representative transaction-based rate. The contingency calculation is triggered in case the number of reporting banks is less than 20 or five banks account for 75% or more of total transaction volumes. In case the €STR is calculated based on the contingency formula it represents a weighted average of the previous day’s €STR and the rate resulting from using the data for the current day. Annual methodology reviews are conducted to confirm that the €STR remains a fair reflection of market movements, that it is backed by sufficient underlying data and that the scope and calculation method selected are therefore adequate.

In a further public consultation, the design of the compounded €STR (C-€STR) and its calculation and publication rules were consulted with the market participants. The C-€STR average rates are entirely computed by using the publicly available historical daily values of the €STR and yields and average rate for the respective tenor over which the €STR values were recorded.

**Determination process**

To ensure a reliable determination and publication of the €STR for the operational implementation a dedicated IT system with high criticality standards was set up. Eurosystem-internal operational procedures provide detailed guidance on the tasks of the operators. These were both tested extensively during a shadow production period of nine months before the start of the publication of the €STR.
The €STR is published on every TARGET2 business day\(^3\) at 8:00 CET based on the eligible transactions traded on the previous TARGET2 business day. The daily shift for the production of the rate starts at 6:30 CET. To minimize the risk of incidents affecting the production and publication of the rate, the proper functioning of the €STR IT system is regularly tested on weekends following IT maintenance activities with potential impact on the €STR production. As the production of the €STR requires staff working on weekends, public holidays and before regular office hours, a weekly shift work system is established. The shift work is performed on a rotating basis by staff members trained on all aspects of the production and publication of the rate. The rotation system also serves to reduce key person risk.

The production and publication of the €STR is highly automated and does not require the use of discretion, minimizing the risk of errors during the process. The collection of the data and the regular functioning of the processes up to the publication of the rate are monitored by the €STR operators who can initiate manual back-up procedures in case of need. The €STR operators also interact with reporting agents as part of the data quality process and perform a final check on the figures to be published.

Audit

Both internal and external audits are conducted in the context of the €STR. Internal audits have so far been conducted on the design and implementation of the €STR as well as on its operation following its go-live. The internal audits are performed in accordance with the ECB Audit Charter and the Audit Charter for the Eurosystem / ESCB and the Single Supervisory Mechanism.

The ECB also appointed an external auditor to independently assess the overall framework used by the ECB to administer the €STR for compliance with the IOSCO Principles. This included the control processes defined in relation to governance, quality, and accountability aspects over the €STR. The outcome of the audit was reflected in an assurance report published by the ECB.

Conclusions

The €STR, which has been published by the ECB since October 2019, is the overnight interest rate benchmark for the euro. The launch of the €STR was part of a global reform of financial benchmarks. The €STR is determined every morning by the ECB on the basis of money market statistical data.

The ECB Directorate General Statistics takes care of the determination process. This is a new use of statistical microdata and a new task beyond the classic statistical functions of central banks, with particular requirements in terms of governance, methodology, determination process and audits. The €STR benefits from a number of operational advantages resulting on being fully based on pre-existing statistical

---

\(^3\) TARGET2 is the real-time gross settlement (RTGS) system owned and operated by the Eurosystem. TARGET2 is open every day, with the exception of: Saturdays; Sundays; New Year’s Day; Good Friday and Easter Monday; 1 May (Labour Day); Christmas Day; and 26 December.
The euro short-term rate (€STR) – the new role of central bank statistics in financial markets

The euro short-term rate (€STR) involves particular requirements and challenges beyond classical statistical functions, which have been addressed through a dedicated legal act for €STR, a specific methodology developed by the ECB’s Directorate General Market Operations and involving public consultations, highly critical IT systems and internal and external audits.

The successful implementation of the €STR is an example of the increasing relevance of statistical microdata and of their usage beyond classical statistical functions and responsibilities.

References


ECB, PricewaterhouseCooper (2020), Independent assurance report of the ECB statement of compliance with control processes defined in relation to governance,
quality and accountability activities over the €STR and its adherence to the IOSCO Principles for Financial Benchmarks, September.


https://www.bis.org/publ/qtrpdf/r_qt1903e.pdf
The euro short-term rate (€STR)

The new role of central bank statistics in financial markets

25/08/2022

L. Amorese, J. Huerga, R. Rühmkorf
Following the global financial crisis, granular data collections and micro-databases became increasingly relevant to produce statistics and beyond.

In the context of the reform of financial benchmarks, the ECB decided in 2017 to develop the euro short-term rate (€STR) based on the Money Market Statistical Reporting (MMSR) data already available to the Eurosystem.

MMSR data were considered of sufficient quality and timeliness to serve the daily production of a reference rate.

The €STR has been published since 2 October 2019, by the ECB as administrator, on every day on which the Eurosystem TARGET2 payment system is open.

Following a transition period, the €STR successfully replaced the euro overnight interest average (EONIA) as the overnight benchmark rate for the euro and it is the fallback in the euro interbank offered rate (EURIBOR) in case of future cessation.

Money market statistical reporting (MMSR) and the euro short-term rate (€STR) are two noteworthy examples of how the ECB uses microdata for market surveillance and other purposes.
The €STR reflects the wholesale euro unsecured overnight borrowing costs of euro area banks.

It is exclusively based on borrowing transactions by MMSR reporters with financial counterparties on fixed rate unsecured overnight deposits denominated in euro, over €1 million.

It is calculated as a volume-weighted trimmed mean of the relevant transactions, removing the top and bottom 25% rates in terms of volume before computing the mean.
Figures published daily in addition to the rate:
- Business volume
- Number of €STR-eligible transactions
- Number of banks reporting €STR-eligible transactions
- Share in the total volume of the largest 5 contributors
- Rate at the 25th and 75th percentile as indicators on the dispersion of the data
- Publication type (standard or republication)
- Calculation method (normal or contingency)
- Compounded €STR average rates and compounded index - backward-looking calculated over tenors 1W, 1M, 3M, 6M, 12M
€STR determination process

- €STR determination process as part of the statistical function provides benefits for the overall daily process and data management.
- Pre-existence of the MMSR data and quality process avoids double reporting and provides a strong basis for the €STR with the reporting agents.
- It avoids submission solely for the purpose of compiling a benchmark.
- The use of actual transactions and automatic calculation avoids expert judgement.
- The data quality management process is complemented by the overall statistical non-compliance framework to ensure compliance with the minimum standards for the reporting of MMSR data.
€STR challenges - Governance

• The ECB, as a central bank, is not in the scope of the IOSCO Principles, nor subject to the EU Benchmark Regulation. Nevertheless, the €STR is in line with international best practice for the provision of financial benchmarks, in particular with the IOSCO Principles.

• EU and ECB regulations provide the legal foundation for the MMSR data, establishing the daily reporting obligations including the scope and timeliness of the data.

• €STR is specifically governed by the ECB €STR Guideline, including the ECB’s responsibility for the administration and oversight of the €STR and the tasks and responsibilities of euro area national central banks (NCBs).

• Internal Eurosystem rules complement the Guideline by specifying the required high criticality of the IT systems as well as detailing the operational tasks to be performed in the €STR calculation.

• The €STR guideline establishes the €STR Oversight Committee which reviews, challenges and reports on all aspects of the determination process.

• The €STR governance framework ensures consistency with the €STR methodology and policies and also guarantees that no expert judgement is involved in the €STR.
€STR challenges - Methodology

• The method of calculation of the €STR is defined in the published €STR methodology and policies. The €STR methodology was developed by the ECB’s Directorate General Market Operations.

• To align the definition and features of new rate with the needs of users, two public consultations were conducted to define the scope and key technical/operational parameters.

• The methodology also includes a contingency formula for calculating the rate in case of insufficient underlying data.

   This could include cases where:

   I. there is a lack of data
   II. there is a possible concentration of inputs or
   III. systems break down

• In case the €STR is calculated based on the contingency formula it represents a weighted average of the previous day’s €STR and the rate resulting from using the data for the current day.

• Annual methodology reviews are conducted
€STR challenges – Determination process and audit

- The €STR IT system with **high criticality standards** to ensure a reliable determination and publication.
- Internal procedures were **extensively tested** in the 9-months shadow production period which anticipated the €STR go-live.
- The ECB daily shift for the rate production starts at 06:30 CET on every TARGET2 business day.
- Additional IT **system testing** is regularly performed on **weekends** to **minimize** the risk of **incidents**.
- The production and publication of the €STR is **highly automated** and does not require the use of discretion, **minimizing** the risk of **errors** during the process.
- The €STR operators monitor the process, **interact** with reporting agents as part of the data quality process, perform manual tasks in case automatic procedures fail and carry out a **final check** on the figures to be published.
- Both **internal** and **external audits** are conducted in the context of the €STR.
  - Internal exercise – in accordance with the ECB Audit Charter and the Audit charter for the ESCB and SSM.
  - External exercise – to independently assess the overall framework used by the ECB as €STR administrator for compliance with the IOSCO principles.
Conclusions(1)

- MMSR and €STR confidential statistical information is **widely used** by ECB DG-M and NCBs for
  - the **definition** and **implementation** of the euro area monetary policy
  - **monitoring** the monetary policy transmission mechanism
  - gathering info on market **expectations** for future trajectory of policy rates
  - **analytical work** supporting Eurosystem policies

- The €STR is determined **every morning** by the ECB on the basis of money market statistical data

- The €STR benefits from a number of operational advantages resulting on being fully based on **pre-existing statistical microdata**

- At the same time the €STR involves **particular requirements** and **challenges beyond classical statistical functions**, addressed through a dedicated legal act for €STR, a specific methodology, highly critical IT systems and internal and external audits

- The successful implementation of the €STR is an example of the **increasing relevance** of statistical microdata and of their **usage** beyond classical statistical functions and responsibilities

---

(1) The opinions contained in this presentation are the sole responsibility of the authors and may not reflect the views of the ECB
Something old, something new: 
a reflexion on the new normal for statistical producers¹

Lígia Maria Nunes,  
Banco de Portugal

¹ This presentation was prepared for the conference. The views expressed are those of the author and do not necessarily reflect the views of the BIS, the IFC or the central banks and other institutions represented at the event.
Something old, something new: a reflection on the new normal for statistical producers

Lígia Maria Nunes

Traditionally central banks answer new statistical requirements based on legal documents such as regulations or guidelines. These are the times when statistical authorities ask reporting institutions for new data, adapt IT systems, strictly apply the established methodologies and when most of the work is done aiming to address mature needs. Then, COVID-19 pandemic came and quaked many routines. With the fear of having no data but the strong belief that decision makers needed to measure what was happening in the world, statistical producers challenged their way of working. COVID-19 turned several impossible into possible: with a deep sense of public service, teams started looking for alternative data and for new methodologies to anticipate the release of existing statistics and to produce new statistical indicators. Institutions cooperated faster than ever and from the trainee to the Board a new mindset came up: the statistical function has to be agile in such way that statistical authorities are capable to answer users’ needs in any circumstance.

The COVID-19 pandemic was an extreme event and the world is looking for its “new normal”. Inspired by some new approaches that were adopted by Banco de Portugal following the outbreak of the pandemic, this paper aims to reflect on what should be the “new normal” for statistical producers: should they go back to the traditional way of doing statistics or should their keep alive the recent learnings?

Keywords: statistical function, corporate culture

JEL classification: M14
1. Introduction

The COVID-19 pandemic disrupted the world. Economic relationships were broken making statisticians facing one of the biggest challenges of their professional lives: how to measure the unmeasurable. Senior statisticians warmed up to such a challenge in 2008 with the outbreak of the great financial crisis but, as we know, routines easily pull us down to our "statistical comfort zone". Here is my billion-dollar question: how many crisis do we need to internalize the “something new”?

This paper is a reflection on how much the Banco de Portugal’s statistics department is internalizing the lessons learned through the COVID-19 pandemic time.

2. Reflection #1: old data vs new data

Central bank statistics are traditionally compiled with data directly reported by institutions (e.g. assets and liabilities broken down by multiple criteria) and administrative data (e.g. public registers, tax files).

The great financial crisis flagged the importance of granular data, which became crucial to take the statistical function to the next level of producing whatever users need. The pandemic crisis confirmed that importance but flagged that it could be not enough. Firstly, because some traditional data – even those with monthly frequency – became insufficiently up to date to be useful at that time; secondly, because the outbreak of the pandemic made statistical authorities feared of not having data from the respondents due to offices closures.

New data was needed and statistical authorities started looking for alternative sources to complement official statistics and replace missing reports. In Banco de Portugal, we looked for “internal” new data sources, since we realised that already existing datasets – especially administrative data that had been collected for many years – were not duly exploited for statistical purposes: sometimes because potential users were not aware of their existence or because its sharing was difficult. Additionally, we looked for external new data, taking benefit from the new digital information, namely web-based indicators.

For instance, combining both strategies, the Banco de Portugal ensured that the publication of external statistics would not be disrupted. In particular, in May 2020, the Bank was able to publish the monthly preliminary indicator on the evolution of the travel component of the balance of payments. Compiled mainly with payment cards information, this preliminary indicator allowed to anticipate in 20 days the evolution of exports and imports of tourism and follow the impacts of COVID-19
pandemic in this sector, which in 2019 generated a direct and indirect contribution of almost 12% to the Portuguese GDP.

What did we learn?

Given its dynamic nature, measuring economic activity can be a never-ending process: there will always be new structural and unexpected developments that are not yet measured but should be. Indeed, the COVID-19 pandemic was a stark reminder that new challenges can constantly emerge that require changes in the statistical information produced and offered. As stated by Tissoet et al (2020) “it urges that statisticians figure out how complementary data sources can be brought into their mainstream statistical frameworks (i) by integrating alternative input sources within conventional methodological processes or (ii) using these additional sources to get supporting and benchmarking data that can act as an “information buffer” in times when conventional official statistics dry up or are lagged significantly”.

What is something new?

It became unquestionable that we need to make a better use of existing data. As mentioned, the COVID-19 context shed light on the existence of various datasets that may not be widely known and this suggests that simple quick wins could be achieved by better exploring and publicising internally the information available. To overcome this issue, in 2022, we implemented an internal data warehouse that centralizes all data received in the statistics department, with non-harmonized formats, turning its sharing easier. Also in 2022, with a department restructuring, an innovation team, highly specialized in mathematics and data science, was created under the integrated information management division to give answer to both challenges: to create mechanisms to turn easier the use of existing data and to explore the so-called alternative or big data sources, that currently mostly fall outside the official statistics boundary, but that could provide timelier, more frequent and complementary insights to traditional statistics.

3. Reflection #2: old data vs new statistics outputs

Traditionally central banks answer new statistical requirements based on legal documents such as regulations or guidelines. During several months, or even years, statistical authorities mature their needs, adapt reports and IT systems and strictly apply the established methodologies. This is the something old.

However, unexpected events, such as the COVID-19 pandemic, quaked all the routines and highlighted the urgency to have new, more timely and frequent statistics.

The need for timely statistics reflected the fact that the speed of the pandemic and the size of the economic disruptions had called for having more rapid data at hand to quickly assess the economic situation. Many central banks simply felt the need to do, whatever they could, to bring forward the publication of several statistics. In 2020, the Banco de Portugal anticipated the publication of monetary and financial
institutions balance sheet items and interest rates, statistics on securities issuance and the annual economic and financial indicators of the non-financial corporations.

At the same time new economic phenomena emerged and it became urgent to measure them. I would like to mention two of them. The first had to do with identifying the effects of the pandemic on business activity. The frequency of official statistics (quarterly and annually) turned them clearly insufficient to recognise new trends and to design policies to mitigate economic impacts on enterprises themselves and on the economy. Having this in mind, the National Statistics Institute (NSI) and Banco de Portugal launched the Fast and Exceptional Enterprise Survey, with the aim of identifying the effects of the pandemic on business activity. This survey was initially conducted every week, and then it was spread out. To launch a new survey could be seen as “the something old”. In fact, this is the traditional way through which statistical authorities ask for new data. However if one considers that the release of the first results took place on 14th April, less than one month after the Portuguese Government had declared the state of emergency and mandatory lockdown, maybe, we can find on it something really new. Faster than ever, the two statistical authorities – the NSI and the Banco de Portugal – worked together from the design of the survey until the publication of the main results.

The second phenomenon had to do with a set of measures approved by the Portuguese Government and by financial institutions to protect families and corporations, aiming to ensure their maintenance of liquidity. One of these key measures was the approval of loans moratoriums, a support to banks’ customers on their loans' obligations, which allowed the extension of the payments’ dates with the creditor institutions without registering a contractual default. At the time, an interdepartmental team was created at Banco de Portugal to monitor and discuss the impact of loans moratoriums under different perspectives – supervision, financial stability, legal and statistics.

When the discussion on how to get numbers started, the huge potential of the Banco de Portugal’s Central Credit Register was reinforced. In fact, through this micro database it was quite easy to request to financial institutions to identify the loans' contracts covered by these schemes. And in April 2020, the Banco de Portugal started compiling monthly data on loans under moratorium, namely, the amount of loans and the number of associated debtors for private individuals and non-financial corporations. During the first couple of months, this information was restricted for internal usage and, in March 2021, it was externally released in BPstat, the Bank’s statistical portal.

What did we learn?

In the “new normal” central banks should shape statisticians to a mindset of deeply understanding that our mission must go beyond the standard offering of official statistics. Statistical authorities must be ready to give timely information on the new issues raised by new events, especially when the traditional statistical apparatus does not properly cover them. It is time to move from this “answer-in-time-of-crisis” to a new modus operandi on central banks statistics.

Besides shaping teams, three additional points should be taken into account.
One has to do with our answer to the trade-off between quality and timeliness. Advances in the compilation processes are sometimes achieved at the price of reducing the quality in the aggregates measured. Are we willing to pay this price? If we are not there already, an alternative could be to look for other timelier indicators built using not the “traditional” but secondary data sources. They could be very useful to shed light on specific economic areas such as earlier estimates on credit granted to the economy, public debt, net lending/net borrowing of the economy, current and capital balance or non-financial corporations’ profitability.

The second point covers the fact that statistical frameworks have to become more flexible to address evolving users’ needs and the sheer uncertainty created by the crisis. As stated by Tissot et al (2020) “this called statisticians to re-assess the relevance and agility of their tools and methods to deliver required statistics. Key was to be able to meet the (new types of) users’ data demands as structural changes were quickly happening. They had in particular to “think the unthinkable”, in order to support policy makers despite the high-level data uncertainty […]. They also had to be innovative, for instance to deal with the methodological issues induced by the new policy measures taken in response”.

Lastly, digital innovation and institutional cooperation are key to help accelerating the production of official statistics. A more automated linking between granular datasets and the macro statistical frameworks could help to compress compilation times, provide more frequent indicators, and reduce revisions. In addition, higher levels of cooperation and coordination among statistical and public authorities could accelerate processes already in place to allow effective exchanges of information, especially when crisis occurred.

What is our something new?

New data explorations are a priority of the Banco de Portugal's current strategic plan. The recognition of the huge potential of micro databases combined with teams increasingly skilled to work with large datasets and the availability of new tools to do it, are powering the development of new analysis and outputs. Web analytics results have been confirming the importance of giving users relevant and relatable data. As an example, the statistical press releases on loans under moratorium was accessed, between April 2021 and January 2022, on average, 450 times per month, while the remaining statistical press releases had, on average, 360 accesses per month.

In the aftermath of the COVID-19 pandemic, the outbreak of a war. With the same sense of public service, teams starting exploring all available data to evaluate the exposure of the Portuguese economy to Russia and Ukraine. In a couple of days, using different databases it was possible to conclude that the exposure of Portuguese corporations and banks to these markets was quite low.

With the war, prices pressures and a new wave of challenges for non-financial corporations. Under this context, the NSI and Banco de Portugal decided to conduct a new edition of the Fast and Exceptional Enterprise Survey in May 2022, taking into account that the pandemic had not yet been overcome and that the effects of the armed conflict in Ukraine had intensified the problems in global supply chains and the increase in prices of energy and other essential products.

The Statistics Department’s strategic plan for 2021-25 encourages divisions in charge of statistical production to look constantly for new statistical indicators and new statistical outputs, especially related to emerging issues.
4. Reflection #3: old data vs new ways of communicating statistics

During the last years, central banks became more aware of the importance of the statistical communication function. In fact, several works have been published supporting the need of moving from a statistical dissemination approach to a communication approach. More than selecting a proper channel or the most creative way of showing data, communication is, and will always be, about putting users at the centre of our activity.

Traditionally central banks publish in their websites, or statistical portals, a large amount of data series that, time to time, are accompanied by some statistical press releases or additional analysis. However, overall, central banks statistical communication practices are still based on the regular data series publication.

The COVID-19 pandemic rocked this poor way of dealing with users. Under so much uncertainty, a sense of supporting users increased without precedents. As mentioned, central banks revisited the services they provided to their stakeholders, for instance by increasing the frequency and/or timeliness of their estimates to better support policy decision and several initiatives were taken to give information to enterprises and to the general public reducing the uncertainty of the that times.

As many central banks, the Banco de Portugal launched in BPstat a dedicated webpage to follow the impact of COVID-19 in several statistical indicators such as GDP, inflation, payment habits, banking loans and deposits, non-financial corporations profitability, public debt, general government financing, interest rates, tourism or current and capital account. Besides the main results, in this webpage the Banco de Portugal explained how to read the numbers and how to understand their evolution under that specific context.

What did we learn?

I dare to say that the COVID-19 pandemic context leveraged the statistical communication function in the Banco de Portugal. In that particular moment, we had not to convince anyone of the urgency to take care of our users. We built a team, the StatsComm Team, with volunteer people from different divisions and departments to discuss which main statistical indicators should be given and explained to users in BPstat. With strong habits of always presenting the same information using the same technical vocabulary, this was the beginning of a challenging journey of putting users above our “statistical comfort zone”. The results were encouraging: between April 2020 and May 2021 the BPstat webpage “The impact of COVID-19 on the Portuguese economy in 2020” was visited 22.5 thousand times, which represents about 40% of all the accesses to all the news published in BPstat during the same period.

What is our something new?

The StatsComm Team is, today, a multidisciplinary team of around 15 people belonging to five different divisions: central balance sheet, monetary and financial statistics, financial accounts, external statistics, financial system microdata and communication and planning division. With a deep understanding of specific statistics and an increasing knowledge in communication and data visualization...
themes, this team is in charge of creating statistical content to publicize and explain central bank statistics and it is committed in promoting statistical literacy among the society.

The communication initiatives promoted by the Banco de Portugal are now, increasingly targeted, to the most searched and relevant economic and statistical subjects at each time. So far, the Banco de Portugal is working directly with five different groups (entrepreneurs, students, teachers, journalists and the general public) and the statistical communication initiatives mainly rely in BPstat and social media – Twitter, YouTube and Instagram.

Under the department restructuring, a communication [and planning] division was created, mirroring how crucial a good communication strategy is to accomplish our mission.

5. Final remarks

Here is my billion-dollar question: how many crisis do we need to internalize the “something new”? Maybe, no more. Maybe, we only need to be challenged from time to time to remember the goals that we can achieve, the tasks we can do and how, with our work, we can make the world a more informed place and, thus, a better place to live.

Maybe we only need to be sure that we cannot turn back from the recent developments: we cannot give up on looking for new data sources, developing new statistical outputs or exploring new methodologies.

Maybe we only need to understand that we cannot give up on our independency to produce opportune and relevant information and that the way to do it is increasingly replacing aggregated reports by microdata.

Maybe we need to understand that whatever we do, we do it for our users, and that it is crucial to remind often our teams that every single one contributes to the overall user experience, even the most backstage worker who seemingly interacts with no one.

Maybe leaders need to increasingly support a mindset of “disruption”, encouraging teams to take calculated risk, using mistakes as learning and growth opportunities to explore new technologies, techniques and approaches.

Maybe it is time to take advantage of our highly qualified teams and the digital transformation process in course.

Maybe this crisis was a “wakeup call to deal with issues that had been neglected for too long” (Tissot et al, 2020). With no more excuses or complexities, it is time to understand and internalize the lessons learnt.

So let us be as flexible as we can, keep learning, keep doing, keep cooperating and building our own path to be prepared to accomplish our mission: to give users whatever they need in any circumstance.
6. References

BIS website - Covid-19 statistical resources

Bruno Tissot, Barend de Beer (2020); “Implications of Covid-19 for official statistics: a central banking perspective”; IFC working papers nº 20

Lígia Maria Nunes, Ana Colaço, Rita Marques, Mariana Oliveira (2021); “When reaching is not enough: 8 tips to engage with Central Banks’ data users; Banco de Portugal

Statistics Portugal (2022); Tourism Statistics 2021
SOMETHING OLD, SOMETHING NEW: A REFLECTION ON THE NEW NORMAL FOR STATISTICAL PRODUCERS

AUGUST 2022 | BASEL

LÍGIA MARIA NUNES | HEAD OF COMMUNICATION AND PLANNING DIVISION
HOW MANY CRISSES DO WE NEED TO INTERNALIZE THE “SOMETHING NEW”?
Reflection #1: OLD VS. NEW DATA

AGGREGATED DATA AND MICRODATA CAN NOT BE ENOUGH

ALTERNATIVE DATA SOURCES

TRAVEL EXPORTS (y-o-y rate of change)

-100% 0% 100% 200% 300% 400% 500%


Final rates (exports) Preliminary indicator (exports)
WHAT DID WE LEARN?

It urges that statisticians figure out how complementary data sources can be brought into their mainstream statistical frameworks (i) by integrating alternative input sources within conventional methodological processes or (ii) using these additional sources to get supporting and benchmarking data that can act as an “information buffer” in times when conventional official statistics dry up or are lagged significantly.

Tissot et al (2020)
2020 the Banco de Portugal anticipated the publication of several statistics.

**Fast and Exceptional Enterprise Survey**

<table>
<thead>
<tr>
<th>% of Corporations with Decreased Turnover during Lockdowns</th>
</tr>
</thead>
<tbody>
<tr>
<td>April 2020 1st fortnight</td>
</tr>
<tr>
<td>98%</td>
</tr>
<tr>
<td>February 2021 1st fortnight</td>
</tr>
<tr>
<td>96%</td>
</tr>
</tbody>
</table>

- **Accommodation and Food Service Activities**: 81% in April 2020, 62% in February 2021.
- **Total of Reporting Corporations**: 98% in April 2020, 62% in February 2021.

**Loans under Moratorium**

- **Private Individuals**: January 2021
  - 20 billion €
  - 8.8% of total loans
  - 408 thousand debtors
  - 16.1% of total debtors

Reflection #2: OLD VS. NEW STATISTICS OUTPUTS
WHAT DID WE LEARN?

SHAPE YOUR TEAMS

OUR MISSION MUST GO BEYOND THE STANDARD STATISTICAL OFFER
WHAT DID WE LEARN?

01
TIMELIER INDICATORS BUILT USING SECONDARY DATA SOURCES

02
MORE FLEXIBLE STATISTICAL FRAMEWORKS
“THINK THE UNTHINKABLE”

03
DIGITAL INNOVATION AND INSTITUTIONAL COOPERATIONS
NEW DATA EXPLORATIONS

- new data exploration tools
- web analytics results confirm the importance of giving users relevant and relatable data
- teams increasingly skilled to work with large datasets
- a Banco de Portugal's priority

OUR SOMETHING NEW
COVID-19 ROCKED THE POOR WAY WE WERE DEALING WITH USERS

OUR SOMETHING NEW

Statscomm team

• multidisciplinary team of 15 people
• a new statistical communication strategy target oriented
• a new communication [and planning] division at the Statistics department
How many crises do we need to internalize the “something new”?  

Maybe, **no more!**

- Keep looking for new data sources, statistical outputs and methodologies
- Don’t go back on the flexibility given by microdata
- Understand that **users matter**
- Encourage teams to **take risk**

**REMEMBER YOUR MISSION:** **GIVE USERS WHATEVER THEY NEED IN ANY CIRCUMSTANCE**
Product level greenhouse gas contents – how to get there?\textsuperscript{1}

Ulf von Kalckreuth,
Deutsche Bundesbank

\textsuperscript{1} This presentation was prepared for the conference. The views expressed are those of the author and do not necessarily reflect the views of the BIS, the IFC or the central banks and other institutions represented at the event.
Product level greenhouse gas contents – how to get there?\textsuperscript{1}

Ulf von Kalckreuth, Deutsche Bundesbank

Abstract

This note presents a system of greenhouse gas (GHG) content indicators for markets and policymakers. The system is lean and informative. It condenses the relevant product and enterprise-specific information into a single number: the GHG value. Like prices, GHG values are easy to understand, manage and communicate. The envisaged scenario is one in which, at all levels of production, goods and services have two tags – the financial price to pay and the GHG value.

GHG values of products are interdependent. The massive information processing this simultaneity involves can be handed over to the market. As in a price system, markets learn collectively and decentralised. A micro-simulation exercise based on existing statistics is carried out, and the results indicate that convergence is fast. With appropriate institutional underpinning, the disclosure of GHG values by producers may become self-sustaining. Ultimately, the note develops a number of specific policy options.

Keywords: greenhouse gas intensities, carbon accounting, green finance

JEL classification: Q56, Q51, C81

Contents

1 The vision............................................................................................................................................ 2
2 Coping with simultaneity.................................................................................................................. 3
   2.1 A reduced form for carbon contents...................................................................................... 3
   2.2 Micro-level information flows............................................................................................. 4
   2.3 Pulling oneself up by one’s bootstrap................................................................................. 5
3 A micro-simulation experiment on decentralised learning....................................................... 5
4 Policy implications........................................................................................................................... 7
References............................................................................................................................................... 10
Annex 1: An IO view on product level carbon contents................................................................. 11
Annex 2: GHG content for product groups, Germany 2018......................................................... 14

\textsuperscript{1} Contact: Ulf von Kalckreuth, Deutsche Bundesbank, Postfach 100602, 60006 Frankfurt am Main, Germany. Phone: +49 (0)69 9566 36010. Email: ulf.von-kalckreuth@bundesbank.de. This policy note is a condensed version of von Kalckreuth, Ulf, Pulling ourselves up by our bootstraps: the greenhouse gas value of products, enterprises and industries, Deutsche Bundesbank Discussion Paper 23/2022. The paper will published both as a SUERF policy note and as a contribution to the 11th IFC Biennial Conference, BIS Basel, 25-26 August 2022. It represents the author’s personal opinions and does not necessarily reflect the views of the Deutsche Bundesbank or its staff.
1 The vision

This note brings together three strands of scientific work: input-output (IO) methodology and its capability to keep track of indirect emissions in interlinked systems of production, the carbon accounting literature on how to evaluate carbon emissions in single companies, making “dual use” of financial and cost accounting methodology, and the work on GHG Protocol emission classes in environmental reporting.

At the heart of environmental problems is a situation in which the effect that producing and using goods has on scarce resources is not properly reflected in the price system. In the case of GHGs, the scarce resource is the capacity of the environment to absorb carbon emissions – or, to be more precise, the maximum permissible quantity of carbon emissions in line with global warming targets.

For a massive reduction of GHG emissions, it is vital that consumers, investors and policymakers be able to properly evaluate the environmental consequences of production activities so that they can make the right choices.

What is it one would ideally expect from an indicator system designed for climate mitigation and specifically for financial sustainability purposes? We need exact quantitative information on the relevant emissions at the level of both firms and products. All emissions, direct and indirect, need to be covered, the latter not as loose estimates, but based on realised material flows and micro-level production interdependencies. Granular information is notably scarce, especially on indirect emissions. But it is indeed granular information that is required to make meaningful distinctions that go beyond favouring products and firms in sectors with a low carbon intensity or selecting stocks that happen to be in high-tech sectors.

A metric that summarises the relevant information needed to make decisions on the production, use and consumption of goods and services is the GHG value, defined in this note as the total amount of carbon equivalents emitted in the course of production of a good or service, either directly or indirectly through the use of intermediate input products. The definition of indirect emissions is recursive, recurring to the GHG values of earlier production stages. The concept has two additional important complements: a process of information exchange between providers and users of intermediate inputs, as described by Kaplan and Ramanna (2021a, 2021b), and micro-level standards for the measuring of direct emissions, such as the one provided by the GHG Protocol.

There is a huge benefit in establishing and maintaining a system of product-level GHG values. Consumers can use them to compare alternatives. If they prefer less carbon-intensive alternatives and are willing to pay the price, this creates competitive pressure towards products with lower GHG contents. The pressure carries over to earlier stages of production: along the entire value chain, buyers of intermediate inputs will opt for less carbon-intensive alternatives. Whereas the effect of a carbon tax is working on the supply side, from the beginning to the end of the value chain,

---

2 There are other terms for the amount of GHG emitted directly and indirectly in the course of production. “GHG content” is mostly used for sectoral IO measures, while the terms “GHG footprint” and “GHG intensity” are general and not tied to any measurement framework. The term “E-liability” is a concept proposed by Kaplan and Ramanna (2021a, b) to characterise a process for collecting, processing and reporting information on GHG emissions in an accounting framework.
the effect of disclosing GHG values takes the other direction, on the demand side, from the final products to the primary inputs. Administration and policymakers can be provided with a solid foundation for classifying firms – for taxes or subsidies, industrial policy or taxonomies for sustainable finance purposes. As an example: GHG value information is precisely what is needed to get EU plans for a carbon border adjustment mechanism off the ground. At each stage of production, the metric captures and carries forward the environmental resources that have been used up to that point. In a peer group of goods that are close substitutes, GHG values allow for the identification of inefficient producers and production technologies. Regarding unrelated goods, consumers and policymakers can compare and weigh their respective usefulness against their consequences for the climate. GHG values are like real rates of exchange between products and their consequences for the environment. It is a quantity structure that makes it possible to trace the price effects of carbon reduction policies at all levels – an important input for monetary policy in the transition to a low carbon economy. It may also be used to derive targets for allocation purposes.

This is all that measurement can give. The rest of this note discusses the methodology that will enable implementation. The key is the recursive nature of the metric, enabling Input-Output (IO) analytics, and decentralised data generation from an exchange of information between buyers and sellers of inputs. The iterative process can be started based on existing statistics!

2 Coping with simultaneity

As mentioned above, the GHG value is defined recursively: it is the sum of direct emissions attributed to a product and the GHG values of all inputs, covering indirect emissions. Indirect emissions are a sum of input requirements, multiplied by their respective unit GHG values, see Annex 1 for the formal definition. This is easy enough, but in order to use the definition directly, we need to have input GHGs in order to compute the GHG of output. How can GHG values of outputs be calculated in a world where not all the GHG values of inputs are known? GHG values are interdependent – the value for any product will depend on the value of all inputs.

2.1 A reduced form for carbon contents

Using a linear setting that starts from input requirements on the product level, we can express the solution of a system of interlinked indicators as a reduced form, using the apparatus of Input-Output (IO) analytics, see Annex 1. This reduced form yields a reference point for product level GHG values and it allows to compute GHG values for product groups, i.e. on the aggregate level. Yet, on the product level, computing the reduced form is not possible, as the required information will never be available centrally. The key message of this note is to show that this is not necessary. Producers

---

3 As part of the European Green Deal, the European Commission intends to put a carbon price on targeted imports by 2026 to avoid “carbon leakage”, i.e. the migration of industries to countries with more relaxed emissions policies. Technically, importers need to buy carbon certificates corresponding to the carbon price that would have been paid if the products had been produced in the European Union; see here official information with further links to the proposed legislation. Without a quantification of carbon content, the WTO may well consider the proposal an illegal tariff.
do not have to be aware of all the stages of the value chain – they only need to know their own technology and the GHG values of the inputs provided by their immediate suppliers.

2.2 Micro-level information flows

Kaplan and Ramanna (2021a, b) have recently argued that for solving the conundrum of production interactions, direct information flows between input providers and the buyers of inputs are needed. Kaplan and Ramanna suggest recording direct emissions as an “environmental liability”, or E-liability, and passing them on to the buyers of inputs, in the same way as a company’s value added is passed on to the buyers of an input. According to their proposal, E-liabilities are created when a company emits greenhouse gas in the course of production. They are acquired when an intermediate input with an E-liability attached is bought. In this case, a GHG account of the seller is credited and the respective account of the buyer is debited. The E-liabilities corresponding to direct emissions and to intermediate inputs will be assigned to products. The E-liability of the output will thus embody the direct emissions of all earlier stages. If the product is sold, either for final use or as an intermediate input to an external client, the company account is credited with the E-liability of that good. The E-liability characterises the product and is attached to it, and it leaves the firm with the output. On the other side, the GHG account of the buyer is debited. At the company level, any change in the E-liability over a given time interval will reflect the GHG content of inventory changes.

E-liabilities are framed as a close, almost perfect analogue to costs. Both are valued resource consumption. The input vectors may figure both in cost accounting and in E-liability accounting, with only the valuation differing – for standard cost calculation it is financial prices, whereas in the context of carbon accounting it is the E-liabilities of inputs. This enables the use of standard accounting techniques, the outcome of centuries of experience with valuation problems. Actually, E-liabilities are fraught with a large number of such valuation issues: emissions from overhead activities such as the heating of production facilities and office buildings, transportation, the E-liabilities of capital goods, or combined production technologies. These require the accounting allocation of company-level costs. The cost accounting solutions that exist simply need to be applied to the task of calculating E-liabilities. For an earlier literature review on carbon accounting, see Stechemesser and Guenther (2012).

Kaplan and Ramanna leave it to the companies to decide just how they wish to allocate costs, provided that the accounting identities are respected and the allocation follows respected accounting principles. With a valuation vector for input goods at hand, it is possible to carry out information aggregation and processing using standard cost accounting software, both at the product and at the enterprise level.

This raises a question we already have encountered. Whenever and wherever this system will start to operate, it will do so in a world where there are no E-liabilities for inputs from outside the company. How can those inputs be evaluated in E-liability accounting? If the flows of inputs in the value chain were unidirectional, from primary raw materials to high-tech products, it would be possible to work out E-liabilities sequentially, but what if a woodcutter needs a chain-saw, or a car, or a computer? Even disregarding circularity, the issue of missing valuations will persist, be it for
imported goods or with regard to producers that will be exempt for a variety of reasons. Thus, in order to become operational, the concept needs to be adapted to circumstances in which input providers cannot or do not want to declare their E-liabilities.

By imposing additional structure, the analytical view developed here will allow us to do so. Note that by considering all GHG value equations jointly and solving them for the reduced form, it is assumed that the definition of inputs is the same over processes. This also means that the allocation rules should be the same or at least comparable. Without this restriction, E-liability measurement is consistent between buyers and sellers of inputs, but not necessarily comparable between firms producing similar or identical goods.

One other thing needed for comparability is a protocol for the measurement of direct emissions. A well-known and widely accepted rulebook for the measurement of direct emissions is provided by the GHG Protocol developed and supported by the World Resources Institute (WRI). The existing ISO norms visibly build on the GHG Protocol.

### 2.3 Pulling oneself up by one’s bootstrap

A key insight of this note regarding incomplete information on input GHG values is the following: Starting from estimates and using the GHG values provided by their suppliers whenever available, the GHG values computed by producers will converge to the true values. Instead of centralised processing, the market will perform the task in a decentralised and iterative manner. This is shown analytically in von Kalckreuth (2022) using a result on matrix power expansion, and the argument is supported by an extensive micro-simulation on the basis of sectoral data from Germany. The result on decentralised learning has a powerful implication: as technology and direct emission intensity change over time, the GHG measures provided by the market system will follow suit, staying informative, without the need for any central institution to take account and intervene.

These results have a simple intuition: GHG values are an analogue to economic value, with direct emissions playing the role of the value added of a production stage and indirect emissions corresponding to the value of intermediate inputs. Just as input prices are processed in cost calculation, the GHG values of activities and products can be passed along the stages of the value chain. We can look at this as social learning – the participants of the production systems are like interconnected, information processing neurons.

### 3 A micro-simulation experiment on decentralised learning

Imagine that producers try to compute GHG values for their products based on their set of input requirements and incomplete information on the GHG values of inputs. If available, they will use GHG value indicators provided by their suppliers. If not, they will substitute estimates derived from existing aggregate statistics into the equation.

---

4 This is analogous to the processing of information on economic scarcity in the price mechanism, as described by Hayek (1945).
They will pass the resulting indicators on to their clients. Thus, in the first round, all information on GHG values comes from sectoral averages. In subsequent stages, input indicators will be product-based. A simulation allows us to see how the system evolves.

On the basis of statistical information on Germany for 2018, the author has simulated production interactions for a set of 7,699 hypothetical goods, each belonging to one of 71 product groups. For these product groups, the information on sectoral production interlinkages is available from Destatis (2021a) IO tables. The model is essentially an inflated and stochastic version of sectoral IO tables and direct emissions data for Germany, calibrated to reproduce the micro-level within-sector heterogeneity of direct emissions and Scope 2 emissions. The details are provided by von Kalckreuth (2022). The input coefficients matrix for the moderate number of products in the simulation has a size of $7,699^2$ cells, almost 60 million. This is enough to see that a centralised approach is not feasible for any realistic set of products.

By computing the reduced form as derived in Annex 1, the model is solved for direct and indirect emissions. These are the ‘true values’, and we can observe how well the first round proxies and the micro level measures of GHG values perform.

Graph 1: GHG values -- simulating 7,699 products based on German data

![Graph 1: GHG values](image)

Author’s estimates and computation, based on Destatis (2021a) and Destatis (2021b).

Graph 1 displays the resulting distribution of direct emission versus indirect emission intensities. Three sectors have been singled out visually. Electricity and heat are the source of Scope 3 emissions, and they are characterised by very high direct and indirect emissions per unit of output. The GHG intensity of agriculture is also high, specifically because of CH$_4$ (methane) due to livestock farming and NO$_2$ (nitrous oxide) due to fertilisers. Lastly, basic iron and steel are characterised by high indirect emissions, mostly due to heavy use of products from the same sector.

In order to get the iterations going, we need proxies as starting values. Group-level GHG contents are calculated from the IO tables and the direct emission statistics. For computations, the author uses the IO matrix for total production plus imports and the emission intensity calculated as national emissions over national production by product group. This assumes that production abroad is carried out with the same technology as national production. The shortcut is not satisfactory, but certainly more
appropriate than considering only national production interlinkages in calculating GHG values. The results for sectoral GHG content are shown as a graph in Annex 2, see von Kalckreuth (2022).5

The simulation then traces the evolution of GHG value measures in a situation where each producer only knows the input coefficients of their own product and the best effort GHG value estimates of others. This can be carried out on a small and decentralised information base, and the micro-simulation allows a study of whether and how fast decentralised learning converges to the true value.

Graph 2: GHG value learning using sectoral GHG contents as initial value

Graph 2 gives, for all product groups, a graphic representation of learning in iterations 1 to 10, depicting the mean absolute distance of estimates from the true GHG values, normalised by the level of GHG value.

Convergence is very fast for almost all products. A major exception is product group CPA 24.1.-24.4, Basic iron and steel, products of the first processing of steel, where convergence is visibly slower. This is related to a large input coefficient of 0.5 for inputs of goods of the same sector. With this, a “wrong” prior set of GHG values will be transmitted with a relatively high weight to the next iteration of the learning process.

4 Policy implications

To a certain extent, GHG value disclosure may rely on voluntary action by producers. There is a distinct commercial interest in obtaining and communicating carbon

---

5 The product group-level GHG contents have been calculated here for the purposes of a simulation exercise, and they are not meant to provide statistical information. Still, as a consistency check they can be compared against the results for Germany in Destatis (2019). This publication gives a detailed evaluation of the carbon content of final use in 2008 to 2015. In spite of the differences in the reference year and some important methodological aspects, the figures compare well for the GHG content of overall production, electricity and heat and the large industrial and service sectors. For methodological differences, see von Kalckreuth (2022, fn 20 and 21.)
accounting information. It is also clear that some firms have good reasons to declare the GHG value of their products either incorrectly or not at all. Just as with financial costs, if there is an opportunity to make products look cheaper than they really are or to avoid talking about costs altogether, some market participants will take it.

Thus, to establish a GHG value system, some reporting obligations will be necessary. This section makes the argument that reporting obligations may not have to be broad-based. Instead, legislation only needs to make sure that a threshold volume of disclosure, e.g. from large companies, is surpassed. Under certain conditions, this will trigger a process that will end in almost universal voluntary disclosure.

If companies subscribe to an E-liability system as envisaged by Kaplan and Ramanna (2021b), it is not in their interest to conceal GHG values – if they choose not to declare their E-liability, they will not get the credit. External auditing is needed to make sure that GHG debits and credits match inventory changes. In addition, basic valuation standards need to be backed up. Producers may rig the valuation of inputs that have no GHG values attached. By distorting the accounting allocation, they can make their output look “too cheap” in terms of GHG or cross-subsidise one product line with GHG-sensitive demand by charging other product lines where demand is inelastic. Thus, as a first component, formal auditing needs to make sure that the GHG value measure is a fair estimate, using the information on direct emissions and production interlinkages existing at the company level. Auditing is carried out against disclosure standards that have to be specified in advance. It is best organised in parallel with financial auditing, with governments having the right to scrutinise dubious statements. In this respect, it is promising that the IFRS is about to change its statutes in order to set up a board on disclosing standards for environmental information.

Second, an information platform is required that makes accessible the information available on:

- industry averages;
- direct emissions from company reports;
- product level GHG values, as far as they exist.

There is a path that leads to voluntary disclosure by (almost) all firms. Suppose that the information platform, in addition to making existing information publicly available, computes estimated average GHG content for firms of a given industry that do not disclose their E-liabilities – from the known industry averages and the known E-liabilities of the firms that do disclose. These estimates will be used to evaluate the average GHG values of inputs produced by non-disclosing firms.

Producers with low GHG values, relative to their peer group, will have a clear incentive to disclose, especially if they are active in GHG-sensitive markets. With low GHG values, they can charge higher prices to their buyers of intermediate or final goods or reap the rewards of positive publicity. This fact will generate a signal value for the decision not to disclose. The signal will be reinforced by calculating sector averages for GHG values conditional on not disclosing. With many companies disclosing, those that do not disclose will look increasingly unattractive. We may envisage an iterative process where first the cleanest firms disclose, then those that are not top tier, but still well above average, etc. In the end, the only firms to not
disclose will be those with rather extreme GHG values, and the fact that they do not disclose will be informative enough.

This unravelling due to taking the average over the ever smaller number of non-disclosing companies is quite similar to the Stiglitz and Weiss (1981) account of the possible breakdown of the credit market under asymmetric information. In the scenario at hand, however, the result is a separating equilibrium with voluntary disclosure. In order to create an incentive that is great enough to trigger this mechanism on a large scale, we may need to overcome a threshold number of participating firms.

Just as in a system of market prices, information on carbon usage can be processed in a decentralised and efficient manner, even without specific disclosure obligation. The key ingredients are micro-level auditing and a centralised information platform. This is where central banks may have an important role to play. They need to collect much of the required information anyway in order to classify their assets and collateral and, in some important cases, to rate companies. In addition, they have the mandate to disseminate statistical information for policy purposes as well as all the necessary infrastructure, experience and working routines.

The following is a list of policy options for central banks resulting from the discussion above:

1. Cooperate with statistical offices in setting up rather disaggregated IO model for economic areas, such as the EU, and larger participant countries. The corresponding direct emission information is also needed. This is very effective in creating a joint framework for condensing data at a product group level.\(^6\)

2. Set up and maintain a dissemination platform for GHG content data at the level of sectors, enterprises and products. Disclosure standards may oblige producers to use GHG contents published on that platform for their inputs. These platforms may also name and make available reference proxies for cases where product-level information is not available, especially in the case of imports.

3. Support development and propagation of disclosure standards and assist in setting them as a basis for comparability and auditing. Those rules can build on the relevant GHG Protocol standards, at least for direct emissions.\(^7\) What inputs to consider and how to evaluate them needs to be determined.\(^8\)

4. Interact with government bodies and the IFRS on potential disclosure requirements. In the case of the EU, the ongoing work on the CSRD is especially relevant. In light of the discussion above, possible disclosure requirements should target large companies so as to overcome a threshold that will induce

---

\(^6\) This matches Recommendation 1 in the suggested work plan for the new G20 Data Gaps initiative.

\(^7\) See, in particular, WRI and WBCSD (2004) as a standard for Scope 1 and 2 emissions, WRI and WBCSD (2011a) for Scope 3 emissions at the enterprise level, WRI and WBCSD (2011b) for corresponding standards for Scope 3 emissions at the product level, and the Technical Guide on Scope 3 emissions in WRI and WBCSD (2013).

\(^8\) See specifically ISO 14064-1:2018 on GHG reporting at the organisation level and ISO 14067:2018 on reporting at the product level.
voluntary disclosure by others, and by producers of primary goods and import goods so as to ensure valid GHG values at the front end of the value chain.9

A limited GHG value system is conceivable, aiming at a targeted subset of products only, such as energy, transportation, agriculture and parts of the manufacturing sector; see Annex 2 for an overview. This is easier to initiate and may still reap large parts of the overall benefit. But the simplicity comes at a cost. For many input goods, GHG value coefficients will have to be imputed permanently, as even in the long run there will be no values from input providers. Good proxies would be more essential than in an encompassing GHG value system. But even in this limited version, granular information would come from producers using their private knowledge on the input composition. This is perhaps the most important feature of a system of GHG values.

References


9 As already mentioned, the upcoming legislation on the carbon border adjustment mechanism and the enhancement of the scope of the carbon emissions trading systems are beneficial in this respect.


WRI and WBCSD, Product Life Cycle Accounting and Reporting Standard: A standardized methodology to quantify and report GHG emissions associated with individual products throughout their life cycle, 2011b.


Annex 1: An IO view on product level carbon contents

The GHG value encompasses both direct and indirect GHG emissions as a consequence of the production of a good or service. Indirect emissions are the result of direct emissions in a chain – or rather a fabric – of other production processes. Those production interlinkages are key for the consistent treatment of indirect emissions. IO analysis is designed for this type of interlinkages, and in fact it has been used in tackling the issue of attributing resource consumption to final output at the sectoral level since the 1970s. IO analysis makes the structure of an interlinked system of GHG values accessible.

To fix ideas, consider the following. In production planning, every process is defined by a bill of material (BoM) that specifies all inputs, plus a route sheet that explains how to combine them. A complex production process may be decomposed into several stages. Consider the BoM of product $k$,

$$ a_k = (a_{i1}, a_{i2}, \ldots, a_{ik})' $$

with $a_{ij}$ being the quantity of good $i$ that enters the production process. There are entries for all input goods in the economy, most of them with a value of zero, of course. Let the amount of GHG emitted directly be given as $d_i$. Let scalar $g_i$ be the GHG value of good $i$, the quantity of GHG that is emitted in the production of one unit. List the GHG values of all input goods in a vector as well:

$$ g = (g_1, g_2, \ldots, g_k)' $$
The GHG value of product \( k \) is then given as the sum of direct and indirect emissions. Importantly, we do not add a definition for indirect emissions, but simply define them recursively as the GHG values of inputs:

\[
g_k = d_k + g' a_k = d_k + \sum_i g_i a_{i_k}.
\]

Indirect emissions are the direct emissions at earlier stages of the value chain. The equation is both perfectly general and encompassing. It relates to products and activities and – for a defined time span – to enterprises and sectors as well.

As it stands, the equation is a definition. It helps us understand the problems associated with gathering and processing information. For actual computation, all the \( g_i \) corresponding to the BoM of product \( k \) are required. If these are known, we can calculate the GHG value of product \( k \) in a straightforward way from direct emissions and the BoM. This is like computing the energy content of food: it is enough that producers know the composition of their product and the energy content of the ingredients.

If the relevant elements of \( g \) are unknown, we can use equation (1) recursively and try to compute the GHG values involved, going up the value chain from more complex intermediate inputs down to primary and primitive inputs. The structure is well known from linear production planning and IO analysis, pioneered by Wassily Leontief, and it was indeed the same author who first proposed using IO models for analysing pollution generation associated with inter-industry activity.\(^{10}\) Conceptually, we can solve for the GHG values of all products simultaneously. Let

\[
A = \begin{pmatrix} a_1 & a_2 & \ldots & a_K \end{pmatrix}
\]

be the matrix of the BoMs for all output goods, \( 1, \ldots, K \). With \( d \) being the column vector of the associated direct emissions, one may write:

\[
g' = d' + g' A
\]

Reordering and postmultiplying the “Leontief inverse” \((I-A)^{-1}\) yields:

\[
g' = d'(I-A)^{-1}
\]

The GHG values of products (product \( k \) and all the others) result from their own direct emissions and the direct emissions of all the intermediate goods used for their production by intermediation of a matrix derived from the BoM that reflects the interlinkages in production. If the coefficients in the GHG equation refer to empirical production technologies actually being used to produce goods, \( 1, \ldots, K \), it can be taken for granted that the inverse exists and all its elements are non-negative.

As simple and beautiful as this relationship is, it is not possible to use it directly. Matrix \( A \) comprises the BoMs for all products in the economy, including those that have been imported, and if a certain input is produced using two different technologies, it should actually have two separate entries. Meanwhile, vector \( d \)

\(^{10}\) Wassily Leontief was awarded the 1973 Nobel Prize for the development of IO analysis. Leontief (1966, 1986) covers much of his work. Leontief (1970) himself introduced pollution by augmenting the technology matrix to include a row of pollution generation coefficients; see Qayum (1994) for a consistent reformulation. For IO analysis in general, see Miller and Blair (2022), and specifically Chapter 10 for environmental IO analysis.
collects the direct emissions that characterise all of these processes. Except for simple cases, this cannot be dealt with at the micro level. However, sector-level approximations of factor intensities using IO models are feasible. And just as the price mechanism is able to process an enormous amount of information in a decentralised way, there are ways to make the coordinated exchange of information between producers do the rest of the work.
### Annex 2: GHG content for product groups, Germany 2018

<table>
<thead>
<tr>
<th>Product groups</th>
<th>Group-level GHG content (CO₂ equivalents, kg/€)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Electricity, distribution services for electricity and air conditioning</td>
<td>4.210</td>
</tr>
<tr>
<td>Water transport services</td>
<td></td>
</tr>
<tr>
<td>Products of agriculture, hunting, and related services</td>
<td></td>
</tr>
<tr>
<td>Air transport services</td>
<td></td>
</tr>
<tr>
<td>Basic iron and steel, products of the first processing of steel</td>
<td></td>
</tr>
<tr>
<td>Coke and refined petroleum products</td>
<td></td>
</tr>
<tr>
<td>Ceramic products, processed stone and clay</td>
<td></td>
</tr>
<tr>
<td>Coal and lignite</td>
<td></td>
</tr>
<tr>
<td>Crude petroleum and natural gas</td>
<td></td>
</tr>
<tr>
<td>Glass and glass products</td>
<td></td>
</tr>
<tr>
<td>Paper and paper products</td>
<td></td>
</tr>
<tr>
<td>Food, beverages, tobacco products</td>
<td></td>
</tr>
<tr>
<td>Basic precious and other non-ferrous metals</td>
<td></td>
</tr>
<tr>
<td>Manufacturing gas; distribution spheres of gaseous fuels through mains</td>
<td></td>
</tr>
<tr>
<td>Wood and of products of wood and cork, except furniture</td>
<td></td>
</tr>
<tr>
<td>Sewageage, waste, material recovery, remediation services</td>
<td></td>
</tr>
<tr>
<td>Metal ores, mining and quarrying products, support services</td>
<td></td>
</tr>
<tr>
<td>Natural water; water treatment and supply services</td>
<td></td>
</tr>
<tr>
<td>Travel agency, tour operator and other transportation services and related...</td>
<td></td>
</tr>
<tr>
<td>Rubber and plastic products</td>
<td></td>
</tr>
<tr>
<td>Land transport services and transport services via pipelines</td>
<td></td>
</tr>
<tr>
<td>Fabricated metal products, except machinery and equipment</td>
<td></td>
</tr>
<tr>
<td>Textiles, wearing apparel, leather and related products</td>
<td></td>
</tr>
<tr>
<td>Constructions and construction works for civil engineering</td>
<td></td>
</tr>
<tr>
<td>Buildings, and building construction works</td>
<td></td>
</tr>
<tr>
<td>Printing and recording services</td>
<td></td>
</tr>
<tr>
<td>Accommodation, food and beverage serving services</td>
<td></td>
</tr>
<tr>
<td>Postal and courier services</td>
<td></td>
</tr>
<tr>
<td>Products of forestry, logging and related services</td>
<td></td>
</tr>
<tr>
<td>Warehousing and support services for transportation</td>
<td></td>
</tr>
<tr>
<td>Pharmaceutical products and preparations</td>
<td></td>
</tr>
<tr>
<td>Motor vehicles, trailers and semi-trailers</td>
<td></td>
</tr>
<tr>
<td>Machinery and equipment n.e.c.</td>
<td></td>
</tr>
<tr>
<td>Other transport equipment</td>
<td></td>
</tr>
<tr>
<td>Repair and installation services of machinery and equipment</td>
<td></td>
</tr>
<tr>
<td>Specialised construction works</td>
<td></td>
</tr>
<tr>
<td>Furniture and other manufactured goods</td>
<td></td>
</tr>
<tr>
<td>Electrical equipment</td>
<td></td>
</tr>
<tr>
<td>Retail trade services, except of motor vehicles and motorcycles</td>
<td></td>
</tr>
<tr>
<td>Computer, electronic and optical products</td>
<td></td>
</tr>
<tr>
<td>Other personal services</td>
<td></td>
</tr>
<tr>
<td>Wholesale trade services, except of motor vehicles and motorcycles</td>
<td></td>
</tr>
<tr>
<td>Repair services of computers and personal and household goods</td>
<td></td>
</tr>
<tr>
<td>Sporting services and amusement and recreation services</td>
<td></td>
</tr>
<tr>
<td>Wholesale and retail trade services of motor vehicles and...</td>
<td></td>
</tr>
<tr>
<td>Telecommunications services</td>
<td></td>
</tr>
<tr>
<td>Residential care and social work services</td>
<td></td>
</tr>
<tr>
<td>Other professional, scientific, technical and veterinary services</td>
<td></td>
</tr>
<tr>
<td>Arts, culture, entertainment and gambling</td>
<td></td>
</tr>
<tr>
<td>Public administration and defence services</td>
<td></td>
</tr>
<tr>
<td>Human health services</td>
<td></td>
</tr>
<tr>
<td>Security, buildings and landscape, office administrative and support services</td>
<td></td>
</tr>
<tr>
<td>Publishing services</td>
<td></td>
</tr>
<tr>
<td>Services furnished by membership organizations</td>
<td></td>
</tr>
<tr>
<td>Scientific research and development services</td>
<td></td>
</tr>
<tr>
<td>Film, video, TV production, sound recording, music publishing; broadcasting</td>
<td></td>
</tr>
<tr>
<td>Advertising and market research services</td>
<td></td>
</tr>
<tr>
<td>Education services</td>
<td></td>
</tr>
<tr>
<td>Architectural and engineering services; technical testing and analysis...</td>
<td></td>
</tr>
<tr>
<td>Insurance and pension funding services, except compulsory social security</td>
<td></td>
</tr>
<tr>
<td>Legal and accounting services; services of head offices; management...</td>
<td></td>
</tr>
<tr>
<td>Financial services, except insurance and pension funding</td>
<td></td>
</tr>
<tr>
<td>Services auxiliary to financial services and insurance services</td>
<td></td>
</tr>
<tr>
<td>Compulsory social security services</td>
<td></td>
</tr>
<tr>
<td>Real estate services</td>
<td></td>
</tr>
<tr>
<td>Employment services</td>
<td></td>
</tr>
<tr>
<td>IT and information services</td>
<td></td>
</tr>
<tr>
<td>Rental and leasing services</td>
<td></td>
</tr>
</tbody>
</table>
Product level greenhouse gas contents – how to get there?

Pulling ourselves up by our bootstraps

Dr. Ulf von Kalckreuth, DG Statistics, Centre of Excellence, Deutsche Bundesbank

11th Biennial IFC Conference
“Post-pandemic landscape for central bank statistics”
BIS, Basel, 25-26 August 2022,
GHG value – the vision

- This talk is on mitigation. For decision making with an eye to emissions, we need **product level carbon content information**!

- **Deutsche Bundesbank Discussion Paper 23/2022**

- **GHG value**: direct and indirect GHG emissions on the product level

- **Implicit definition**: GHG value depends on direct emissions, inputs and their GHG value!

- GHG use is indicated **on every stage of production and passed on**, like a second price tag!

- **Producers, investors, consumers and political authorities** can have the information needed for decision making. Competition among producers may induce rapid adjustment!

- Carbon taxes work on the supply side, from the beginning to the end of the supply chain. The GHG value works on the demand side, from the end to the beginning of the value chain!

With GHG values for all products, the system can work smoothly. But for most goods, there aren’t. **How do we get there?**
Consider the vector of inputs of product $k$, with $a_{k,i}$ being the quantity of good $i$ embodied in the production process of one unit:

$$a_k = (a_{k1} \ a_{k2} \ ... \ a_{kK})'$$

Let $d_k$ be the amount of GHG directly emitted and $g_i$ be the GHG value of input $i$. Then the GHG value of good $k$ is given as the sum of direct and indirect emissions:

$$g_k = d_k + g'a_k = d_k + \sum_i g_i a_{ki}$$

If the $g_i$ are known, we can calculate the GHG value of product $k$ directly, based on our knowledge of direct emissions and technology.
The reduced form

If the \( g_i \) are unknown, the GHG value is still defined. The equation is recursive. Eq. (1) is an IO model for production. We can solve for the GHG values of all products simultaneously. Let

\[
A = (a_1 \ a_2 \ ... \ a_K)
\]

be the matrix of the Input coefficients for all produced goods. With \( d \) the vector of direct emissions for products 1,\ldots, \( K \), we may write:

\[
g' = d' + g'A
\]

and solving for \( c \) yields

\[
g' = d'(I - A)^{-1}
\]

---

GHG values of all goods  
Direct emissions for all goods  
Leontief inverse, reflecting production interlinkages  

Sectoral level: this can be calculated from existing data.  
Micro level: we do not need to compute this solution. Let decentralised information processing do the work!
(2) Accounting and sparse micro level communication

**Robert Kaplan (Harvard) and Karthik Ramanna (Oxford):** Harvard Business Review
November/December 2021

K&R propose to treat emissions as a liability -- **E-liability** -- that is moved forward with input supply along the supply chain and allocated over products. E-liabilities can be a measure of total (direct and indirect) GHG content – conceptually, emissions are collected over the value chain.

**Introduce a sparse information flow** from input providers to producers! The relevant information is revealed, but neither inputs nor technology!

**Standard techniques and routines** can be used to process E-liability information. Allocation of emissions to products largely unconstrained, left to producer.

To address the **issue of missing input information, circular value chains** and to ensure **comparability**, the linear structure outlined above is extremely useful.
The key result

Producers do not need to know the GHG values of the entire economy, only those of their own providers, or estimates thereof!

Information processing very effectively on a decentral basis. Iterating equation (1) will lead to the correct GHG values!

Operational version of the definition equation

\[ \tilde{g}_k = d_k + \tilde{g}'a_k \]

Input coefficients, known by producers

Direct emissions, known by producers

GHGVs stated by input providers or estimates as initial values

Producers' GHGV estimate

This is shown both analytically and by simulation, based on production interactions in Germany.

Jumpstart the system and boot it bottom up! Hayek and the „man on the spot“!
Simulation – market learning of the GHG value system

Simulating GHGVs for 7699 products

Evolution of mean absolute error

GHG values based on IO information for Germany
A simulation experiment based on 71 product groups

Convergence of GHG value measures by product group
Mean absolute error relative to true GHG values -- informative prior

The first step (use of private information on input structure) is the most important
Signalling and a role for institutions

Signalling

There is a path that leads to voluntary disclosure by (almost) all firms:

Disclosure is a signalling game

• Producers with low GHG value will have an incentive to disclose, as they can charge higher prices.
• Signal value for the decision not to disclose
• Reinforced by disseminating disclosed GHG values on a central data platform
• Reinforced further by calculating sector averages conditional on not disclosing

Auditing and an information platform

Auditing to make sure that the GHG values are a fair estimate

Centralised platforms to make available the existing information
  • on industry averages
  • on GHG values on a product and on the company level, if available

Compute estimated carbon content for firms of a given industry that do not disclose their GHG values, from the known industry averages and the known GHG values of the firms that do disclose.

Strong incentive for disclosure!
Policy options for central banks and international organisations

The GHG value is a **decentral information system**, but it needs **institutional support**, eg from central banks and international organisations

1. Co-operate with Statistical Institutes in setting up a **rather disaggregated IO-models**. This will **give us useful group level GHGVs immediately** and serve as a basis for product level GHGVs. Already suggested for the **Data Gaps Initiative within the G20**.

2. Set up and maintain a **dissemination platform for GHG value data** on the level of sectors, enterprises and products (eg with CPA classification system as a basis)

3. Support development of **disclosure standards**, as a basis for comparability and auditing. For direct emissions, those rules can build on the relevant GHG Protocol standards.

4. Interact with supervisory authorities and the IFRS on **disclosure and auditing requirements**, eg regarding the CSRD. Possible disclosure requirements should target large companies, as well as producers of primary goods and importers.

In addition, **support for ongoing field studies** may be very useful!
A characterisation of financial assets based on their cash-flow structure\(^1\)

Celestino Girón,
European Central Bank

---

\(^1\) This presentation was prepared for the conference. The views expressed are those of the author and do not necessarily reflect the views of the BIS, the IFC or the central banks and other institutions represented at the event.
A characterisation of financial assets based on their cash-flow structure

Celestino Giron

Abstract

In the International Economic Accounting Standards, the classification criteria for financial assets aim to bring out the economic substance underlying the specific arrangements adopted to establish the links between borrowers and lenders. Differences in the cash flows associated to the assets, like how the payments are determined or their distribution over time, are key to their economic effects. While these cash flow differences are implicitly considered by the Standards, they are only sometimes used explicitly for asset classification prescriptions. In this paper we provide a taxonomy of cash-flow structures and a formal characterisation of assets on that basis. This taxonomic discussion can inform the current process of review of the Standards and, beyond that, contribute to the classification of assets in practice.

Keywords: International Economic Accounting Statistical Standards, 2008 System of National Accounts, Financial accounts, cash-flows, vectorial spaces

JEL classification: E01, G23, C18

Contents

1. Introduction ....................................................................................................................................... 2
2. Financial assets in macroeconomic statistics ........................................................................ 3
3. Cash-flow taxonomy....................................................................................................................... 4
   Cash flow vector space .................................................................................................................. 7
4. Cash-flows and the classification of financial assets in macroeconomic statistics ......................................................................................................................... 8
5. Conclusion ........................................................................................................................................ 10
1. Introduction

The statistical classification of complex financial assets is sometimes subject of controversy among macroeconomic statisticians. Thus, it is not always clear for instance whether a specific asset should be treated as a loan or as other account payable, or whether it is a forward derivative or an option.

These difficulties are not due to the absence of appropriate classification criteria in the statistical standards, but rather to the fact that those criteria are based on institutional features that are not always easy to identify. Moreover, growing complexity in the design of financial assets is progressively disconnecting the institutional features from the underlying economic substance, which is the fundamental concept lying behind the classification criteria.

A fundamental part of the economic substance behind an asset lies in the characteristics of the cash flows associated to it, including their distribution across time. They determine the specific role of the asset in order to carry value forward, i.e. to act as a store of value, and the nature of the liquidity, maturity and risk services embedded in them; these features explain the asset price. However, the statistical standards pay little attention to the cash flows and their time profile to characterise financial assets, beyond broad distinctions such as between debt and equity instruments depending on whether a predetermined stream of payments is present or not.

This note presents a taxonomy of cash flows aimed to assist the classification of financial assets in macroeconomic statistics. Under the approach in this note a financial asset is a collection of cash flows, an array of payments associated to discrete moments in time (although a generalisation as continuous payments can be made). Algebraically, they are then vectors in an infinite multidimensional space given by time. The individual payments consist in the provision of goods and services, or other financial assets, i.e. other cash flows. The asset stock, the value of the cash flow, is, in this algebraical presentation, a function defined on the arrays containing the cash payments.

The taxonomy is not intended to replace the rich classification criteria available in the standards, but to complement it. It can help decide on the classification of borderline cases by checking which cash flow taxonomical category the asset belongs to, or approximates the most. At the same time, the taxonomy can be used to define the assets in the standards themselves and thus contribute to clarity and certainty in the methodological prescriptions.

Section 2 briefly summarises the classification criteria available in the standards, Section 3 is the core of the paper and presents the methodological approach taken and the taxonomy proposed. Section 4 illustrates the application of the taxonomy to some specific difficult classification issues identified in the context of the review of the Economic Accounting Statistical Standards. Section 5 concludes.
2. Financial assets in macroeconomic statistics

The 2008 edition of the System of National Accounts (2008 SNA) defines assets as economic objects which carry value forward from one period to the next i.e. (act as store of value) and on which property rights can be exercised. Financial assets can be defined as the subset of assets characterised by the fact that they have been specifically designed to be store of value, by virtue of a legal contract or a social consensual agreement, as opposed to non-financial assets, which act as store of value not by design, but as a side effect of their main role in production processes (fixed assets), rent source (non-produced assets) or artistic or sentimental motives (valuables).

2008 SNA enumerates the categories of assets that would be classified as financial assets. These are “financial claims, shares or other equity in corporations plus gold bullion held by monetary authorities as a reserve asset” (2008 SNA 3.36). Most financial assets are “financial claims”, which are economic relations between economic agents by virtue of which one party is obliged to provide payments to the other (2008 SNA 3.33 to 3.35). Equity consists in rights on the residual value of corporations and do not entail any obligation of predetermined payments from the corporation to the holders (2008 SNA 11.81). Finally, gold is considered as a financial asset under certain circumstances and its role as store of value is backed only by an implicit consensual contract.

Financial claims are in turned broken down into various categories on the basis of institutional features, including their negotiability and the classification of the unit for which the claim represent a liability. The main categories of financial assets envisaged, including those that are not financial claims, are Monetary gold and SDRs, Currency and deposits, Debt securities, Loans, Equity and investment fund shares, Insurance, pension and standardized guarantee schemes, Financial derivatives and employee stock options and Other accounts receivable. These categories are further broken down in sub-categories.

Individual assets are placed into one category or the other on the basis of their economic substance, rather than their legal form, and how this substance aligns with the corresponding classification categories. Thus, for instance repurchase agreements and securities lending with cash collateral are seen as the same kind of claims in substance irrespective of their formal differences and classified in both cases as deposits (if liabilities of a deposit taking corporation) or loans (otherwise).

Moreover, the emphasis if put on the economic effects, rather than on the economic purposes. For instance, the classification does not distinguish between financial derivatives used for hedging from those motivated by trading, completely disregarding the purpose of the deal. An exception to the economic effect principle is the functional classification of cross-border transactions in the Balance of Payment, where the focus is turned to the economic purposes.

At the same time, the classification of individual assets into the macroeconomic categories is not always without difficulties. A recent example of this is the debate around the classification of balances arising in cash collateral agreements, where the

global statistical community is still struggling to agree on the appropriate
classification for them, either as deposits, loans or other accounts receivable\(^3\).

3. Cash-flow taxonomy

We define a financial asset at time \( t \), \( a^t \) as an array of future cash payments \( a^t = \{a^t_{i+1}\}, i = 0, 1, 2, \ldots \) where \( i \) is a temporal index pointing at future payments amounting \( a^t_{i+1} \). As time passes, i.e. \( t \) increases, past flows are disappearing from the array. For financial claims, any \( a^t \) is accompanied by another array \( l^t \) such as \( l^t = -a^t ; l^t_{i+1} = -a^t_{i+1} \) representing the liability view of the claim. When \( t \) is evaluated at the time of claim inception, \( l^t_{i+0} = -a^t_{i+0} \) are the funds transferred to the debtor by the creditor when the claim is initiated.

We classify cash flow arrays according to their temporal profile and whether the payments \( a^t_{i+1} \) are prefixed or not. This would give us the following broad categories:

**General financial claims**

These cash flow streams present a flow at inception representing a payment from the creditor to the debtor when the claim is initiated (for instance the funds raised in the issuance of a debt security) and several payments from the creditor to the debtor in the subsequent (periods covering coupon payments and redemption of principal).

All flows are predetermined. These are the sign restrictions:

\[ a^t_{i+0} < 0 ; a^t_{i+1} \geq 0 \forall i > 0 . \]

Figure 1 shows the profile of one such assets with 7 cash payments and another with an infinite series of payments.

The profile of 1.2 correspond to loans, non-transferable deposits and debt securities (other than zero-coupon debt). Figure 1.2 is the profile of perpetual debt.

**Dual payments**

A particular case of the category above is one where only two payments exist, one at inception and one at redemption. Depending on whether the initial payments or equal or not to the one at the end, this would be the profile of a zero-coupon bond

A characterisation of financial assets based on their cash-flow structure

(payment at the end larger than at inception, Figure 2.1 for a 7-period asset) or of other payables (Figure 2.2, showing a short-term payable)

<table>
<thead>
<tr>
<th>Dual payments</th>
<th>Figure 2</th>
</tr>
</thead>
<tbody>
<tr>
<td>Figure 2.1</td>
<td>Figure 2.2</td>
</tr>
<tr>
<td><img src="image1.png" alt="Diagram" /></td>
<td><img src="image2.png" alt="Diagram" /></td>
</tr>
</tbody>
</table>

**Equity**

Equity cash flows have the same sign restrictions as financial claims (\( a_{t+0} < 0; a_{t+i} \geq 0 \forall i > 0 \)), but the amount of the flows is not predetermined. Similar as for perpetual debt, the time horizon is in principle infinite, but with the prospect of having here at some point a higher flow that would terminate the payment stream. Figure 3 shows this marking the stochastic nature of the flows in green.

<table>
<thead>
<tr>
<th>Equity</th>
<th>Figure 3</th>
</tr>
</thead>
<tbody>
<tr>
<td><img src="image3.png" alt="Diagram" /></td>
<td></td>
</tr>
</tbody>
</table>

This is the profile of a corporate stock, the flows after inception representing dividend payments, and the final payment being the corporate residual value (which could be zero, but never negative).

**Forwards**

This cash flow is characterised by no flow at inception and stochastic flows, which can be positive or negative, over the rest of its life. The sign restrictions only apply to the first flow then (\( a_{t+0} = 0 \)). Figure 4.1. shows the corresponding profile for a 7-period stream assuming that the stochastic payments take place at any time during those periods, but it can also be the case that single payment takes place at the end of the cash-flow life.
We find this cash-flow profile in forward contracts, like swaps or other futures (which can lead to periodic payments as in 4.1 if there is periodic settlement, or to a single payment).

**Options**

As opposed to forwards, here a negative payment is made at inception and a stochastic non-negative payment is made at the end of the cash stream. The sign restrictions are then as follows: \( a_{t+0} < 0 \); \( a_{t+k}^i \geq 0 \) for \( a \); \( a_{t+i}^i = 0 \) \( \forall i \neq 0, k \). Figure 4.2 shows the corresponding profile.

Naturally, this is the cash-flow presented by call and put options, the initial flow corresponding to the premium and the last payment to the gain obtained for exercising the option (which can never be negative as the holder can decide not to exercise the option if the difference between market and strike price lead to a loss).

Actual cash flows might fall under one of the categories above or have a mix nature having features of some of them. At the same time, this approach allows to express assets (cash flows here) as linear combination of other cash flows, which enables complex, non-standard assets to be expressed as combinations of assets falling under the categories above (see a more formal discussion in the Box on the cash flow vectoral space). In fact, the categories above can in same cases be expressed on the basis of some of the other categories proposed. For instance, the generic financial claim case in Figure 1.1 can be split into the sum of four zero-coupon cash-flows as in 2.1. Likewise, the profile of equity (Figure 3) can be seen as closely related to that of an option (Figure 4.2). Section 4 elaborates further on the latter.

The concept of asset stock in this context can be formally introduced by defining a function on the cash flow arrays that would convert them into scalars. This function can be seen as being the traditional “net present value” using a discount rate that is given by the market, in case this exist, which would be a market of cash flows. In formal terms: \( s(a^i) = npv (\{a_{t+i}^i\}, r) = \sum_i a_{t+i}^i (1+r)^t \), \( s() \) denoting stock and \( r \) being the discount rate\(^4\). When evaluated at inception, \( s() \) yields zero by construction for all cases presented above (unless the flows are unbalanced in terms of market.

\(^4\) A generalization can be made by having a different discount rate \( r \) for each of the payments. When the payments are uncertain \( a_{t+i}^i \) must be understood as the mathematical expectation of the payment.
perception and the whole cash-flow has an initial value; this is known in the market practice as being “out of the money”, as opposed to being “in the money”). As time elapses and in particular the initial payment is dropped from the subsequent arrays, the market value and the stock valuation are different from zero.

**Cash**

A last category of cash flow is a degenerate case that presents just a single certain payment in any future moment $i$ as decided by the creditor. As opposed to all other categories presented above, the stock function does not evaluate to zero at inception, but to the value of that one and only flow, as by convention the flow is discounted as if it would take place at $i=0$.

To this category belong assets without liabilities, like monetary gold, or to transferable assets like banknotes, coins or transferable deposits.

The individual payments forming generic cash streams are typically assets of this kind (cash streams of a single flow). The payments can also consist in the provision of goods and services, or of non-financial assets. Finally, the individual flows in the cash streams can also be generic financial assets, i.e. generic cash flows entailing more than one cash payment. This evokes the inspiring image that financial assets are payment streams in which each payment is itself a payment stream, and so on in infinite recursion.

---

**Cash flow vector space**

The universe of cash flows has the algebraical structure of an infinite-dimensional vector space over the field of the Real numbers. In this box, we apply the standard vector space concepts to this specific vector space.

The two binary operations of the vector space are defined as follows:

\[
\begin{align*}
  a^t + b^t &= \{a^t_{i+t} + b^t_{i+t}\}, i = 0, 1, 2 \ldots; \quad a^t = \{a^t_{i+t}\}, b^t = \{b^t_{i+t}\} \\
  a a^t &= \{aa^t_{i+t}\}, i = 0, 1, 2 \ldots; \quad a^t = \{a^t_{i+t}\}, a \text{ a real number}
\end{align*}
\]

Linear combinations of assets/cash flows applying the definitions above are also assets $c^t = \alpha a^t + \beta b^t$ in the vector space, which formalises the statement in the man text that cash flows can be expressed as combinations of other cash flows.

Given a group of assets, the set of all assets that can be constructed by means of linear combinations of the group also has the structure of vector space, and constitutes a sub-space of the larger cash flow vector space. It is said that the group of assets spans the sub-space.

The converse is not true, and a given asset might not be able to be expressed as a combination of another set of assets (i.e. the former is not in the span of the later). More in general, a set of assets is said to be linearly independent if none of the assets in the set can be expressed as a linear combination of the others.

Interesting sets of assets are those whose components are linearly independent and span the whole cash-flow vector space. A set like this is called a basis and can be used to express any other asset. As the cash flows are generally arrays of infinite length, such basis should have infinite members, and the expression of a generic asset in the basis would be:

\[
a^t = \alpha_1 u^t_1 + \alpha_2 u^t_2 + \cdots + \alpha_n u^t_n + \cdots
\]

where $(u^t_1, u^t_2, \ldots, u^t_n, \ldots)$ is a basis and the set $(\alpha_1, \alpha_2, \ldots, \alpha_n, \ldots)$ are called the components of $a^t$ in the basis.

The so-called canonical basis is the set of assets $c^t_i = \{c^t_{i+x}\}, c^t_{i+j} = 1$ $c^t_{i+j} = 0 \forall i \neq j$ and the components of any asset $a^t$ in that basis are nothing but the individual cash payments $a^t_{i+t}$.
More interesting are basis other than the canonical one. Beyond the trivial representations resulting from the canonical basis, assets can be expressed as combinations of proper assets included in generic basis. An interesting line of research is the construction of a basis of options (as defined in Section 3), so that we can conclude that options are the building blocks of any other financial asset (see a related discussion concerning equity in Section 4).

Although the generic cash flow vector space is infinite-dimensional, interesting sub-spaces cover assets of finite payment streams. Those sub-spaces are spanned by finite basis, and the corresponding assets are represented with finite components sets. The change of representation from one finite basis into another one can be done by constructing a matrix formed by columns having the components in the old basis of the assets in the new basis; the components of an asset in the new basis results from pre-multiplying the components of the asset in the old basis by the inverse of the matrix above (this is the so-called change of basis matrix).

4. Cash-flows and the classification of financial assets in macroeconomic statistics

The taxonomy in Section 3 constitutes an instrument for the classification of financial assets in macroeconomic statistics. It can serve to improve the definitions in the statistical standards and to resolve borderline cases when other criteria do not suffice to decide on classification. Moreover, it can inform alternative classification structures.

An example of a context where the taxonomy might be of use for deciding on the classification of a category of assets is the discussion on the classification of Credit Default Swaps (CDS) that has taken place in the context of the review of the statistical standards. A CDS is a contract between two parties whereby one of the parties (the investor or the protection buyer) to buy protection against the risk of a credit event (typically bankruptcy) on a specified entity. The protection buyer pays a series of fees (premiums) to the protection seller, and, if the credit event takes place, the protection seller pays a compensation to the buyer for the loss. The premium payments are called “the premium leg”, and the compensation payment “the contingent leg”.

The controversy on the CDS classification comes from the fact that some of the characteristics of CDS are those of a forward contract (market value switching from positive to negative and vice versa, multiple payments along the life of the contract) and some others are those of an option (payment at inception, presence of a contingent non-negative payment in any case), which makes it difficult to decide whether they belong to one category or the other.

A close examination of the cash flows of a CDS against the background of the categories put forward in Section 3 can help take a decision. A CDS cash flow presents a series of fixed payments (the premium leg) and a contingent inflow only if the credit event takes place (which is never negative). The sign restrictions are then $\delta_{t+k} \geq 0$.

A characterisation of financial assets based on their cash-flow structure

0 for a k; a_{i+k}^* ≤ 0 ∀i ≠ k. Panel 5.1 shows this cash flow associated to a CDS contract and compares it with the categories forward and option of the taxonomy.

The CDS cash stream is similar to that of an option and has little in common with that of a forward; in particular, it doesn’t present the flipping contingent cash payments during the contract life. The sign restrictions are not those of a forward. Actually, one can even say that the option cash stream is a particular case of the CDS one, where the premium leg has a single payment at inception in the case of an option. This analysis would support a classification of CDS together with generic options.

While the cash flow of a CDS contract can be seen as a generalization of the option cash flow where the premium leg has more than one payment, the equity cash flow (Figure 3) can be seen as well as a generalization of the option cash flow where

---

6 One of the arguments put forward to support the classification of CDS as forwards is the fact that the market value of the CDS can change signs, while that of an option is always non-negative. However, this happens only due to the relatively unimportant difference that the premium leg of an option has a single payment, while that of a CDS has many payments. We can divide the cash streams of both options and CDS into two sub-streams, one with the premium leg and another with the contingent payment, and the value of the total stream would be the sum of the value of the two sub-streams (see Box for a formal discussion on decomposition of cash flows into sub-streams). The value of the contingent leg is always non-negative during the lifetime of the contract both for options and CDS. However, the value of the premium leg is zero after the initial payment for options (because no other premium payment is made), and always negative for CDS. Depending on whether the value of the contingent leg, which changes across time depending on the changes in the market perception of the likelihood of the credit event is higher or lower than that of the premium leg (in absolute value), the overall value of the CDS will be positive or negative. For options, being the premium leg zero, the value of the total cash stream will always be non-negative.
the contingent leg has more than a single payment at the end of the cash flow.
Another way to say the same thing: the equity cash flow in case there is no interim payment (no dividend payment) is indistinguishable from an option cash flow. This similarity between options and equity is reminiscent to the well-known statement by Merton (1974) that equity of a levered company can be seen as a call option on the assets of the corporation, where the strike price is the value of its debt. This would suggest an alternative classification hierarchy where options and equity are placed next to each other.

Another issue where the analysis of the associated cash flow gives some insights on the statistical classification is that of Crypto Assets without liabilities designed as a general medium of payment (CAWLM). While the core of the discussion on these assets, which include bitcoin and similar cryptocurrencies, is whether they are financial or non-financial, the fact that they present a “cash” payment stream profile would suggest a classification closer to the assets having that profile (monetary gold, coins, banknotes and transferable deposits) should they be finally classified as financial assets.

5. Conclusion

The cash flow properties of the financial assets embed most of their economic substance, to the extent that it is analytically useful to look at financial assets just as such cash streams. The tentative taxonomy presented in this paper puts forward a few broad categories that can work as a classification system. We have also suggested ways to look at assets as linear combination of other assets and therefore a way for breaking them down into building blocks that can fit into the categories proposed.

While thinking about assets as cash flows is very common in many analytical fields, such as corporate structure theory, the International Economic Accounting Standards do not often make explicit use of cash-flow features to provide classification guidance. It would add certainty and clarity to the classification criteria if some reasoning such as the one followed in this paper is considered in preparing new guidance. This would also facilitate the use of macroeconomic statistics for certain analytical purposes.

Moreover, and irrespective of the explicit inclusion of a cash flow approach in the statistical standards or not, the analysis of the cash flows can always provide useful insights to decide on the classification of borderline cases.

At the same time, some work is needed to further formalise the approach suggested in this paper. This applies in particular to the mathematical underpinning sketched in Box 1, but also to the granularity of the taxa forming the taxonomy.

A characterization of financial assets based on their cash-flow structure

Celestino Girón

11th Biennial IFC Conference on Post-pandemic landscape for central bank statistics

26 August 222
What is this paper proposing?

✓ taxonomy of cash-flow structures
✓ characterization of financial assets on that basis …

to provide new classification criteria and contribute to the resolution of borderline cases
Designed to carry forward economic value

Encompasses:
✓ Financial claims
✓ Equity
✓ Financial Gold

- Currency and deposits
- Debt securities
- Loans
- Insurance technical reserves
- Financial derivatives
- Other accounts

Individual assets classified on the basis of economic substance (no legal form) and economic effects (no economic purpose)

No (much) explicit mentioning to cash flow characteristics
Assets seen as cash flows

Arrays of expected cash flows

\[ a^t = \{a_{t+i}^t\}, \ i = 0, 1, 2 \ldots \]

Also arrays of expected cash flows

\[ a_{t+i}^t, \text{ expected flow in } t+i \]

Assets can be expressed as linear combinations of other (linearly independent) assets

\[ a^t = \alpha_1 u_1^t + \alpha_2 u_2^t + \cdots + \alpha_n u_n^t + \cdots \]

For financial claims, a “mirror” cash flow exists (a liability)

\[ l^t = -a^t; \ l_{t+i}^t = -a_{t+i}^t \]
Rubric

Taxons

Generic claim

Zero coupon

Perpetual debt

Equity

Certain flow

contingent flow
The case of CDS

Are CDS \textbf{forwards} (market value switching from positive to negative, multiple payments along the life of the contract) or \textbf{options} (payment at inception, presence of a contingent non-negative payment in any case)?
Are options and equity the same kind of assets?

Indeed, these two asset-liability configurations reflect the same economic substance, the same corporate structure.
✓ Cash flow structures give relevant information on the economic substance of financial assets

✓ A taxonomy of cash flows can give answers to borderline classification problems and inform alternative asset classification systems

✓ Further research is needed to formalise the mathematical structure of cash flows (vector space), including its extension to the continuous case
A decision-making rule to detect insufficient data quality – an application of statistical learning techniques to the non-performing loans banking data

Paolo Cimbali, Marco De Leonardis, Alessio Fiume, Barbara La Ganga, Luciana Meoli and Marco Orlandi,
Bank of Italy

---

1 This presentation was prepared for the conference. The views expressed are those of the authors and do not necessarily reflect the views of the BIS, the IFC or the central banks and other institutions represented at the event.
A decision-making rule to detect insufficient data quality

an application of statistical learning techniques to the non-performing loans banking data

Barbara La Ganga, Paolo Cimbali, Marco De Leonardis, Alessio Fiume, Luciana Meoli and Marco Orlandi

Abstract

The data quality level generally follows an improving trend thanks to subsequent corrections submitted by reporting agents; however, a data quality worsening may occur when data production is affected by exogenous and unpredictable events, such as the pandemic or changes in the reporting requirements. Using banks Non-performing loans regulatory requirements, the paper proposes a decision-making rule to improve the detection of these cases by defining a synthetic data quality indicator based on past evidence from data quality management activity: outliers, their severity level and received confirmations of underlying data, the latter considered for estimating the expected confirmations through a logistic regression model.

Keywords: outliers, non-performing loans, data quality, supervised machine learning, logistic regression.

JEL classification: C18, C81, G21.

Contents

1. Introduction ................................................................. 2
2. From data collection to the release of the information .............................. 3
3. The NPL dataset and model selection ............................................. 6
4. Main results for the selected model .............................................. 8
   4.1 Logistic regression: estimation of ‘remark confirmed, Yes/No’ .......... 8
   4.2 Application of the decision-making rule ..................................... 9
Conclusions ............................................................................. 11
References .................................................................................. 12
1. Introduction

The Bank of Italy collects a large array of statistical and supervisory data from banks and other financial institutions on a regular basis to support its analyses and policy decisions. The data are organized in datasets, also called reports, each obeying different reporting regulations. Assessing the Data Quality Level of reports (DQL) is key to enabling users to carry out thorough and robust analyses.

Reporting Agents (RAs) are required to ensure high-quality data. Errors need to be promptly corrected by the RAs with a new report submission, whereas the unusual, yet correct, data have to be confirmed by the RAs. A Data Manager is in charge of evaluating the received confirmations and establishing a dialogue with the RAs in order to collect further information on unusual data patterns. This process creates a data quality cycle where RAs are required to submit new reports as long as the DQL is not sufficient (Casa et al., 2022).

Through subsequent data submissions, DQL is expected to improve over time; however, the data reliability may be impaired by outliers that are actual errors. A worsening in the DQL could occur especially when data production is affected by exogenous and unpredictable events. These events include RAs’ IT malfunctions in the production of statistical reports, errors occurring when applying changes in the reporting requirements, or extraordinary events, such as the pandemic, that can affect the organisation of the RAs (Schnabel, 2020; Tissot and De Beer, 2020). Prompt detection of a substantial DQL worsening is key for reducing costs in the Data Quality Management (DQM) process and getting fit-for-use data.

In literature on data quality, this topic is typically approached by monitoring some stand-alone dimensions, such as timeliness, accuracy and consistency, all evaluated through metrics and indicators (Damia and Aguilar, 2006; Pipino et al., 2002). In the international statistical context, some Institutions have developed specific strategies and frameworks (IMF Data Quality Assessment Framework, 2012; ESS QAF European Commission and Eurostat, 2019). A peculiar approach to DQM is envisaged by the application of the Benford’s law, also known as the ‘first digit law’. This approach allows identifying errors as elements that do not follow an empirical regularity in data describing naturally-occurring phenomena (Gonzalez-Garcia and Pastor, 2009).

In this paper, with reference to the accuracy, completeness and consistency of reports sent by RAs, we investigate the improvement of the DQL monitoring by defining a new decision-making rule which could support the Data Manager in assessing whether a revision by an RA of previously transmitted data improves or worsens the DQL. The rule is based on a synthetic data quality indicator computed through past evidence from the DQM activity: the outliers, their severity level and the received confirmations of underlying data. These three metrics are considered to evaluate the DQL; furthermore, since the confirmation status is unknown once a new data submission is received, the confirmations related to remarks just generated are

---

The authors are grateful to Gianluca Cubadda and Alessio Farcomeni (University of Tor Vergata, Rome), Laura Mellone, Francesca Monacelli and Roberto Sabbatini (Bank of Italy) for their useful comments and fruitful discussions on a preliminary draft of the paper. The views expressed herein are those of the authors and do not necessarily reflect those of the Bank of Italy.
A decision-making rule to detect insufficient data quality

estimated through statistical learning techniques using the information on the past DQM activity.

The methodology, applied to the Non-Performing Loans (NPL) dataset collected by the Bank of Italy, shows an improvement in the DQL monitoring process.

The paper is organized as follows. Section 2 describes the logical path that leads from the DQM process to the definition of the decision-making rule. Section 3 describes the data used in the empirical part of the study and illustrates the main arguments that guided the choice of the final model by comparing different alternatives. Section 4 shows the empirical results also validating the consistency of the decision rule with the data corrections that take place in practice.

2. From data collection to the release of the information

In the Bank of Italy the process spanning from the collection to the dissemination of monetary and financial data is structured in steps. As soon as a dataset is submitted by a RA, it undergoes the application of a set of automatic data quality checks (DQCs) to detect anomalies. When an outlier is identified, a notification (called ‘remark’) is generated and transmitted to the RA for its assessment and possibly for action to be taken.

Depending on the type of DQC, the generated remarks can request either a confirmation or a correction of the original data. DQCs that detect only reporting errors are defined as ‘non-confirmable’ since the data related to remarks generated (‘non-confirmable remarks’) cannot be confirmed. Each DQC is characterised by a ‘severity level’ defined in advance according to a scale increasing from 0 to 10 based on an a priori evaluation by data quality experts of the impact of errors on the quality of the underlying data. According to the ‘severity’ of DQCs, the generated remarks are also classified as ‘serious remarks’ and ‘non-serious remarks’, respectively. In particular, the ‘serious remarks’ are characterised by a severity level equal to 9 or 10; the ‘non-serious remarks’ by a severity level below 9.

RAs are required to assess outliers and to correct or confirm them by providing a suitable explanation for the underlying anomalous pattern. Based on the severity of the outstanding outlier as well as of the analysis of the explanations provided for the confirmations, the Data Manager may conclude that the DQL is still not adequate and therefore continue the dialogue with the RA open until the DQL reaches an overall satisfactory level. At each iteration, the Data Manager faces a trade-off: on the one hand, there is the need to make the information promptly available to users which implies that the above data quality cycle must be kept short; on the other hand, there is the need for the data to be as free as possible from significant errors. This assessment represents the core issue of the present study.

In detail, a RA submits the $k$th dataset related to a specific reference date $t$. After the first round of DQCs, the RA receives a set of remarks by the Data Manager. After checking the evidence in its internal systems, the RA corrects the erroneous data by submitting the $(k+1)^{th}$ report. In so doing, the revisions can either enhance or, in some extreme but still plausible cases, introduce a new error which worsens the DQL. Then, the Data Manager faces the situation illustrated in Graph 1.
A decision-making rule to detect insufficient data quality

<table>
<thead>
<tr>
<th>$k^{th}$ submission</th>
<th>(k+1)$^{th}$ submission</th>
</tr>
</thead>
<tbody>
<tr>
<td>Not-released</td>
<td>D</td>
</tr>
<tr>
<td>Released</td>
<td>A</td>
</tr>
<tr>
<td></td>
<td>Released</td>
</tr>
</tbody>
</table>

- **CASE A**: the $k^{th}$ data submission is considered fit-for-use since the automatic DQCs have generated only non-serious remarks and information can be disseminated to the users. With the (k+1)$^{th}$ data submission of the same dataset, the new data worsen the DQL because of the presence of serious remarks, which prevent data to be made available to users.

- **CASE B**: both $k^{th}$ and (k+1)$^{th}$ data submissions are considered fit-for-use because the automatic DQCs have generated only non-serious remarks. In this case, the latest data are always available to the users.

- **CASE C**: this event is the inverse situation of CASE A and it shows an improvement of DQL between the $k^{th}$ and (k+1)$^{th}$ data submissions. Only the latest data are released to the users.

- **CASE D**: both $k^{th}$ and (k+1)$^{th}$ data submissions are not released to the users because of the presence of serious remarks. The Data Manager and the RA carry on their interaction in order to resolve the data issues.

In cases A, C and D, the Data Manager’s decision is straightforward. Since in case B the (k+1)$^{th}$ data submission could contain some anomalies such that they reduce the DQL in comparison with the $k^{th}$ submission, it is key that the Data Manager has an instrument to early detect the worsening of the DQL from one report to the next.

Graph 2 depicts the DQM process in presence of a decision rule based on the change of the DQL.

---

**DQM - flow chart**

**AT EACH DATA SUBMISSION**

- CASE A/D
- CASE B
- CASE C

**END**

**NOT RELEASED**

**RELEASED**

---

Graph 1

Graph 2
In brief, the decision-making algorithm must mirror the DQL assessment made by the Data Manager when a subsequent data submission is transmitted by a specific RA and reference date. Then, we consider the DQL variation between two consecutive reporting submissions, the $k^{th}$ and the $(k+1)^{th}$, of the same dataset ($\Delta DQL_{k+1} = DQL_{k+1} - DQL_k$) when both submissions do not bear serious remarks and are in principle fit-for-use. Hence, the rule needs to take as inputs the number of remarks generated at each round of DQCs, the severity level for each of them and the number of confirmed remarks.

Let us define the dummy variable $R$ (remark) taking the value 1 if the DQC $c$, applied at the reference date $t$ for the $k^{th}$ data submission sent by the RA $p$, is violated and 0 otherwise

$$R_{t,p,c,k} = \begin{cases} 1, & \text{if } c \text{ is violated} \\ 0, & \text{if } c \text{ is satisfied} \end{cases}$$

and the dummy variable $Conf$ taking the value 1 if the remark $R$ is confirmed by the RA and 0 otherwise

$$Conft,p,c,k = \begin{cases} 1, & \text{if } R_{t,p,c,k} \text{ is confirmed} \\ 0, & \text{otherwise} \end{cases}$$

In case of non-confirmable remarks, $Conf$ is by construction equal to zero.

The synthetic data quality indicator for the $k^{th}$ data submission is defined as follows:

$$I_{t,p,k} = \sum_{c=1}^{C} \tau_c (R_{t,p,c,k} - Conft,p,c,k)$$

where $C$ denotes the number of remarks generated and $\tau_c$ represents the severity level of the DQC.

The higher the value of $I_{t,p,k}$, the lower the $DQL_k$. In particular, when the $k^{th}$ data submission is not impaired by remarks not confirmed, the indicator is equal to zero.

Our decision rule (1) is such that the release of the $(k+1)^{th}$ data submission takes place when

$$I_{t,p,k+1} \leq I_{t,p,k} \quad (1)$$

The decision rule (1) assumes that the evidence of a confirmed remark is already known. However, the submission of an explanation from an RA is subsequent to the generation of the related remark, then this status, named $Conf$, has to be forecasted through the estimation of the probability $p(Conf)$ that a remark is confirmed by a given RA, on a specific reference date:

$$Conft,p,c,k+1 = \begin{cases} 1, & \text{if } p(Conft,p,c,k+1) > \text{cut-off} \\ 0, & \text{otherwise} \end{cases}$$

where $\text{cut-off}$ is a threshold lying within $(0, 1)$. In order to obtain the estimated probability $p(Conf)$ machine learning techniques are applied to the available dataset including confirmable remarks actually observed in the previous reference dates; a cross-validation method allows to assess the cut-off level.
3. The NPL dataset and model selection

In this section we define the binary classification problem in order to predict whether a remark is confirmed by the RAs using the Italian NPL dataset.

The NPL dataset comprises detailed information on banking non-performing exposures and on the status of their credit recovery procedures transmitted by the parent company of a banking group that reports on a consolidated basis and by the individual banks on a stand-alone basis. Reporting follows a half-yearly periodicity with reference dates 31st December and 30th June (Banca d’Italia, 2016).

The dataset created for the study includes five reference dates from 30th June 2017 to 30th June 2019. In this time frame 445 RAs reported over 17 million of records assessed by a set of 37 automatic confirmable DQCs that generated 65,705 remarks.

The data used to estimate the probability of confirmation consider, as observations, all the individual confirmable remarks generated for each RA and reference date; more specifically, a remark is considered only once even if it is pending in more than one data submission. Remarks generated by non-confirmable DQCs are not used in the model estimation but are considered during the application of the decision-making rule.

As shown in Table 1, the total number of remarks detected in the given period amounts to 65,705, out of which 5,083 are ‘confirmable remarks’; in turn, 4,643 of the confirmable remarks, related to the reference dates from 30th June 2017 to 31st December 2018, are used for the model estimation (‘training set’) and 440, related to the reference date 30th June 2019, for its assessment (‘validation set’).

<table>
<thead>
<tr>
<th>Reference date</th>
<th>Number of remarks (total)</th>
<th>Number of confirmable remarks</th>
<th>Number of non-confirmable remarks</th>
</tr>
</thead>
<tbody>
<tr>
<td>2017-06-30</td>
<td>31,306</td>
<td>758</td>
<td>30,548</td>
</tr>
<tr>
<td>2017-12-31</td>
<td>8,679</td>
<td>857</td>
<td>7,822</td>
</tr>
<tr>
<td>2018-06-30</td>
<td>18,706</td>
<td>1,398</td>
<td>17,308</td>
</tr>
<tr>
<td>2018-12-31</td>
<td>5,576</td>
<td>1,630</td>
<td>3,946</td>
</tr>
<tr>
<td>2019-06-30</td>
<td>1,438</td>
<td>440</td>
<td>998</td>
</tr>
<tr>
<td>Total</td>
<td>65,705</td>
<td>5,083</td>
<td>60,622</td>
</tr>
</tbody>
</table>

Sources: NPL dataset – Bank of Italy

The dataset contains the dummy variable for the observed confirmation of remarks (Conφ) and the regressors. In particular, there are 430 dummy variables, of which 15 for DQCs and 415 for RAs, and numeric variables related to the imbalances defined as the differences among quantitative aggregates of the remark, the average number of records sent by an RA for a specific reference date, and the reference dates.

A wide set of commonly used supervised approaches were considered: ridge logistic classifier (Hoerl and Kennard, 1970; Le Cessie and Van Houwelingen, 1992; Schaefer et al., 1984; Cule and De Iorio, 2012; Van Wieringen, 2020); linear and quadratic discriminant analysis (LDA and QDA); decision tree classifier; k-neighbors
classifier and random forest (James et al., 2013; Hastie et al., 2009). The results of the application of these approaches to the training set are reported in Table 2.

### Model comparison in the training set and in the validation set

<table>
<thead>
<tr>
<th>Model</th>
<th>Logistic regression</th>
<th>Ridge logistic classifier ($\lambda=1$)</th>
<th>Linear discriminant analysis</th>
<th>Decision tree classifier</th>
<th>Quadratic discriminant analysis</th>
<th>K-neighbors classifier</th>
<th>Random forest</th>
</tr>
</thead>
<tbody>
<tr>
<td>Training set</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Optimal cut-off</td>
<td>0.41</td>
<td>0.69</td>
<td>0.50</td>
<td>0.52</td>
<td>0.49</td>
<td>0.53</td>
<td>0.71</td>
</tr>
<tr>
<td>Accuracy</td>
<td>0.8283</td>
<td>0.8277</td>
<td>0.7265</td>
<td>0.7252</td>
<td>0.7256</td>
<td>0.7271</td>
<td>0.5044</td>
</tr>
<tr>
<td>Recall</td>
<td>0.9516</td>
<td>0.9168</td>
<td>0.9902</td>
<td>0.9355</td>
<td>0.9938</td>
<td>0.9807</td>
<td>0.3944</td>
</tr>
<tr>
<td>Precision</td>
<td>0.8347</td>
<td>0.8558</td>
<td>0.7293</td>
<td>0.7483</td>
<td>0.7274</td>
<td>0.7330</td>
<td>0.8346</td>
</tr>
<tr>
<td>Negative predictive value</td>
<td>0.7980</td>
<td>0.7303</td>
<td>0.5541</td>
<td>0.5023</td>
<td>0.5435</td>
<td>0.5390</td>
<td>0.3325</td>
</tr>
<tr>
<td>Validation set</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Accuracy</td>
<td>0.8068</td>
<td>0.7841</td>
<td>0.7636</td>
<td>0.7523</td>
<td>0.7795</td>
<td>0.7545</td>
<td>0.7795</td>
</tr>
<tr>
<td>Recall</td>
<td>0.9417</td>
<td>0.8455</td>
<td>0.9738</td>
<td>0.9038</td>
<td>0.9942</td>
<td>0.9359</td>
<td>1.0000</td>
</tr>
<tr>
<td>Precision</td>
<td>0.8325</td>
<td>0.8735</td>
<td>0.7786</td>
<td>0.8031</td>
<td>0.7821</td>
<td>0.7887</td>
<td>0.7795</td>
</tr>
<tr>
<td>Negative predictive value</td>
<td>0.6154</td>
<td>0.5093</td>
<td>0.1818</td>
<td>0.3889</td>
<td>0.5000</td>
<td>0.3333</td>
<td>NA</td>
</tr>
</tbody>
</table>

Sources: NPL dataset – Bank of Italy

The good performance of the logistic ridge regression with respect to other methods is not new in the literature (Bradley, 1996). Similar results have been obtained when the considered dataset had a large set of uncorrelated variables (Cornell-Farrow and Garrard, 2020), as it is the case of the NPL dataset. The logistic regression, as a special case of the ridge logistic classifier, has been selected for the model estimation since this method outperforms the others in terms of accuracy, recall and precision. The LDA, QDA and the random forest do not predict properly the probability of confirmation of a remark in the validation set, as (about) all the remarks are expected to be confirmed. QDA should exhibit a better performance in case of non-linear decision boundary, and in our empirical analysis its performance turned out to be only slightly higher than the LDA considering the validation set. A decision tree classifier and a k-neighbors classifier provide similar results than the previous models, but they still underperform the ridge logistic classifier. A few studies suggested that the latter models may improve their performance when filtering or clustering of features is applied (Ala’raj et al., 2020; Rajaguru and Chakravarthy, 2019). However, this may be associated with enhanced skewness of the results, increasing their sensitivity and reducing their specificity.

Following Le Cessie and Van Houwelingen (1992) and considering that in our case the binary response $Conf$ and a set of predictors $X$ of $U$ dimensions are measured on remarks, the estimated probability that a remark is confirmed is given by $p(Conf)$, where the probability function $p$ follows the logistic regression model:

$$ p(Conf) = \frac{\exp(X\beta)}{1 + \exp(X\beta)} $$

The ridge logistic classifier is obtained by maximizing the likelihood function $l(\beta)$ with a penalized parameter $\lambda$ applied to all the coefficients $\beta$ except the intercept ($\beta_0$); in such case the estimator will be:
The logistic ridge regression estimator depends on the choice of a tuning parameter $\lambda \geq 0$ to be determined separately (Cule and De Iorio, 2013; Le Cessie and Van Houwelingen, 1992).

In order to maximize the model’s performance, $\lambda$ and the cut-off parameter were optimized by cross-validated grid-search. The best results are reached when the cut-off is 0.41 and $\lambda$ is set to zero, i.e. when the standard logistic regression is applied.

4. Main results for the selected model

The results for the estimation of future confirmation of the remarks through the logistic regression model are shown in Section 4.1; the results related to the application of the proposed decision-making algorithm are presented Section 4.2.

4.1 Logistic regression: estimation of ‘remark confirmed, Yes/No’

The logistic regression model used to estimate the probability that a remark is confirmed is defined as follows:

$$
\text{Logit}(p(\text{Conf})) = \beta_0 + \beta_1 \cdot \log(\text{Imbalance}) + \beta_2 \cdot \log^2(\text{Imbalance}) + \beta_3 \cdot \log(\text{Records}) + \\
+ \beta_4 \cdot \text{Reference date} + \sum_{c=1}^{15} \gamma_c \cdot \text{DQC}_c + \sum_{p=1}^{415} \delta_p \cdot \text{RA}_p + \varepsilon
$$

The estimated confirmation $\hat{\text{Conf}}$ of a remark, generated by a DQC $c$ for the $(k+1)^{th}$ data submission sent by the RA $p$ for the reference date $t$, is given by the following:

$$
\hat{\text{Conf}}_{t,p,c,k+1} = \begin{cases} 
1, & \text{if } p(\text{Conf}_{t,p,c,k+1}) > 0.41 \\
0, & \text{otherwise}
\end{cases}
$$

Where 0.41 is the optimal value for the cut-off since it maximizes the accuracy of the model. The model was estimated both on the training set and validation set. In general, all the measures computed in the training and validation sets suggest a satisfying performance of the chosen classification model (Table 3). The measures selected for assessing the goodness of the classification show high and similar values in the two sets, except for the negative predictive value that is 0.7980 and 0.6154, respectively, in the training set and in the validation set.

Since the purpose of the decision rule is to determine whether the DQL has improved by means of a subsequent data transmission, cases where remarks are predicted as confirmed when they should not (false positive) should be limited. This is a precautionary approach which however may lead to consider the DQL related to the new data submission as insufficient and request further analysis of remarks to the RA.

The results show that the model estimation achieves a high level of precision (0.8347 in the training set; 0.8325 in the validation set) and, at the same time, an
A decision-making rule to detect insufficient data quality (163 in the training set; 20 in the validation set).

Confusion matrix computed in the training set and the validation set

<table>
<thead>
<tr>
<th>Conf</th>
<th>Value</th>
<th>Measure</th>
<th>Value</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>No</td>
<td>Yes</td>
<td>Total</td>
</tr>
<tr>
<td>No</td>
<td>644 (TN)</td>
<td>163 (FN)</td>
<td>807</td>
</tr>
<tr>
<td>Yes</td>
<td>634 (FP)</td>
<td>3,202 (TP)</td>
<td>3,836</td>
</tr>
<tr>
<td>Total</td>
<td>1,278</td>
<td>3,365</td>
<td>4,643</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Conf</th>
<th>Value</th>
<th>Measure</th>
<th>Value</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>No</td>
<td>Yes</td>
<td>Total</td>
</tr>
<tr>
<td>No</td>
<td>32 (TN)</td>
<td>20 (FN)</td>
<td>52</td>
</tr>
<tr>
<td>Yes</td>
<td>65 (FP)</td>
<td>323 (TP)</td>
<td>388</td>
</tr>
<tr>
<td>Total</td>
<td>97</td>
<td>343</td>
<td>440</td>
</tr>
</tbody>
</table>

Note: TN true negative; FN false negative; TP true positive; FP false positive; Conf and Conf are the actual and predicted values of the variable Conf, respectively.
Sources: NPL dataset – Bank of Italy

4.2 Application of the decision-making rule

As presented in Section 2, we analyse the application of the proposed decision rule to the ‘case B’ where two consecutive submissions are considered as releasable since they are not impaired by serious remarks. In particular, the decision-making algorithm (1) identifies whether a new report $k+1$ improves or worsens the DQL respect to the previous report $k$, sent by an RA $p$ for the same reference date $t$.

An application on the NPL dataset has been performed by comparing in terms of DQL indicator all the 275 pairs of consecutive data submissions in the ‘case B’ with reference dates from 2017 to 2018 and the 14 ones referred to June 2019 (Table 4).

The results of the application of the decision rule confirm that the proposed method prevents the Data Manager from releasing non-fit-for-use data to the users since it automatically identifies, respectively for the reference dates 2017-2018 and June 2019, 29 and 1 cases where the DQL decreases with the subsequent submission (Table 5).

As expected the limited number of cases of DQL worsening is consistent with the RAs’ data correction process that is typically characterised by an improvement of the
A decision-making rule to detect insufficient data quality DQL over time. This consistency is further corroborated by the results related to a possible application of the decision rule to the other cases ‘A’, ‘C’ and ‘D’.

### Classification of submissions in the current approach

#### Table 4

<table>
<thead>
<tr>
<th>$k^{th}$ submission</th>
<th>$(k+1)^{th}$ submission</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Not-released</td>
<td>Released</td>
<td>Total</td>
</tr>
<tr>
<td>Not-released</td>
<td>269</td>
<td>407</td>
</tr>
<tr>
<td>Released</td>
<td>51</td>
<td>275</td>
</tr>
<tr>
<td>Total</td>
<td>320</td>
<td>682</td>
</tr>
</tbody>
</table>

#### Reference date June 2019

<table>
<thead>
<tr>
<th>$k^{th}$ submission</th>
<th>$(k+1)^{th}$ submission</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Not-released</td>
<td>15</td>
<td>23</td>
</tr>
<tr>
<td>Released</td>
<td>1</td>
<td>14</td>
</tr>
<tr>
<td>Total</td>
<td>16</td>
<td>37</td>
</tr>
</tbody>
</table>

Sources: NPL dataset – Bank of Italy

### Application of the decision-making algorithm to ‘case B’

#### Table 5

<table>
<thead>
<tr>
<th>Results of the decision rule</th>
<th>Reference dates between 2017 and 2018</th>
<th>Reference date of June 2019</th>
</tr>
</thead>
<tbody>
<tr>
<td>Released submission</td>
<td>246 (89%)</td>
<td>13 (93%)</td>
</tr>
<tr>
<td>Additional Not-released submission</td>
<td>29 (11%)</td>
<td>1 (7%)</td>
</tr>
<tr>
<td>Total</td>
<td>275 (100%)</td>
<td>14 (100%)</td>
</tr>
</tbody>
</table>

Sources: NPL dataset – Bank of Italy

In order to assess the appropriateness of the classification of the submissions as released and additional not-released according to the decision rule, a comparison with a benchmark is carried out. In particular, the benchmark considered is obtained from the application of the decision rule when the observed variable $Conf$ is considered instead of the estimated value according to the logistic model of Section 4.1.

\[
\hat{I}_{t,p,k+1} \leq I_{t,p,k} \quad (2)
\]

where $I_{t,p,k+1}$ is computed using the estimated $\hat{Conf}_{t,p,k+1}$.

The application of rule (1) is compared with (2) and the results are shown in Table 6 for the time period from 2017 to 2018 and June 2019, respectively. The results show that for the reference dates from 2017 to 2018, in 97% of cases the prediction

\[ \text{See Appendix B from La Ganga et al., 2022.} \]
is effective; while, considering the reference date of June 2019, the decision is definitely the same in both cases.

Verification of the decision rule: comparison with the benchmark decision rule that considers the observed Conf

<table>
<thead>
<tr>
<th>Reference dates between years 2017 and 2018</th>
<th>Benchmark decision on the ((k+1)^{th}) submission (using Conf)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Decision on the ((k+1)^{th}) submission considering Conf</td>
<td>Not-released</td>
</tr>
<tr>
<td>Not-released</td>
<td>22</td>
</tr>
<tr>
<td>Released</td>
<td>2</td>
</tr>
<tr>
<td>Total</td>
<td>24</td>
</tr>
</tbody>
</table>

Accuracy = 0.9673

<table>
<thead>
<tr>
<th>Reference date June 2019</th>
<th>Benchmark decision on the ((k+1)^{th}) submission (using Conf)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Decision on the ((k+1)^{th}) submission considering Conf</td>
<td>Not-released</td>
</tr>
<tr>
<td>Not-released</td>
<td>1</td>
</tr>
<tr>
<td>Released</td>
<td>0</td>
</tr>
<tr>
<td>Total</td>
<td>1</td>
</tr>
</tbody>
</table>

Accuracy = 1

Sources: NPL dataset – Bank of Italy

Conclusions

In literature on data quality, this topic is generally approached by monitoring different dimensions, e.g. timeliness, accuracy and consistency of data. This paper proposes a decision-making rule that improves the monitoring of the accuracy, completeness and consistency of data sent by RAs to the Authorities and supports the Data Manager in deciding whether to immediately release the data to internal and external users.

More broadly, especially in all cases where data reporting is made more challenging for RAs due to exogenous and unpredictable events, such as also the pandemic fallout on organisational issues, the proposed DQL indicator can play an important role in making more efficient the DQM activity.

Using machine learning techniques based on the results of the automatic validation process and Data Manager’s past evaluations of explanations received by RAs, we assess the variation of the DQL between two consecutive data submissions sent by the same RA and for a specific reference date. From a methodological point of view, a logistic regression model has been used to predict the confirmation probability of a single remark in order to estimate the difference in DQL between two versions of the same dataset (the second carrying the revisions of the first). The final model has been selected by comparing different available alternatives and optimizing the goodness of fit. In our case, the logistic regression model outperforms several models that are commonly used in machine learning studies.
In the second phase of the study, a decision-making rule has been developed by taking into account the estimated confirmation probability of a single remark, which emerged from the previous step, together with the total number and the severity of the remarks. The results of the application on the NPL dataset are remarkable, since in 97 per cent of cases the decision rule provides coherent results to those observed when the actual status is known.

In the current practice, the comparison between two consecutive reports submitted by the same RA is left to the judgment and expertise of the Data Manager and the interaction with the RA. The proposed decision-making rule leads to a less time-consuming and more harmonized approach. Moreover, it supports the Data Manager to determine whether data revisions do improve the DQL and it provides guidance for prioritizing DQM activities. Although the methodology is applied to a specific (granular) dataset – the banks’ Non-Performing Loans reporting of the Bank of Italy – the method can be applied to other datasets in order to improve the DQL assessment, provided that past evidence from the DQM activity consisting of outliers, their severity and the confirmation is available.

References


Banca d’Italia (2016), ‘Instructions for the editing of reports on non-performing exposures’.


Cule, E., De Iorio, M. (2012), ‘A semi-automatic method to guide the choice of ridge parameter in ridge regression’, Imperial College London and University College London.


A decision-making rule to detect insufficient data quality

an application of statistical learning techniques to the
non-performing loans banking data

Barbara La Ganga, Paolo Cimbali, Marco De Leonardis, Alessio Fiume, Luciana Meoli and Marco Orlandi

Banca d’Italia
Motivation

- It is key to count on an efficient and effective monitoring of the quality level of the data transmitted by Reporting Agents (RAs) in order to provide users with high-reliable data to carry out thorough and robust analyses.

- Data Quality Level (DQL) generally follows a positive trend thanks to subsequent corrections submitted by RAs; however, a data quality worsening may occur especially when data production is affected by exogenous and unpredictable events, such as RAs’ IT malfunctions, changes in the reporting requirements or operative tensions and staff shortage (also as seen during the pandemic).

- The aim of the study is to define a decision-making rule:
  - to speed up the detection of DQL worsening;
  - to provide a synthetic measure of the DQL.
At each data submission, the reliability of the data is assessed upon arrival by the Banca d’Italia by using a set of automatic Data Quality Checks (DQCs).

A severity level from 0 to 10 is assigned to each DQC.

When a DQC detects plausible errors (outliers) or deterministic errors, remarks are sent to the RA to request for:

- corrections of erroneous data by sending a new data submission
  or
- confirmations of the data. These can be, in turn, accepted or refused by the Data Manager.
Based on the severity level of the DQC, the generated remarks are classified as “serious” and “non-serious”. If at least 1 serious remark is generated, the data submitted are kept on hold to be examined by the Data Manager (hence not immediately released to the users).

Considering 2 subsequent data submissions sent by an RA for a specific reference date, the possible cases are as follows:

- In cases A, C and D, the Data Manager’s decision is straightforward; in case B the (k+1)th submission may worsen the DQL.
- The proposed decision-making rule is applied to case B to detect the unexpected worsening of the DQL.

<table>
<thead>
<tr>
<th>kth submission</th>
<th>(k+1)th submission</th>
</tr>
</thead>
<tbody>
<tr>
<td>Not-released</td>
<td>D</td>
</tr>
<tr>
<td>Released</td>
<td>A</td>
</tr>
<tr>
<td></td>
<td>C</td>
</tr>
<tr>
<td></td>
<td>B</td>
</tr>
</tbody>
</table>
Definition of the proposed decision-making rule

- The proposed rule is based on a synthetic data quality indicator computed through past evidence from the Data Quality Management (DQM) activity:
  - number of remarks (R) generated by the DQC
  - severity level (\( \tau \))
  - number of confirmations (Conf)

- Definition of a synthetic data quality indicator \( I_k \) for the \( k^{th} \) data submission sent by an RA for a specific reference date:

\[
I_k = \sum_c \tau_c \cdot (R_{c,k} - Conf_{c,k})
\]

- If the DQC detects deterministic errors precisely (non-confirmable DQCs), \( Conf_{c,k} \) is by construction equal to 0.
- The higher the value of \( I_k \), the lower the DQL of the \( k^{th} \) data submission.
- The proposed decision-making rule is defined as follows:

AT EACH DATA SUBMISSION

CASE A/D

CASE B

CASE C

Decision-making rule:

\( I_{k+1} > I_k \) ?

YES

CASE A/D

NOT RELEASED

CASE B

RELEASEd

CASE C

NO
What quantities are available for the calculation of $I_k$ and $I_{k+1}$?

- Let us assume we want to compare the DQL of the $(k+1)^{th}$ data submission with the DQL of the $k^{th}$. Once the $(k+1)^{th}$ data submission is received, the availability of the information for the calculation of $I_k$ and $I_{k+1}$ is as follows:

<table>
<thead>
<tr>
<th>Number of remarks</th>
<th>✓</th>
<th>✓</th>
</tr>
</thead>
<tbody>
<tr>
<td>severity level</td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td>Number of confirmations</td>
<td>✓</td>
<td>✗</td>
</tr>
</tbody>
</table>

- The number of confirmations related to remarks, generated by the confirmable DQC $c$ for the $(k+1)^{th}$ data submission, is estimated:

$$\text{Conf}_{c,k+1} = \sum_{r=1}^{R_{c,k+1}} \text{Conf}_{c,k+1, r}$$

where:

$$\text{Conf}_{c,k+1, r} = \begin{cases} 1, & \text{if } p(\text{Conf}_{c,k+1, r}) > \text{cut-off} \\ 0, & \text{otherwise} \end{cases}$$

- **cut-off** is a threshold lying within $(0, 1)$ assessed with a cross-validation method.

- The estimation of the probability $p(\text{Conf})$ is derived applying machine learning techniques to a dataset including remarks generated by confirmable DQCs actually observed in the previous reference dates.
Dataset and Model selection

- **Dataset:** Banks Non-performing loans dataset (NPL), collected by Banca d’Italia on a biannual basis
  - over 17 million of records between 30th June 2017 and 30th June 2019
  - about 65K remarks generated, of which 5,083 by confirmable DQCs
  - 15 dummy variables for DQCs and 415 for RAs
  - numeric variables: differences among quantitative aggregates of remarks, number of records sent and reference dates

- **Model selection:** the logistic regression model outperforms.

<table>
<thead>
<tr>
<th>Model</th>
<th>Logistic regression</th>
<th>Ridge logistic classifier (λ=1)</th>
<th>Linear discriminant analysis</th>
<th>Decision tree classifier</th>
<th>Quadratic discriminant analysis</th>
<th>K-neighbors classifier</th>
<th>Random forest</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Optimal cut-off</strong></td>
<td>0.41</td>
<td>0.69</td>
<td>0.50</td>
<td>0.52</td>
<td>0.49</td>
<td>0.53</td>
<td>0.71</td>
</tr>
<tr>
<td><strong>Training set</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>• from June 2017 to December 2018</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>• 4,643 remarks</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Accuracy</strong></td>
<td>0.83</td>
<td>0.83</td>
<td>0.73</td>
<td>0.73</td>
<td>0.73</td>
<td>0.73</td>
<td>0.50</td>
</tr>
<tr>
<td><strong>Recall</strong></td>
<td>0.95</td>
<td>0.92</td>
<td>0.99</td>
<td>0.94</td>
<td>0.99</td>
<td>0.98</td>
<td>0.39</td>
</tr>
<tr>
<td><strong>Precision</strong></td>
<td>0.83</td>
<td>0.86</td>
<td>0.73</td>
<td>0.75</td>
<td>0.73</td>
<td>0.73</td>
<td>0.83</td>
</tr>
<tr>
<td><strong>Negative predictive value</strong></td>
<td>0.80</td>
<td>0.73</td>
<td>0.55</td>
<td>0.50</td>
<td>0.54</td>
<td>0.54</td>
<td>0.33</td>
</tr>
<tr>
<td><strong>Validation set</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>• June 2019</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>• 440 remarks</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Accuracy</strong></td>
<td>0.81</td>
<td>0.78</td>
<td>0.76</td>
<td>0.75</td>
<td>0.78</td>
<td>0.75</td>
<td>0.78</td>
</tr>
<tr>
<td><strong>Recall</strong></td>
<td>0.94</td>
<td>0.85</td>
<td>0.97</td>
<td>0.90</td>
<td>0.99</td>
<td>0.94</td>
<td>1.00</td>
</tr>
<tr>
<td><strong>Precision</strong></td>
<td>0.83</td>
<td>0.87</td>
<td>0.78</td>
<td>0.80</td>
<td>0.78</td>
<td>0.79</td>
<td>0.78</td>
</tr>
<tr>
<td><strong>Negative predictive value</strong></td>
<td>0.62</td>
<td>0.51</td>
<td>0.18</td>
<td>0.39</td>
<td>0.50</td>
<td>0.33</td>
<td>NA</td>
</tr>
</tbody>
</table>

Sources: NPL dataset – Banca d’Italia
Considering the subsequent submissions of the case B, the decision-making rule allows the Data Manager to **automatically and promptly identify cases where the DQL decreases** and it **prevents the users to use non-fit-for-use data.**

<table>
<thead>
<tr>
<th>Reference dates between years 2017 and 2018</th>
<th>$(k+1)^{th}$ submission</th>
<th>Released submissions</th>
<th>Additional Not-released submissions</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Not-released</td>
<td>Released</td>
<td>Total</td>
</tr>
<tr>
<td><strong>$k^{th}$ submission</strong></td>
<td>269</td>
<td>407</td>
<td>696</td>
</tr>
<tr>
<td>Released</td>
<td>51</td>
<td>275 (Case B)</td>
<td>326</td>
</tr>
<tr>
<td>Total</td>
<td>320</td>
<td>682</td>
<td>1,002</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Reference date of June 2019</th>
<th>$(k+1)^{th}$ submission</th>
<th>Released submissions</th>
<th>Additional Not-released submissions</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Not-released</td>
<td>Released</td>
<td>Total</td>
</tr>
<tr>
<td><strong>$k^{th}$ submission</strong></td>
<td>15</td>
<td>23</td>
<td>38</td>
</tr>
<tr>
<td>Released</td>
<td>1</td>
<td>14 (Case B)</td>
<td>15</td>
</tr>
<tr>
<td>Total</td>
<td>16</td>
<td>37</td>
<td>53</td>
</tr>
</tbody>
</table>

**Sources:** NPL dataset – Banca d’Italia
Conclusions

- The proposed decision-making rule **improves the current DQL monitoring** by promptly detecting additional cases of DQL worsening.

- The synthetic data quality indicator $I_k$ provide a **synthetic measure of the overall quality of data** transmitted by the RAs.

- **The decision-making rule is accurate.** It was assessed by comparing its results with the outcome resulting from an application of the decision-making rule based on the real status of the remarks confirmability: in 97% of cases the conclusions coincide.

- The proposed method can be **flexibly applicable to various data collections.**

- For the NPL dataset, the **implementation of the decision-making rule** in the Banca d’Italia’s collection system is **ongoing.**
Thank you for your attention!

barbara.laganga@bancaditalia.it
Data quality management of entity group data: relevance and tools to address current challenges

Bruno Carreiras,
European Central Bank

1 This presentation was prepared for the conference. The views expressed are those of the author and do not necessarily reflect the views of the BIS, the IFC or the central banks and other institutions represented at the event.
Data quality management of entity group data: relevance and tools to address current challenges

Bruno Carreiras

The COVID-19 crisis brought renewed challenges for statisticians in Central Banks, especially those dealing with granular datasets. This paper addresses some of those challenges in the context of entity group data by proposing a metric to identify “relevant” changes in group compositions (joiners and leavers) and rank them by importance. We show how such an indicator can provide a useful tool for an effective prioritisation of data quality management activities within the statistical community. As a complement, we also show how the same indicator can be used as an evaluation technique for corporate changes.

Keywords: granular data, data management, data quality, groups, master-data, statistics, network data, central banking.

Contents

Data quality management of entity group data: relevance and tools to address current challenges ........................................................................................................................... 1
Introduction ............................................................................................................................................... 2
1. Background ........................................................................................................................................... 3
2. Quality assurance of group data in RIAD .................................................................................. 4
3. Modelling relevance and prioritising group movements ................................................... 5
   3.1 Significance Multiplier to measure relevance of groups and group movements .............................................................. 5
   3.2 Prioritising and evaluating movements under a dynamic business environment ............................................................................. 6
4. Implementation: Significance Multiplier in practice ............................................................. 8
   4.1 Data preparation and construction of the dataset .................................................... 8
   4.2 Prioritisation of data quality follow-up on movements ........................................... 8
   4.3 Valuing changes in corporate structures..................................................................... 11
5. Conclusions ......................................................................................................................................... 14

1 European Central Bank, DG-Statistics, Analytical Credit and Master Data Division, Research Analyst. I would like to thank my colleagues from the RIAD team in DG-Statistics for their input and guidance. The views expressed in this paper are those of the author and do not necessarily reflect the views of the ECB.
Introduction

Data quality is one of the most crucial topics in the world of increasing data volumes in central banking. As datasets become more granular and data is collected more frequently, data quality issues increase and can be more difficult to assess. This combined with the need for accurate statistical indicators in real time, amplified with the recent COVID-19 crises (De Beer, Tissot 2020), creates an environment where data must be managed efficiently.

Some argue that innovation can be a tool to respond to these challenges. The usage of machine learning and other innovative techniques to manage data quality has already proven useful to some of the ECB’s granular datasets (see Romano et al, 2021). This work intends to be an addition to that innovative wave by proposing a new way to efficiently prioritise the data quality management of the entity group structures stored in the RIAD, the ECB master data platform on entities.

RIAD groups are ‘network trees’ (company structures) which are generated each day based on relationship data between entities. These groups are monitored every day for any changes in their structure (movements) and any quality issues are followed up. With the steady increase in the number of relationships and consequently in the number of groups, more movements need to be assessed by data quality managers. Under this context, it is important to know which group movements are the most relevant so to follow-up in an efficient manner.

This paper tries to respond to that need by proposing a metric, the Significance Multiplier (SM), that ranks movements inside group structures according to their impact on supported business processes. This metric relies on an input-based model, where users can decide “what is important for them”, depending on their specific loss function. As a by-product, the same tool allows evaluating group changes in corporate structures motivated by real business cases (e.g. mergers or acquisitions). By being a complementary calculation to the group structure and an addition to the dataset, the metric can also be refreshed on a daily basis. Thus, it can also be used to identify a priori the impact of certain movements and support decision makers.

The paper is structured as follows: in Section 1 the RIAD dataset and the RIAD group structures are described. In Section 2, the data quality checks that are applied to RIAD relationship data are laid out. Section 3 introduces a baseline concept for group importance/relevance and related data quality prioritisation. Moreover, that same baseline concept is applied into the context of the group members and movements. Still in Section 3, the baseline scenario is extended to the input-based approach that aims to internalise specific business needs in the significance multiplier.

---

2 “Don’t take it for granted: the value of high-quality data and statistics for the ECB’s policymaking”- Schnabel (2020)

3 The potential role of innovation is mentioned in Rosolia et al (2021) or De Beer, Tissot (2021)

4 Register of Institutions and Affiliates Data, see Background section for more information.
In Section 4, two applications are laid out, related to data quality prioritisation (via priority ranking) and to the evaluation of movements under a specific business case. Section 5 draws the main conclusions and puts forward possible future extensions.

1. Background

The Register of Institutions and Affiliates Data (RIAD)\(^5\) is the ESCB shared master dataset on entities serving several central banking and Single Supervisory Mechanism (SSM) (even daily) business processes and statistical data collections. RIAD is operated in the ECB’s Directorate General Statistics and provides master data (e.g. name, address, legal form, institutional sector) on various types of Organisational Units (legal entities and their affiliates, such as branches and subsidiaries). An important feature of RIAD is that it provides the relationships (bilaterally) linking between the entities.

RIAD data is collected from National Central Banks (NCBs) and other National Competent Authorities (NCAs)\(^6\) compliant with several legal acts\(^7\), and provided by the latter based on various data sources available at national and international level.

The RIAD system allows data providers to create, change and delete data (including relationships) on any Organisational Unit (or ‘entity’, hereinafter) on a daily basis. This flow provides the flexibility needed for a central system to gather information coming from around Europe and swiftly replicate business changes, hence providing always up-to-date information needed for the performance, inter alia, of key central banking activities (e.g. risk management and collateral management). Based on (entity-to-entity) relationship data provided by all relevant sources, RIAD calculates (different types of) group structures every day.

RIAD groups are based on navigation algorithms that start from the valid relationships collected in the system and generate a ‘network tree structure’. Figure 1 shows how a typical RIAD group network could be represented graphically.

These algorithms were designed to support various stakeholders with respective business processes: as a result, there are currently four different group calculations in RIAD\(^8\), they are used to perform tasks such as assessing the links between institutions

---

\(^5\) As defined in Guideline (EU) 2018/876 of the European Central Bank on the Register of Institutions and Affiliates Data (RIAD)

\(^6\) National competent authorities in the context of SSM data collection and as defined in: Guideline (EU) 2020/497 of the European Central Bank on the recording of certain data by national competent authorities in the Register of Institutions and Affiliates Data

\(^7\) RIAD data provision is based on the following legal acts: Guideline (EU) 2018/876 of the European Central Bank on the Register of Institutions and Affiliates Data (RIAD); Guideline (EU) 2019/1335 of the European Central Bank amending Guideline (EU) 2018/876 on the Register of Institutions and Affiliates Data (RIAD); Guideline (EU) 2020/497 of the European Central Bank on the recording of certain data by national competent authorities in the Register of Institutions and Affiliates Data

\(^8\) For this paper only the algorithm based on ownership relationships is analysed, any reference to number of groups only relates to that algorithm. In short, those groups are calculated in the following way: if entity A has an ownership relationship towards two entities (B and C), RIAD would calculate a group structure where A is the head and B and C are its children. If B would own another entity D, then the group structure would comprise four entities (adding one to the bottom). Additionally, one-to-one relationships are also considered a group (with two members).
pledging collateral for monetary policy operations, monitoring supervised banking groups, or calculating consolidated (loans/securities) exposures. As a result, any data quality issue affecting RIAD groups data (and, before that, the underlying relationships) may negatively affect such important tasks.

Currently there are ca 130,000 groups in RIAD, based on a universe of ca 300,000 relationships. As explained above, any changes on the universe of relationships may create movements (genuine or not) in the group structures. Assessing the changes in such groups, and ensuring data quality, is a clear challenge given limited resources and user-specific priorities.

![Figure 1 - Graphical representation of a RIAD group structure.](image)

2. Quality assurance of group data in RIAD

Relationships in RIAD (the source data for group calculation) are subject to automated quality checks embedded in the system and serving as a check for the consistency, completeness, and accuracy of the data. Relationships which do not pass such controls are not considered in the calculation. On the other hand, valid relationships enter the daily calculation and generate the group population.

It often occurs that the daily group calculation is affected by the dynamic nature of the RIAD collection environment. At any point in time, an entity may leave or join a group and create a change in its structure. These are the so-called group movements. A movement is triggered by a change in the underlying data that can be either genuine or erroneous. This brings the need for an extra layer of quality checks on groups: **post-calculation monitoring**.

Let’s take as an example the case of a new relationship being uploaded into RIAD on a certain day: if the new relationship is added to the top of a group signalling a new head, it can be assessed either of two ways:

1. This change is legitimate and in fact the group has a new head (e.g. a new holding at the top of the group).
2. This change is not legitimate and reflects a mistake in the upload of a relationship (e.g. wrong parent/child direction).

Case b) is a clear inconsistency, while case a) is a genuine change. In the end, experience proved that there would always need to be a subjective “eye” that would define legitimacy in both cases.
In a large and growing group population, if a data provider/data quality manager would need to confirm the legitimacy of every movement, there would be a significant operational effort related to group monitoring. To mitigate this effort, it is important to define a prioritisation measure that can point a data quality manager to which cases should be assessed first.

3. Modelling relevance and prioritising group movements

3.1 Significance Multiplier to measure relevance of groups and group movements

The number of groups and the complexity of the underlying data are challenging to manage. To efficiently deal with the quality management of such an amount of information there is a need to classify all potential issues so to allow DQ managers to focus on the most relevant changes/movements. In other words, there is a need to define a concept of “significance” or “relevance” in the context of RIAD group structures.

The starting premise is that the baseline criteria should be the size of the group (S) affected, i.e. number of group members. If \( S_{G1} > S_{G2} \), then G1 is more relevant than G2 and data quality should be prioritised for G1. By assessing the size of the group (\( S_G \)), one can rank all existing groups by importance and define a baseline prioritisation metric.

However, the size of group networks as a first level of prioritisation can be reductive. The number of members in a group does not necessarily represent the best measure of how relevant a movement inside the group is.

If a group member moves, the group network is affected at every level that is dependent on that group member, i.e. a group member always leaves with all its children (if any). Based on this, one can state that a movement is as significant as the group members that it involves. Thus, the importance of a group member/movement can then be defined as the size of the movement in its group resulting from a single event affecting its own presence, or significance multiplier for RIAD group structures (SM). In simpler terms, the SM is the size of the sub-group headed by a certain group member.

Under the assumptions outlined above, looking at figure 2, the exit at T+1 of group member X is less important as it impacts only its own presence (SM=1). While in figure 3, the movement of group member X (as leaver), affects the presence of three entities, signalling that this is a more relevant group member (SM=3).

---

9 This metric is an application of a weighted centrality measure for network data to the RIAD data model and its algorithms for a concrete business purpose.
The SM provides a good measure for importance of a member inside the group population, i.e. you can compare entities inside the same group and across groups. If the SM of \( i \) in group \( G \) is higher than the SM of \( j \) in group \( G \), then \( i \) is more important than \( j \), by the same logic any movement that starts in \( i \) is more relevant than any movement that starts in \( j \). Under the baseline scenario that the amount of entities affected is what matters, one can also compare the SM across groups. This suffers from the limitation that groups with more members will necessarily have more significance (this also applies to its movements). This raises the challenge of finding a metric that is more accurate in defining relevance in the whole population.

3.2 Prioritising and evaluating movements under a dynamic business environment

The usage of RIAD group structures has increased significantly in recent years. Many stakeholders are interested in performing their analysis with RIAD data; besides that, several key central bank functions (e.g. risk management, collateral management) rely on RIAD data for their (daily!) operational processes. All RIAD clients have different expectations and needs and hence different criteria to identify “relevant” groups. As an example, in the context of AnaCredit\(^{10}\) such relevance may be measured with a credit-related property (e.g. outstanding nominal amount of the loan), while for the

---

\(^{10}\) As defined in: Regulation (EU) 2016/867 if the European Central Bank on the collection of granular credit and credit risk data
climate change purposes\textsuperscript{11} this could be based on the value of variables such as the carbon intensity or the main economic activities (as a proxy for that). The nature of these differences raises another question: how can one homogeneously apply measures of importance that are suitable to all clients’ expectations and business needs while at the same time coping with limited resources in a shared master dataset?

As defined in the previous section, the importance of a group at its very baseline is measured by the number of its group members. In principle, this would provide a very good proxy of relevance for a group. However, it would suffer from the immediate limitation that relevant groups with a small number of members would be deemed as less important. To tackle this, an \textbf{extension based on a criterion applied to specific business cases ($\alpha$)} is proposed. The business specific criteria would be a pre-defined (input based) variable selected by the stakeholder. This criterion should allow to \textbf{accurately quantify the importance of a group for a certain business process}.

At full group network level, the \textbf{definition of size ($S$) becomes dynamic}, i.e. size should be measured differently than a pure count of group members but rather with a relevant quantifier ($aS_G$):

$$aS_G = \sum_{i \in G} \alpha_i$$

Where $G$ is the group structure, $i$ is a group member belonging to $G$ and $\alpha$ is a business specific criteria for a group member. In short, this would mean that all the business specific criteria for every group member would be summed up.

Under these new conditions, it is now possible that a smaller group (e.g. with 20 members) has more importance than a larger group (e.g. with 200 members), if the sum of its $\alpha$ is larger.

Naturally, the same logic can be applied to group members. While in section 3 the SM is dependent on the number of members affected by an event on one given member, under this extension it is dependent on the sum of $\alpha$ affected by an event on one given member.

$$aSM_i = \sum_{i \in G_k} \alpha_i$$

Where $G$ is the group structure, $G_k$ is the sub-group of $G$ impacted by the single event on $k$. In practice, a group member would be valued as the sum of the business criterion of all the entities it has below it (the “sub-group”). In figure 3, the $aSM$ of X would be the sum of the $\alpha$ of X, Y and Z.

The business dedicated method allows for a clear definition of importance of a member inside the whole group universe by defining a \textbf{criteria ($\alpha$) that provides a universal measure of comparison (cross-group)}. As a result, you can accurately compare two movements (or group members). This is the advantage of the $\alpha$SM in comparison to the original SM.

All in all, the $\alpha$SM for DQM defines relevance of movements by positioning (by importance) each group member in the full universe of group structures.

\textsuperscript{11} New unit in the ECB announced in 2021:
Thus, we are prioritising data quality actions by defining a ranking of importance for every group member. If a group member is touched and triggers a movement, this movement will be ranked by importance, as the αSM incorporates the impact.

The fact that this measure is applied equally to every group member provides a rather important by-product. Besides providing this DQM efficiency gain, it also has a use for different business purposes. The value that this metric allocates to every group member can, for example, be used to evaluate corporate changes, i.e. mergers, splits, acquisitions, in a wide range of topics (the α can be modified to fit different business cases). For example, if a branch of a certain group would be acquired by another group, the αSM would provide the value that would move across groups. Thus, the business specific SM metric provides both a priority and an evaluation technique. In the next section it will be laid out how this metric can work in practice.

4. Implementation: Significance Multiplier in practice

4.1 Data preparation and construction of the dataset

To show an application of the model described in this paper, the RIAD dataset was used. For the two use cases below, data from two group snapshots was used, end of January and end of May 2022¹².

Group data in RIAD is modelled as a network tree data, i.e. a dataset that replicates the tree graph and the nodes and edges inside it. The nodes are represented by the RIAD ID (identifier at database level). To complement these data, some additional reference data (RIAD Code, Name, ESA Sector) and balance sheet data was added.

With this dataset prepared, the main dataset¹³ which includes the application of the SM model is constructed. This adds three main variables: size of the group, Significance Multiplier (SM) of each group member (and head), business specific SM of each group member (and head).

4.2 Prioritisation of data quality follow-up on movements

The tool used for this example flags all the daily movements in banking group structures (similarly to what is depicted in figure 2). The movements are then given a ranking of importance based on the business specific SM (αSM). The criterion (α) used is the individual (domestic) balance sheet¹⁴ of banking group members. The use of this ranking is the focus of this example.

---

¹² Note that these data are available in the ECB for those who have sufficient confidential clearing. Meaning that one with the same accesses can replicate the dataset.

¹³ Dataset description can be found in the Annex

¹⁴ Collected under Guideline (EU) 2018/876 of the European Central Bank on the Register of Institutions and Affiliates Data (RIAD) – Annex I
In the SM model, the baseline theory is the starting point. To prove the usefulness of the business specific criteria, the measures of the model with the baseline values (pure counts) and the business specific values (sum of $a$) are compared.

The outcome will be twofold:

- **A ranking at group level**: all group heads will be ranked by importance ($S_G - baseline, aS_G - business specific$)

- **A ranking at group member level**: all group members (excluding heads) will be ranked by importance ($SM - baseline, aSM - business specific$)

The following results apply to the banking group sector following the baseline criteria:

**Ranking of RIAD banking groups by Size**

<table>
<thead>
<tr>
<th>Group Identifier</th>
<th>Group Name</th>
<th>Size</th>
<th>Ranking</th>
</tr>
</thead>
<tbody>
<tr>
<td>G1</td>
<td>Banking group 1</td>
<td>841</td>
<td>1</td>
</tr>
<tr>
<td>G2</td>
<td>Banking group 2</td>
<td>838</td>
<td>2</td>
</tr>
<tr>
<td>G3</td>
<td>Banking group 3</td>
<td>781</td>
<td>3</td>
</tr>
<tr>
<td>G4</td>
<td>Banking group 4</td>
<td>765</td>
<td>4</td>
</tr>
<tr>
<td>G5</td>
<td>Banking group 5</td>
<td>696</td>
<td>5</td>
</tr>
<tr>
<td>G6</td>
<td>Banking group 6</td>
<td>668</td>
<td>6</td>
</tr>
<tr>
<td>G7</td>
<td>Banking group 7</td>
<td>363</td>
<td>7</td>
</tr>
<tr>
<td>G8</td>
<td>Banking group 8</td>
<td>348</td>
<td>8</td>
</tr>
<tr>
<td>G9</td>
<td>Banking group 9</td>
<td>348</td>
<td>9</td>
</tr>
<tr>
<td>G10</td>
<td>Banking group 10</td>
<td>298</td>
<td>10</td>
</tr>
<tr>
<td>G11</td>
<td>Banking group 11</td>
<td>252</td>
<td>11</td>
</tr>
<tr>
<td>G12</td>
<td>Banking group 12</td>
<td>246</td>
<td>12</td>
</tr>
<tr>
<td>G13</td>
<td>Banking group 13</td>
<td>221</td>
<td>13</td>
</tr>
<tr>
<td>G14</td>
<td>Banking group 14</td>
<td>211</td>
<td>14</td>
</tr>
<tr>
<td>G15</td>
<td>Banking group 15</td>
<td>186</td>
<td>15</td>
</tr>
</tbody>
</table>

1 The above data is anonymised due to confidentiality reasons.

Sources: RIAD group data complemented with author’s own calculations

**Ranking banking group members by Significance Multiplier**

<table>
<thead>
<tr>
<th>Entity Identifier</th>
<th>Entity Name</th>
<th>SM</th>
<th>Ranking</th>
</tr>
</thead>
<tbody>
<tr>
<td>M1</td>
<td>Member 1</td>
<td>559</td>
<td>1</td>
</tr>
<tr>
<td>M2</td>
<td>Member 2</td>
<td>556</td>
<td>2</td>
</tr>
<tr>
<td>M3</td>
<td>Member 3</td>
<td>425</td>
<td>3</td>
</tr>
<tr>
<td>M4</td>
<td>Member 4</td>
<td>371</td>
<td>4</td>
</tr>
<tr>
<td>M5</td>
<td>Member 5</td>
<td>247</td>
<td>5</td>
</tr>
<tr>
<td>M6</td>
<td>Member 6</td>
<td>240</td>
<td>6</td>
</tr>
<tr>
<td>M7</td>
<td>Member 7</td>
<td>235</td>
<td>7</td>
</tr>
<tr>
<td>Group Identifier</td>
<td>Group Name</td>
<td>(a)Size (Bln €)</td>
<td>Size</td>
</tr>
<tr>
<td>------------------</td>
<td>-------------------</td>
<td>-------------------</td>
<td>-------</td>
</tr>
<tr>
<td>G5</td>
<td>Banking group 5</td>
<td>&gt; 2,000</td>
<td>696</td>
</tr>
<tr>
<td>G7</td>
<td>Banking group 7</td>
<td>&gt; 2,000</td>
<td>363</td>
</tr>
<tr>
<td>G8</td>
<td>Banking group 8</td>
<td>&gt; 2,000</td>
<td>348</td>
</tr>
<tr>
<td>G230</td>
<td>Banking group 230</td>
<td>&gt; 1,000</td>
<td>137</td>
</tr>
<tr>
<td>G15</td>
<td>Banking group 15</td>
<td>&gt; 1,000</td>
<td>186</td>
</tr>
<tr>
<td>G9</td>
<td>Banking group 9</td>
<td>&gt; 500</td>
<td>348</td>
</tr>
<tr>
<td>G231</td>
<td>Banking group 231</td>
<td>&gt; 500</td>
<td>97</td>
</tr>
<tr>
<td>G1</td>
<td>Banking group 1</td>
<td>&gt; 500</td>
<td>841</td>
</tr>
<tr>
<td>G2</td>
<td>Banking group 2</td>
<td>&gt; 500</td>
<td>839</td>
</tr>
<tr>
<td>G10</td>
<td>Banking group 10</td>
<td>&gt; 100</td>
<td>298</td>
</tr>
<tr>
<td>G232</td>
<td>Banking group 232</td>
<td>&gt; 100</td>
<td>114</td>
</tr>
<tr>
<td>G233</td>
<td>Banking group 233</td>
<td>&gt; 100</td>
<td>50</td>
</tr>
<tr>
<td>G234</td>
<td>Banking group 234</td>
<td>&gt; 100</td>
<td>5</td>
</tr>
<tr>
<td>G224</td>
<td>Banking group 224</td>
<td>&gt; 100</td>
<td>246</td>
</tr>
<tr>
<td>G12</td>
<td>Banking group 12</td>
<td>&gt; 100</td>
<td>45</td>
</tr>
</tbody>
</table>

1 The above data is anonymised due to confidentiality reasons.

Sources: RIAD group data complemented with author’s own calculations

Thus, according to pure size (Table 1), Banking group 1 would be considered the most important banking group as it is the one with the highest number of members (841). Member 1 would be considered the top group member with 559 entities as its direct or indirect children.

To make a reasonable cross-group comparison, one should investigate the results applied to the banking groups under the business specific criterion:

### Ranking banking groups by business specific Size

**Comparison of RIAD group structures on its business specific size**

\( (aS_G - \text{balance sheet}) \)

<table>
<thead>
<tr>
<th>Group Identifier</th>
<th>Group Name</th>
<th>(a)Size (Bln €)</th>
<th>Size</th>
<th>New Ranking</th>
<th>Old Ranking</th>
</tr>
</thead>
<tbody>
<tr>
<td>G5</td>
<td>Banking group 5</td>
<td>&gt; 2,000</td>
<td>696</td>
<td>1</td>
<td>5</td>
</tr>
<tr>
<td>G7</td>
<td>Banking group 7</td>
<td>&gt; 2,000</td>
<td>363</td>
<td>2</td>
<td>7</td>
</tr>
<tr>
<td>G8</td>
<td>Banking group 8</td>
<td>&gt; 2,000</td>
<td>348</td>
<td>3</td>
<td>8</td>
</tr>
<tr>
<td>G230</td>
<td>Banking group 230</td>
<td>&gt; 1,000</td>
<td>137</td>
<td>4</td>
<td>-</td>
</tr>
<tr>
<td>G15</td>
<td>Banking group 15</td>
<td>&gt; 1,000</td>
<td>186</td>
<td>5</td>
<td>15</td>
</tr>
<tr>
<td>G9</td>
<td>Banking group 9</td>
<td>&gt; 500</td>
<td>348</td>
<td>6</td>
<td>9</td>
</tr>
<tr>
<td>G231</td>
<td>Banking group 231</td>
<td>&gt; 500</td>
<td>97</td>
<td>7</td>
<td>-</td>
</tr>
<tr>
<td>G1</td>
<td>Banking group 1</td>
<td>&gt; 500</td>
<td>841</td>
<td>8</td>
<td>1</td>
</tr>
<tr>
<td>G2</td>
<td>Banking group 2</td>
<td>&gt; 500</td>
<td>839</td>
<td>9</td>
<td>2</td>
</tr>
<tr>
<td>G10</td>
<td>Banking group 10</td>
<td>&gt; 100</td>
<td>298</td>
<td>10</td>
<td>10</td>
</tr>
<tr>
<td>G232</td>
<td>Banking group 232</td>
<td>&gt; 100</td>
<td>114</td>
<td>11</td>
<td>-</td>
</tr>
<tr>
<td>G233</td>
<td>Banking group 233</td>
<td>&gt; 100</td>
<td>50</td>
<td>12</td>
<td>-</td>
</tr>
<tr>
<td>G234</td>
<td>Banking group 234</td>
<td>&gt; 100</td>
<td>5</td>
<td>13</td>
<td>-</td>
</tr>
<tr>
<td>G224</td>
<td>Banking group 224</td>
<td>&gt; 100</td>
<td>246</td>
<td>14</td>
<td>-</td>
</tr>
<tr>
<td>G12</td>
<td>Banking group 12</td>
<td>&gt; 100</td>
<td>45</td>
<td>15</td>
<td>-</td>
</tr>
</tbody>
</table>

1 The above data is anonymised due to confidentiality reasons.

Sources: RIAD group data complemented with author’s own calculations

### Ranking banking group members by business specific Significance Multiplier

**Comparison of RIAD group members (excluding heads of groups) on its business specific significance multiplier**

\( (aSM_i - \text{balance sheet}) \)

1. The above data is anonymised due to confidentiality reasons.

Sources: RIAD group data complemented with author’s own calculations
Notice that with the change in the criteria, even though the pure size is a good proxy for importance (as referred in previous sections), the new criteria (Table 3) provides a different view on relevancy. A good example is the inclusion of Banking group 231 which has only 97 members but the sum of the balance sheet of its members positions it on the 7th place of the group priority ranking.

On group member level, Table 4 clearly shows the change from the baseline situation. The group members have their importance ranked in a significantly different way via the business criterion. A group movement on any of these “arms” of the group would result in a sharp decrease of the consolidated value of the network. Note that some of the members have a very low number of children below them but still an high importance. This would be the exact impact that could be missed under the baseline scenario or without a priority ranking.

In conclusion, the application of the business specific significance multiplier allows for a better approximation of the importance of movements under a specific business case. Its usage as a data quality criterion allows to identify which movements are most relevant. This is very helpful for regular data quality procedures with the aim of following-up on possible data quality issues as it prioritises and protects important group members.

4.3 Valuing changes in corporate structures

With the new international landscape, relationship data on banking groups was changed in RIAD to reflect various events. Namely, for the case of Banking Group X, there was a change in structure resulting from the actions taken during the first quarter of 2022.
Applying the SM model to this group, the analysis can be broken down by group member and give them their relative importance inside the group (base on a pre-defined criterion). This will be helpful to understand which of the movements were the most impactful.

By comparing the same group structures in different snapshots (end of January against end of May) one can find the entities that moved away from the group (marked in red in the table 5). To measure these changes, the business specific significance multiplier ($\alpha SM$) at the time before the movements (end-Jan) was collected. As in the previous example, the domestic balance sheet was used; however, any other quantitative metric could be used. The results are described in Table 5.

Naturally, the movements are always reflected in the $\alpha SM$ of the head of the group as they represent an accumulated loss. The advantage of using the $\alpha SM$ on all the group members instead of looking just into the head is that you can evaluate how relevant were each of the “movers” in the total impact in the group. In this example, the impact is only on the X5 “sub-group” and is then reflected in two entities: the X12 and X10 branches. Objectively, you could infer that the $\alpha SM$ would rank the X12 branch as the most relevant and the X10 as the least relevant. Of course, looking into the data subjectively, and as the values are so close together, one could argue that they are virtually of the same relevance. Another important point is that one could, in January, know the full amount of balance sheet value that would leave the group by summing up the $\alpha SM$ of the leaving entities. The big advantage is that any dependent entity (children of these branches) would already be accounted in the measure.

This example shows a movement in a small group but picturing a larger movement in a larger group and spread out across the structure, the $\alpha SM$ could provide the same level of insight. Additionally, if such a metric would be calculated daily, one could have an immediate response to the question: What would be the impact if a certain entity leaves its group? In the example above, what is the impact in group structure X if the X12 branch leaves the group? With the measure explained in this paper, such an answer could be given as a data point in a dataset, addressing needs from decision makers.
How did the Banking Group X structure change? (Table 5)

Evaluating the changes in Banking Group X structure using the business specific significance multiplier (αSM, domestic balance sheet)

<table>
<thead>
<tr>
<th>Entity Identifier</th>
<th>Group Name</th>
<th>Present in May</th>
<th>αSM (Mln €) (Jan)</th>
<th>αSM (Mln €) (May)</th>
<th>Total loss (αSM, Mln €)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Member X1</td>
<td>X1</td>
<td>Yes</td>
<td>&gt; 300,000</td>
<td>&gt; 300,000</td>
<td>&gt; - 2,000</td>
</tr>
<tr>
<td>Member X2</td>
<td>X2</td>
<td>Yes</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>Member X3</td>
<td>X3</td>
<td>Yes</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>Member X4</td>
<td>X4</td>
<td>Yes</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>Member X5</td>
<td>X5</td>
<td>Yes</td>
<td>&gt; 5,000</td>
<td>&gt; 2,000</td>
<td>&gt; - 2,000</td>
</tr>
<tr>
<td>Member X6</td>
<td>X6</td>
<td>Yes</td>
<td>&gt; 500</td>
<td>&gt; 500</td>
<td>-</td>
</tr>
<tr>
<td>Member X7</td>
<td>X7</td>
<td>Yes</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>Member X8</td>
<td>X8</td>
<td>Yes</td>
<td>&gt; 0.1</td>
<td>&gt; 0.1</td>
<td>-</td>
</tr>
<tr>
<td>Member X9</td>
<td>X9</td>
<td>No</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>Member X10</td>
<td>X10</td>
<td>No</td>
<td>&gt; 1,000</td>
<td>-</td>
<td>&gt; -1,000</td>
</tr>
<tr>
<td>Member X11</td>
<td>X11</td>
<td>Yes</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>Member X12</td>
<td>X12</td>
<td>No</td>
<td>&gt; 1,000</td>
<td>-</td>
<td>&gt; -1,000</td>
</tr>
<tr>
<td>Member X121</td>
<td>X121</td>
<td>No</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>Member X13</td>
<td>X13</td>
<td>Yes</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>Member X14</td>
<td>X14</td>
<td>Yes</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>Member X15</td>
<td>X15</td>
<td>Yes</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>Member X151</td>
<td>X151</td>
<td>Yes</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>Member X152</td>
<td>X152</td>
<td>Yes</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
</tbody>
</table>

1 The above data is anonymised due to confidentiality reasons.

Sources: RIAD group data complemented with author's own calculations
5. Conclusions

In this paper we propose an approach to assess the relevance of groups and their group members in a large master-dataset on entities. The method is based on the group data model of the ESCB’s Register of Institutions and Affiliates Data (RIAD), maintained in the ECB’s Directorate General Statistics in close cooperation with EU NCBs and NCAs, and allows efficiency gains in the data quality management of such information. The proposed metric:

- Provides a **prioritisation measure** for data quality management on group structures, allowing data quality managers to focus on the most relevant cases from the point of view of the impact on data users;

- Provides a **business-driven measure to evaluate** important movements inside group structures, allowing experts to know the impact of certain decisions beforehand.

This approach is currently used as a prioritisation tool for data quality management on RIAD group data and has proved valuable in correctly identifying the right priorities in data quality work, which is essential in a context of scarce resources. Besides the two use cases detailed in Section 4, the tool was tested with further business criteria and proved successful in the calculation of the significance multiplier of the group population. Also, it showed that changing the criteria has an impact on how the relevance is distributed. For future studies, one could consider showing a comparison between the results under different business criteria.

The metric also proved useful to value subsidiaries inside groups and show their relative importance. An introduction of such a measure, as a complement to a dataset, could result in efficiency gains for experts across the ECB (and other institutions) when evaluating the impact of real business changes. Another possible extension could be to use this metric to evaluate the impact of corporate actions over time or in the context of the climate change work (e.g. using the carbon footprint as a business specific criteria).

Finally, one could reflect on the possibility of applying this approach to other network data, i.e. group data collected for other statistical datasets. As this tool is based on a network dataset, any other data that has the same structure could essentially use the same technique.
## Annex

### Dataset variables and description

<table>
<thead>
<tr>
<th>Variable</th>
<th>Description</th>
<th>Original dataset</th>
</tr>
</thead>
<tbody>
<tr>
<td>Group ID</td>
<td>Unique identifier of groups in the RIAD system</td>
<td>RIAD group dataset</td>
</tr>
<tr>
<td>Group member RIAD ID</td>
<td>Unique identifier of the entities that belong to the groups</td>
<td>RIAD group dataset</td>
</tr>
<tr>
<td>Parent RIAD ID</td>
<td>Unique identifier of the entities that are parents in group structures</td>
<td>RIAD group dataset</td>
</tr>
<tr>
<td>Group member RIAD Code</td>
<td>RIAD Code of the entities that belong to the groups</td>
<td>RIAD reference data dataset</td>
</tr>
<tr>
<td>Group member Name</td>
<td>Name of the entities that belong to the groups and have a parent</td>
<td>RIAD reference data dataset</td>
</tr>
<tr>
<td>Group member ESA Sector</td>
<td>ESA Sector of the entities that belong to the groups</td>
<td>RIAD reference data dataset</td>
</tr>
<tr>
<td>Group member Domestic Balance Sheet</td>
<td>Domestic balance sheet of the entities that belong to the groups</td>
<td>RIAD reference data dataset</td>
</tr>
<tr>
<td>Size of the group</td>
<td>The group size calculated as described in section 3</td>
<td>Author’s calculations</td>
</tr>
<tr>
<td>Significance multiplier</td>
<td>The baseline significance multiplier based on pure counts for each group member</td>
<td>Author’s calculations</td>
</tr>
<tr>
<td>Business specific significance multiplier</td>
<td>The business specific significance multiplier based on the balance sheet value for each group member</td>
<td>Author’s calculations</td>
</tr>
</tbody>
</table>
References

Schnabel, I. (2020): Don’t take it for granted: the value of high-quality data and statistics for the ECB’s policymaking
Tissot, B (2019): Financial big data and policy work: opportunities and challenges
Bruno Carreiras
DG-Statistics/AMA/MAM, European Central Bank

Data quality management of entity group data
Relevance and tools to address current challenges

IFC 11th Biennial Conference
Post-pandemic landscape for central bank statistics
25-26 August 2022, Basel
# Overview

1. Context
2. RIAD Group Networks
3. Prioritising DQ, the Significance Multiplier (SM)
4. Implementation
5. Conclusions

This paper should not be reported as representing the views of the ECB. The views expressed in this paper are those of the author and do not necessarily reflect those of the European Central Bank.
• The pandemic increased the need to deliver data to decision-makers quickly and with quality.
• Managing these expectations is challenging, especially when dealing with large granular datasets as the Register of Institutions and Affiliates Data (RIAD).
• Being the ESCB shared entity master-dataset, RIAD is managed within a highly dynamic environment where attributes and relationship data can be changed every day.
• Very large datasets combined with regular changes will necessarily need measures to monitor and ensure data quality efficiently:

This paper tries to respond to that need by presenting a metric, the Significance Multiplier, applied to the RIAD Group Networks that aims at prioritising potential data quality issues by evaluating movements in group structures.
How are groups created in RIAD?

In RIAD, relationship data is collected from NCBs and NCAs (based on national data sources) and disseminated to users daily.

- Based on such relationships, group network trees are calculated each day (dynamic).

RIAD Group Networks

~300,000 ownership relationships
~130,000 groups
Available to users via a large network dataset
Annex – Users of RIAD Groups

RIAD Groups

- Banking Statistics and Balance of Payments
- Supervision and Risk Analysis
- Monetary Statistics (Climate Change)
- Macroprudential Analysis and Financial Stability
Ensuring groups’ data quality: group movements

DQM for groups is performed in two ways:
- Consistency checks on the relationship data (pre-calculation)
- Monitoring movements in group structures: real business changes vs erroneous changes (post-calculation)

In a universe of 120,000 groups that may change every day, it is important to evaluate and prioritise the monitoring and verification of movements.

Possible mistakes impact users → movements need to be assessed.

B is a leaver of Group A and Group B is generated.
**Finding the most relevant movements**

In the context of the RIAD groups: (1) movements always start in a single event on a group member and (2) if a member leaves it takes all its direct and indirect children.

At the baseline, the relevance of a group member can be given by the count of members affected by an event on that member or, **Significance Multiplier** → the size of **(sub)group headed by the affected entity**.

![Diagram of group members and their significance multipliers]

<table>
<thead>
<tr>
<th>Group member</th>
<th>Significance Multiplier</th>
</tr>
</thead>
<tbody>
<tr>
<td>A</td>
<td>9</td>
</tr>
<tr>
<td>B</td>
<td>3</td>
</tr>
<tr>
<td>C</td>
<td>4</td>
</tr>
<tr>
<td>D</td>
<td>1</td>
</tr>
<tr>
<td>E</td>
<td>1</td>
</tr>
<tr>
<td>F</td>
<td>1</td>
</tr>
<tr>
<td>G</td>
<td>2</td>
</tr>
<tr>
<td>H</td>
<td>1</td>
</tr>
<tr>
<td>I</td>
<td>1</td>
</tr>
</tbody>
</table>

Measure of relevance for group member and movements intra and cross group

Larger groups and members with more children are always considered more relevant
What is relevant for you?

RIAD Groups are used by multiple stakeholders → important to account for their priorities and needs in any data quality procedure and related DQM prioritisation.

- **Extension**: the valuation should be based on a criterion selected by the interested business area (α) rather than a pure count of the movements. The Significance Multiplier stops being the count of all the members of a subgroup but the sum of its α (αSM)

### Prioritisation ranking

For data quality is based on a defined criterion and not on a naïve count

### By-product

Users can evaluate every group member for their business purpose (can be calculated daily)
Defining a priority ranking for DQ

SM for priority ranking → relevance of groups and group members gives an indication on which movements should be assessed first.

**By group**

<table>
<thead>
<tr>
<th>Group Head</th>
<th>SM</th>
<th>αSM (Bln €)</th>
<th>SM Ranking</th>
<th>αSM Ranking</th>
</tr>
</thead>
<tbody>
<tr>
<td>Banking group 5</td>
<td>696</td>
<td>&gt;2,000</td>
<td>5</td>
<td>1</td>
</tr>
<tr>
<td>Banking group 7</td>
<td>363</td>
<td>&gt;2,000</td>
<td>7</td>
<td>2</td>
</tr>
<tr>
<td>Banking group 8</td>
<td>348</td>
<td>&gt;2,000</td>
<td>8</td>
<td>3</td>
</tr>
<tr>
<td>Banking group 17</td>
<td>137</td>
<td>&gt;1,000</td>
<td>17</td>
<td>4</td>
</tr>
<tr>
<td>Banking group 15</td>
<td>186</td>
<td>&gt;1,000</td>
<td>15</td>
<td>5</td>
</tr>
<tr>
<td>Banking group 6</td>
<td>348</td>
<td>&gt;500</td>
<td>9</td>
<td>6</td>
</tr>
<tr>
<td>…</td>
<td>…</td>
<td>…</td>
<td>…</td>
<td>…</td>
</tr>
<tr>
<td>Banking group 233</td>
<td>50</td>
<td>&gt;100</td>
<td>36</td>
<td>12</td>
</tr>
</tbody>
</table>

**By group member**

<table>
<thead>
<tr>
<th>Group Member</th>
<th>SM</th>
<th>αSM (Bln €)</th>
<th>SM Ranking</th>
<th>αSM Ranking</th>
</tr>
</thead>
<tbody>
<tr>
<td>Member 320</td>
<td>21</td>
<td>&gt;2,000</td>
<td>&gt;120</td>
<td>1</td>
</tr>
<tr>
<td>Member 11</td>
<td>194</td>
<td>&gt;2,000</td>
<td>11</td>
<td>2</td>
</tr>
<tr>
<td>Member 321</td>
<td>2</td>
<td>&gt;2,000</td>
<td>&gt;500</td>
<td>3</td>
</tr>
<tr>
<td>Member 323</td>
<td>1</td>
<td>&gt;2,000</td>
<td>&gt;500</td>
<td>4</td>
</tr>
<tr>
<td>Member 324</td>
<td>25</td>
<td>&gt;1,000</td>
<td>&gt;120</td>
<td>5</td>
</tr>
<tr>
<td>Member 325</td>
<td>5</td>
<td>&gt;1,000</td>
<td>&gt;500</td>
<td>6</td>
</tr>
<tr>
<td>…</td>
<td>…</td>
<td>…</td>
<td>…</td>
<td>…</td>
</tr>
</tbody>
</table>

Movements are given the right priority! → movements on relevant branches are assessed first

Applied to: RIAD Banking groups, end-May
Business specific criteria (α): Individual Balance Sheet
Valuation of corporate changes – the Banking Group X

X12 and X10 branches are closed in RIAD reflecting real world changes!

What is the impact of each of the changes on banking group X balance sheet?

\[ \alpha SM_{X_{12}} = \alpha_{X_{12, Jan22}} + \alpha_{X_{121, Jan22}} > 1,500 \text{ Mln } \varepsilon \]

\[ \alpha SM_{X_{10}} = \alpha_{X_{10, Jan22}} < 1,500 \text{ Mln } \varepsilon \]

Changes are ranked by relevance intra-group!

Valuation is already available before the event (Jan)!

Applied to: RIAD data on banking group X, end-Jan vs end-May
Business specific criteria (\(\alpha\)): Individual Balance Sheet
Conclusions

• The **Significance Multiplier** provides benefits in two ways:
  – Provides a **prioritisation measure** for data quality management of changes in group networks;
  – Provides an **evaluation technique** for corporate changes for assessing the impact of movements in *real time*.

• Already being used to prioritise data quality work on RIAD group networks.

• **Possible extensions:**
  – Application to other business cases: climate change (carbon intensity) or credit data;
  – Application to other network datasets under the same business conditions (dynamic environment).
Thank you! Questions?
Bruno.Carreiras@ecb.Europa.eu
Big data analytics on payment system data for measuring household consumption in Indonesia\(^1\)

Renardi Ardiya Bimantoro, Mohammad Khoyrul Hidayat, Muhammad Abdul Jabbar and Alvin Andhika Zulen,
Bank Indonesia

\(^1\) This presentation was prepared for the conference. The views expressed are those of the authors and do not necessarily reflect the views of the BIS, the IFC or the central banks and other institutions represented at the event.
Big Data Analytics on Payment System Data for Measuring Household Consumption in Indonesia

Alvin Andhika Zulen ¹, Mohammad Khoyrul Hidayat ², Muhammad Abdul Jabbar³, Renardi Ardiya Bimantoro ⁴

Abstract

Consumer spending is one of the main indicators to measure state of the economy in Indonesia. However, those data are published on a quarterly basis with a publication lag of one month. This study examines the use of retail payment system data as a proxy for household consumption indicators in Indonesia. By utilizing Big Data Analytics methodology, we are able to construct an indicator, which is available within a few days after the end of the reference period. This indicator can be used as initial proxy for household consumption, which is indicated by a good correlation with the official data.

Keywords: GDP; household consumption; payment system; big data

JEL classification: B22, C55, E21

¹ Statistics Department – Bank Indonesia; e-mail: alvin_az@bi.go.id
² Statistics Department – Bank Indonesia; e-mail: moh_khoyrul@bi.go.id
³ Statistics Department – Bank Indonesia; e-mail: muhammad_abdul@bi.go.id
⁴ Statistics Department – Bank Indonesia; e-mail: renardi_ardiya@bi.go.id

The views expressed here are those of the authors and do not necessarily reflect the views of Bank Indonesia
Contents

1. Background .................................................................................................................. 3

2. Literature Review ......................................................................................................... 4
   2.1 Utilization of Big Data for Macroeconomic Indicators .............................................. 4
   2.2 Utilization of Payment System Data ........................................................................ 5

3. Methodology ................................................................................................................ 5
   3.1 Data ............................................................................................................................ 5
   3.2 Workflow ................................................................................................................... 7
      3.2.1 Data Preprocessing ............................................................................................... 7
      3.2.2 Data Extraction .................................................................................................... 8
      3.2.3 Data Validation .................................................................................................... 8

4. Result and Analysis ..................................................................................................... 8
   4.1 Evaluation of Classification Model ............................................................................ 8
   4.2 Result Validation ....................................................................................................... 9

5. Conclusion and Future Work ...................................................................................... 11
   5.1 Conclusion ............................................................................................................... 11
   5.2 Future Work ............................................................................................................ 11

References ................................................................................................................... 12
1. Background

The growth of an economy can be measured by Gross Domestic Product (GDP) data. GDP can be calculated through three approaches, i.e. production, income, and expenditure approach. The expenditure approach is a crucial aspect for Bank Indonesia in carrying out its mandate as the monetary, macroprudential, and payment system authority.

The expenditure approach can be analysed through several indicators. One of the indicators is the household consumption expenditure indicator, which represents the spending on goods and services by resident households for final consumption purposes. Household consumption is the most significant contributor (±55%, Figure 1) to Indonesia’s GDP based on the expenditure in 2020.

Contribution of Indonesia’s GDP Components by Expenditure in 2020

![Figure 1](image)

Source: BPS

Bank Indonesia, in synergy with the government, constantly strives to formulate the appropriate policies in its 3 (three) main objectives of Bank Indonesia, which are monetary, financial system stability, and payment system. For the central bank, it is significant to know the current economic condition (state of the economy) which will be used as the basis for predicting the economy’s growth in the future. Considering Bank Indonesia’s policies are aimed at influencing the expenditure sides, Bank Indonesia needs to observe the movement of household consumption as the most significant expenditure component in Indonesia’s GDP as early as possible. However, data related to household consumption indicators in GDP are published and available on a quarterly basis with a publication lag of one month.

Nowcasting is a method used to predict the direction of the economic movement. Through nowcasting, policymakers can assess the direction of the economic movement by using representative high-frequency data to capture the dynamics of the reference indicators (i.e. GDP). Many researches related to nowcasting have been published, and several nowcasting models have been widely used to estimate various indicators (Tarsidin, Idham, & Rakhman, 2016). In addition, this method may help policymakers to formulate policy responses while waiting for the official release of macroeconomic indicators.
The Covid-19 pandemic since the beginning of 2020 has directly impacted the world economy, and Indonesia is no exception. Indonesia has fallen into its first recession in 22 years as the Covid-19 pandemic continues to take its toll. In response to this situation, policymakers need to project macroeconomic indicators to formulate appropriate policies. During this pandemic, the analysis of household consumption indicators as one of the macroeconomic indicators is very fundamental, especially in helping the central bank and government to see the growth and predict household consumption behavior.

On the other hand, technological advancement and the widespread use of cashless payment systems in the digital era have opened up the opportunities to explore large dataset of payments data for monitoring economic activity. For example, current technological advancement allow us to utilize Big Data Analytics for processing large dataset and estimating economic indicators in advance (Buono et al., 2018). This study aims to examine the use of retail payment system data, particularly from the Bank Indonesia National Clearing System (SKNBI), which has a high availability frequency, as a proxy for household consumption indicators in Indonesia.

2. Literature Review

2.1 Utilization of Big Data for Macroeconomic Indicators

Along with the technological advancement, various parties have taken the advantage of Big Data Analytics more broadly. To be more specific, near real-time and faster data processing is urgently needed, especially for supporting policy formulation during the current Covid-19 pandemic. By collecting and processing data on a large scale and high frequency, central bank can determine the current state of the economic in advance as a basis for policy formulation. In addition, literature studies related to the use of Big Data Analytics in the economic field have been developed with various methodologies.

Kapetanios & Papailias (2018) discusses the potential use of Big Data Analytics in nowcasting GDP and other macroeconomic indicators in the UK. In this study, the authors describe various initiatives related to Big Data Analytics in nowcasting macroeconomic indicators. The research also describes the benefit of using Big Data Analytics, which makes it possible to process monthly, weekly, daily, or higher frequency data on a large scale.

In another study, Buono et al. (2018) discusses GDP projection by utilizing various type of data: (i) macroeconomic data with monthly frequency, i.e., core consumer prices, consumer price index, house prices, job vacancies index; (ii) financial data with weekly frequency, i.e., Interest rates, equity indexes, and (iii) uncertainty indicators based on keyword searches in Google. In this study, the author concludes that the results of the uncertainty indicator from Big Data contribute in reducing the RMSFE (root mean squared forecast error) between the nowcasting results and the actual value.
2.2 Utilization of Payment System Data

Several studies have proven that high-frequency data, i.e., data from the payment system, can be used to estimate macroeconomic indicators. By utilizing high-frequency data, which is available faster, we can produce an earlier estimate of current economic conditions. Through this approach, the policy-making authorities benefit by being able to obtain prompt indicators for supporting policy assessment and formulation.

For example, Galbraith & Tkacz (2015) conducted an assessment in using large-scale datasets from the payment system, i.e., debit card, credit card, and cheque transactions, as a proxy for GDP growth in Canada. This study found that by using payment system data as one of the input variables can reduce the nowcasting error by 65%, compared to only using macroeconomic indicators as the input variables.

Recent research was also conducted by Dunn et al. (2020) to measure the impact of the Covid-19 pandemic on consumer spending by utilizing payment transaction data. This study shows a high correlation between official survey data and payment transaction data, especially for retail, accommodation, and restaurant sectors. In terms of data availability, the payment transaction data can be available daily with a lag of 3 (three) days, much higher in frequency when compared to data from monthly surveys that have a publication time lag of 1 (one) month. This study concludes that payment transaction data can be used as alternative data and initial proxy for consumption indicators.

Thus, this study is expected to answer the following research questions:

1. Can payment system data be used as a data source to measure household consumption indicators in Indonesia?
2. Can the resulting household consumption indicators complement the existing indicators?

3. Methodology

3.1 Data

The data source used in this study was obtained from the National Clearing System of Bank Indonesia (SKNBI). This National Clearing System of Bank Indonesia (SKNBI) is a Retail Value Payment System (RVPS) infrastructure operated by Bank Indonesia to process electronic financial data for fund transfer services, debit clearing services, regular payment services, and regular billing services (Regulation of Member of Board of Governors Number 21/12/PADG/2019). Since September 2019, SKNBI has been able to process payment transactions with less than Rp. 1,000,000,000.00 (one billion rupiah) in amount.

The scope of data used in this study is SKNBI fund transfer transaction data from July 2015 to December 2021. Fund transfer service is a service within SKNBI that facilitates the transfer of funds between participant banks, with an average number of transactions reaching ±13 million transactions per month. In the fund transfer service, there are several types of transaction:

1. 50: Transfer of funds between participants on behalf of the customer;
2. 51: Transfer of funds between participants related to Government’s Treasury Single Account (TSA);
3. 52: Transfer of funds between participants and Bank Indonesia’s Treasury Single Account;
4. 53: Transfer of funds between participants that is not for the customers’ needs;
5. 54: Transfer of funds between participants on behalf of the customer without accounts;
6. 55: Transfer of funds between participants on behalf of the customer related to money remittance; and
7. 59: Refund of fund transfers and payments (reversal).

The share for each transaction type is shown in Figure 2, which shows that the largest share of transaction (±92%) is fund transfer transactions between participants on behalf of the customer (code 50).

Nominal and Frequency of SKNBI Transactions Based on Transaction Type

![Figure 2](source)

Source: SKNBI (processed)

The data structure obtained from the fund transfer service transactions is as shown in Table 1. Although the data structure of the SKNBI fund transfer is quite comprehensive, there are still issues related to the data validity. For example, we found that the customer’s type code does not always match the customer’s name category, both in the sender and recipient fields.

SKNBI Fund Transfer Data Structure

<table>
<thead>
<tr>
<th>No.</th>
<th>Data Field</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>DKE ID</td>
</tr>
<tr>
<td>2</td>
<td>BATCH ID</td>
</tr>
<tr>
<td>3</td>
<td>TRANSACTION DATE</td>
</tr>
<tr>
<td>4</td>
<td>SENDER BANK CODE</td>
</tr>
<tr>
<td>5</td>
<td>SENDER LOCATION</td>
</tr>
<tr>
<td>6</td>
<td>BENEFICIARY BANK CODE</td>
</tr>
<tr>
<td>7</td>
<td>BENEFICIARY LOCATION</td>
</tr>
<tr>
<td>8</td>
<td>AMOUNT</td>
</tr>
<tr>
<td>9</td>
<td>TRANSACTION TYPE CODE</td>
</tr>
<tr>
<td>10</td>
<td>SENDER CUSTOMER’S NAME</td>
</tr>
<tr>
<td>11</td>
<td>SENDER CUSTOMER’S ACCOUNT NUMBER</td>
</tr>
</tbody>
</table>
3.2 Workflow

In constructing household consumption indicators from SKNBI fund transfer data, we develop a text mining model with a rule-based approach for processing unstructured information, backed up by parallel computing technology in Apache Spark – Hadoop. We use Python as the programming language. In general, the workflow in this study consists of data preprocessing, data extraction, and data validation.

3.2.1 Data Preprocessing

The data preprocessing stage is carried out to prepare the raw SKNBI fund transfer transaction data so that they can be further processed at the next stage. The process is as follows:

1. **Filter transactions that are not on behalf of customers.**
   The data that will be further processed is only data with transaction type code 50 (transfer of funds between participants on behalf of the customer) and 54 (transfers of funds on behalf of the customer without accounts).

2. **Classification of customers’ categories.**
   For each transaction, we classify sender and beneficiary customers into business entities, governments, and others. This process is critical since customers’ type code has validity issues. Classification is conducted using a rule-based approach with rules as shown in Table 3.

3. **Filter transactions that do not have a description.**

<table>
<thead>
<tr>
<th>No.</th>
<th>Data Field</th>
</tr>
</thead>
<tbody>
<tr>
<td>12</td>
<td>SENDER CUSTOMER’S ADDRESS</td>
</tr>
<tr>
<td>13</td>
<td>SENDER CUSTOMER’S ID NUMBER</td>
</tr>
<tr>
<td>14</td>
<td>SENDER CUSTOMER’S TYPE</td>
</tr>
<tr>
<td>15</td>
<td>BENEFICIARY CUSTOMER’S NAME</td>
</tr>
<tr>
<td>16</td>
<td>BENEFICIARY CUSTOMER’S ACCOUNT NUMBER</td>
</tr>
<tr>
<td>17</td>
<td>BENEFICIARY CUSTOMER’S ADDRESS</td>
</tr>
<tr>
<td>18</td>
<td>BENEFICIARY CUSTOMER’S ID NUMBER</td>
</tr>
<tr>
<td>19</td>
<td>BENEFICIARY CUSTOMER’S TYPE</td>
</tr>
<tr>
<td>20</td>
<td>DESCRIPTION</td>
</tr>
</tbody>
</table>

### Rules for Classification of Customers’ Categories

<table>
<thead>
<tr>
<th>CUSTOMER CLASSIFICATION</th>
<th>SAMPLE KEYWORDS</th>
</tr>
</thead>
<tbody>
<tr>
<td>Others (individual)</td>
<td>Does not contain business entities and governments keywords.</td>
</tr>
</tbody>
</table>

Table 3
3.2.2 Data Extraction

There is limited information on the description of fund transfer transactions in SKNBI, in which there is no standard format/reference code for the information written in the description field (free text). We develop a text mining model to analyze the transaction data to handle this issue. The resulting data from the previous stage (section 3.2.1) are used as input for this stage with the following process:

1. **Classification of the purposes of SKNBI fund transfer transactions.**

   Classification is done using a rule-based approach based on predefined keywords, as shown in Table 4.

2. **Data aggregation.**

   After classifying the purpose of the transaction, data can be aggregated as indicators for each transaction purposes, e.g. household consumption, household income, and business. As for transactions with transfer purposes other than those three categories are classified as “Others”.

   **Table 4: Rules for Classification of Transaction Purposes**

<table>
<thead>
<tr>
<th>SENDER</th>
<th>BENEFICIARY</th>
<th>KEYWORD IN TRANSACTION DESCRIPTION</th>
<th>CLASSIFICATION</th>
</tr>
</thead>
<tbody>
<tr>
<td>Business Entities</td>
<td>Business Entities</td>
<td></td>
<td>Business</td>
</tr>
</tbody>
</table>

3.2.3 Data Validation

The resulting indicators from the previous stage (section 3.2.2) are then validated with the GDP data - Household Consumption (current prices) as the reference indicator. The monthly indicators from the SKNBI are converted into quarterly data by accumulating the nominal amount of transactions in each quarter. This step is required to obtain indicators with the same frequency as household consumption data in GDP. After that, validation is conducted by calculating the correlation value between those two data.

4. Result and Analysis

4.1 Evaluation of Classification Model

After we develop the classification model with the rule-based approach in the previous section, we need to evaluate our model to find out how accurate the model
is in classifying the transaction purposes, particularly for consumption. In this study, we use F1-score\(^5\) as an evaluation metric. The evaluation was carried out on ±4,000 transactions (random sampling) during the period of 2018 to 2021.

The evaluation results in Table 5 show us a good value of the overall F1-score (84.5%). However, the F1-score for predicting consumption transaction is still relatively low compared to the results for other categories. If we analyze using the confusion matrix in Table 6, there are still plenty of false positive cases, i.e. transactions predicted to be "consumption" but should be included in other categories. We suspect that the prediction error could be caused by the accuracy of the customer categories classification, which still need to be improved.

<table>
<thead>
<tr>
<th>TRANSACTION PURPOSES</th>
<th>RECALL</th>
<th>PRECISION</th>
<th>F1-SCORE</th>
<th>AVERAGE OF F1-SCORE</th>
<th>ACCURACY</th>
</tr>
</thead>
<tbody>
<tr>
<td>Consumption</td>
<td>80.2%</td>
<td>62.6%</td>
<td>70.3%</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Income</td>
<td>85.4%</td>
<td>98.8%</td>
<td>91.6%</td>
<td>84.5%</td>
<td>88.3%</td>
</tr>
<tr>
<td>Business</td>
<td>75.6%</td>
<td>98.2%</td>
<td>85.4%</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Others</td>
<td>97.4%</td>
<td>84.5%</td>
<td>90.5%</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

4.2 Result Validation

As previously explained in section 3.2.3, the household consumption indicator from the SKNBI (growth, y.o.y) is validated with the GDP-household consumption indicator at current prices (growth, y.o.y). The correlation of the two indicators can be seen in Table 7 and the graph visualization in Figure 4.

The validation results show a high correlation between the two indicators since 1\(^{st}\) quarter of 2019, including during the Covid-19 pandemic. These results indicate that consumption indicators from SKNBI transaction data can be used as a proxy for household consumption. Moreover, it can be available earlier, i.e. 2 (two) days lag after the end of the period, both weekly and monthly. The availability of this indicator

\[^{5}\text{F1} : 2 \times \text{precision} \times \text{recall} / (\text{precision} + \text{recall})\]

\[^{6}\text{Precision} : (\text{true positive}) / (\text{true positive} + \text{false positive})\]

\[^{7}\text{Recall} : (\text{true positive}) / (\text{true positive} + \text{false negative})\]
is much faster than the publication of GDP data which has a time lag of more than 1 (one) month.

## Correlation Table

### Table 7

<table>
<thead>
<tr>
<th>INDICATORS</th>
<th>Q1-2018 to Q4-2021</th>
<th>Q1-2019 to Q4-2021</th>
<th>Q1-2020 to Q4-2021</th>
</tr>
</thead>
<tbody>
<tr>
<td>Total of Household Consumption</td>
<td>55.2%</td>
<td>85.2%</td>
<td>91.2%</td>
</tr>
<tr>
<td>Clothes, Health &amp; Education, Restaurant &amp; Hotel, and Others</td>
<td>61.3%</td>
<td>85.2%</td>
<td>90.4%</td>
</tr>
<tr>
<td>Health &amp; Education</td>
<td>67.7%</td>
<td>93.3%</td>
<td>90.3%</td>
</tr>
<tr>
<td>Restaurant &amp; Hotel</td>
<td>65.9%</td>
<td>83.4%</td>
<td>60%</td>
</tr>
</tbody>
</table>

Source: SKNBI, BPS (processed)

### SKNBI Consumption Growth and Household Consumption Growth – GDP (percent, y.o.y)

Source: SKNBI, BPS (processed)
5. Conclusion and Future Work

5.1 Conclusion

In this study, we have proposed a new approach in utilizing payment system transaction data as a proxy for household consumption indicators. Using text mining methodology with rule-based model, we can classify customer categories and transaction purposes from the SKNBI fund transfer data. Based on the evaluation of the model, the average F1-score of the model is 84.5%.

Using this methodology, we can obtain a proxy for household consumption indicators from high-frequency payment system data more quickly, compared to the official publication of GDP data. The validation results show a high correlation between these two indicators, which indicates that the household consumption indicator from the SKNBI fund transfer data can be used as a proxy for household consumption indicators.

5.2 Future Work

There are several improvements in the methodology that can be applied for future works.

1. Improving the methodology for classifying customers' categories and transaction purposes, including the use of machine learning algorithms.

2. Utilizing Bank Indonesia’s Fast Payment (BI-FAST) transaction data (implemented since the end of 2021) as additional data source in constructing consumption indicators from retail payment system.

3. Using consumption indicators from the payment system, e.g. fund transfers from SKNBI and payment transactions via cards, with other macroeconomic variables, to construct the nowcasting model of household consumption.
References


Big Data Analytics on Payment System Data for Measuring Household Consumption

Alvin Andhika Zulen, Mohammad Khoyrul Hidayat, Muhammad Abdul Jabbar, Renardi Ardiya Bimantoro
Statistics Department – Bank Indonesia
Email: alvin_az@bi.go.id, moh_khoyrul@bi.go.id, muhammad_abdul@bi.go.id, renardi_ardiya@bi.go.id

August 2022

The views expressed here are those of the authors and do not necessarily reflect the views of Bank Indonesia
OUTLINE

1. Background
2. Data Source
3. Methodology
4. Result & Analysis
5. Conclusion
Household consumption is one of the main indicators to measure state of the economy in Indonesia (largest contributor, 55%, in Indonesia’s GDP). However, GDP data (incl. household final consumption expenditure) are published and available on a quarterly basis with a publication lag of one month.

Bank Indonesia provides retail value payment system, SKNBI (The National Clearing System), that can generate data related to fund transfers, including household transactions.

Advancements of technology and widespread use of payment systems have opened the opportunity to explore large dataset of payment data for monitoring economic activity.

**OBJECTIVE**

Developing a high frequency measure of household consumption in Indonesia from retail value payment system data (SKNBI), by utilizing Big Data Analytics methodology, particularly text mining.
Fund Transfer of SKNBI:
Credit transfer transaction between participants (banks) on behalf of the customers.

- **Total Transactions**: ≅ 13 mio trx/month
- **Nominal Transactions**: ≤ Rp. 1 Billion/trx
- **Availability Period**: July 2015 s.d. December 2021

### COLUMN NAME

<table>
<thead>
<tr>
<th></th>
<th>COLUMN NAME</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>DKE ID</td>
</tr>
<tr>
<td>2</td>
<td>BATCH ID</td>
</tr>
<tr>
<td>3</td>
<td>TRANSACTION DATE</td>
</tr>
<tr>
<td>4</td>
<td>ORIGINATING BANK CODE</td>
</tr>
<tr>
<td>5</td>
<td>SENDER LOCATION</td>
</tr>
<tr>
<td>6</td>
<td>BENEFICIARY BANK CODE</td>
</tr>
<tr>
<td>7</td>
<td>RECEIVER LOCATION</td>
</tr>
<tr>
<td>8</td>
<td>AMOUNT</td>
</tr>
<tr>
<td>9</td>
<td>TRANSACTION TYPE CODE</td>
</tr>
<tr>
<td>10</td>
<td>SENDER CUSTOMER'S NAME</td>
</tr>
<tr>
<td>11</td>
<td>SENDER CUSTOMER'S ACC NUMBER</td>
</tr>
<tr>
<td>12</td>
<td>SENDER CUSTOMER'S ADDRESS</td>
</tr>
<tr>
<td>13</td>
<td>SENDER CUSTOMER'S ID NUMBER</td>
</tr>
<tr>
<td>14</td>
<td>SENDER CUSTOMER'S TYPE CODE</td>
</tr>
<tr>
<td>15</td>
<td>BENEFICIARY CUSTOMER'S NAME</td>
</tr>
<tr>
<td>16</td>
<td>BENEFICIARY CUSTOMER'S ACC NUMBER</td>
</tr>
<tr>
<td>17</td>
<td>BENEFICIARY CUSTOMER'S ADDRESS</td>
</tr>
<tr>
<td>18</td>
<td>BENEFICIARY CUSTOMER'S ID NUMBER</td>
</tr>
<tr>
<td>19</td>
<td>BENEFICIARY CUSTOMER'S TYPE CODE</td>
</tr>
<tr>
<td>20</td>
<td>DESCRIPTION</td>
</tr>
</tbody>
</table>
METHODOLOGY

OVERALL WORKFLOW

Data Acquisition

1. Data Preprocessing
   - Filtering type of transaction.
   - Classification of customer categories into household and business entities.
   - Removing any transactions without description.

2. Data Extraction
   - Classification of transactions into consumption, income, and other transactions (using rule-based*).
   - Data aggregation and constructing consumption and income indicator.

3. Data Validation
   Validation with official indicators (i.e. GDP – household consumption)

*) e.g.: Consumption is SKNBI Fund Transfer with transaction detail containing keywords related to household consumption, e.g.: ‘buy’, ‘shop’, ‘pay’, ‘paid off’, ‘installments’, etc.
We use rule-based (keyword) approach for classifying customer categories.

Example:

<table>
<thead>
<tr>
<th>Sender</th>
<th>Sender Category</th>
<th>Beneficiary</th>
<th>Beneficiary Category</th>
</tr>
</thead>
<tbody>
<tr>
<td>INDO PRIMA SEMESTA</td>
<td>Business Entity</td>
<td>PT. UNITED FAMILY FOOD</td>
<td>Business Entity</td>
</tr>
<tr>
<td>RPKBUNP. SPAN BNI</td>
<td>Government</td>
<td>CV. TANJUNG AGUNG</td>
<td>Business Entity</td>
</tr>
<tr>
<td>TRI AMALIA</td>
<td>Others</td>
<td>PT. MAJU MOBILINDO</td>
<td>Business Entity</td>
</tr>
</tbody>
</table>
We use rule-based (keyword) approach for classifying transaction purpose (consumption, income, business)

**Household Income Keywords**
- 'gaji', 'honor', 'upah', 'payroll', 'salary', 'remunerasi', 'insentif', 'wage', 'sales', 'pensiun', 'lembur', 'overtime', 'dividen', 'kompensasi', 'bagi hasil', 'bonus', 'claim', 'claim', 'payoneer', 'komisi', 'tukin', 'uang makan'

**Household Consumption Keywords**

**Keywords that are not related to consumption/income**
- 'retur', 'return', 'tabungan', 'refund', 'saving', 'reimburse', 'pemindahbukan', 'nabung', 'tsa', 'span', 'pemerintah', 'pajak', 'sp2d', 'sppd', 'pendes', 'dana bos'
In this study, we use F1-score as an evaluation metric. The evaluation was carried out on ±4,000 transactions (random sampling) during the period of 2018 to 2021. **The evaluation results show us a good value of the overall F1-score (84.5%).**

### Confusion Matrix Evaluation Result

<table>
<thead>
<tr>
<th>Purpose of Transaction</th>
<th>Actual</th>
<th>Prediction</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Consumption</td>
<td>Income</td>
</tr>
<tr>
<td>Consumption</td>
<td>154</td>
<td>0</td>
</tr>
<tr>
<td>Income</td>
<td>17</td>
<td>569</td>
</tr>
<tr>
<td>Business</td>
<td>42</td>
<td>0</td>
</tr>
<tr>
<td>Others</td>
<td>33</td>
<td>7</td>
</tr>
</tbody>
</table>

### F1-score Calculation

<table>
<thead>
<tr>
<th>Purpose of Transaction</th>
<th>Recall</th>
<th>Precision</th>
<th>F1-score</th>
<th>F1-score (average)</th>
<th>Accuracy</th>
</tr>
</thead>
<tbody>
<tr>
<td>Consumption</td>
<td>80,2%</td>
<td>62,6%</td>
<td>70,3%</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Income</td>
<td>85,4%</td>
<td>98,8%</td>
<td>91,6%</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Business</td>
<td>75,6%</td>
<td>98,2%</td>
<td>85,4%</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Others</td>
<td>97,4%</td>
<td>84,5%</td>
<td>90,5%</td>
<td>84,5%</td>
<td>88,3%</td>
</tr>
</tbody>
</table>
The validation results show a high correlation between the two indicators since quarter 1-2019, including during the Covid-19 pandemic.

<table>
<thead>
<tr>
<th>Indicators</th>
<th>Correlation of SKNBI Consumption Growth Rate with GDP</th>
</tr>
</thead>
<tbody>
<tr>
<td>Total of Household Consumption</td>
<td>Q1-2018 s.d. Q4-2021 / Q1-2019 s.d. Q4-2021 / Q1-2020 s.d. Q4-2021</td>
</tr>
<tr>
<td>Clothes, Health &amp; Education, Restaurant &amp; Hotel, and others</td>
<td>55,2% / 85,2% / 91,2%</td>
</tr>
<tr>
<td>Clothes, Health &amp; Education and others</td>
<td>61,3% / 85,2% / 90,4%</td>
</tr>
<tr>
<td>Health &amp; Education</td>
<td>67,7% / 93,3% / 90,3%</td>
</tr>
<tr>
<td>Restaurant &amp; Hotel</td>
<td>65,9% / 83,4% / 60%</td>
</tr>
</tbody>
</table>

The validation results show a high correlation between the two indicators since quarter 1-2019, including during the Covid-19 pandemic.
Conclusion

1. We have **proposed a new approach** in utilizing high-frequency payment system transaction data, **using text mining methodology through a rule-based model**, to construct a proxy indicator for household consumption. **Based on the evaluation of the model, the average F1-score of the model is 84.5%**.

2. Our consumption indicator can be **generated from payment system data more quickly** compared to household consumption indicators in GDP publications. The validation results show a **high correlation** between our consumption indicator from payment system and publication of GDP data, which indicates that the indicator from payment system can be used as a **proxy for household consumption indicators**.

Future Works

1. **Improving the methodology for classifying customer categories and transaction purposes**, including the use of machine learning algorithms.

2. **Utilizing Bank Indonesia’s Fast Payment (BI-FAST) transaction data** (implemented since the end of 2021) as additional data source in constructing consumption indicators from retail payment system.

3. Using consumption indicators from payment system, e.g. funds transfer from SKNBI customers or payment transactions via cards, with other macroeconomic variables, to **construct the nowcasting model of household consumption**.
Quantitative analysis of haircuts: evidence from the Japanese repo and securities lending markets

Kana Sasamoto and Kazuya Suzuki, Bank of Japan

---

1 This presentation was prepared for the conference. The views expressed are those of the authors and do not necessarily reflect the views of the BIS, the IFC or the central banks and other institutions represented at the event.
Quantitative Analysis of Haircuts: Evidence from the Japanese Repo and Securities Lending Markets

Kazuya Suzuki and Kana Sasamoto

Abstract

Given the absence of comprehensive studies on market structure and haircuts for repo and securities lending transactions, this study provides a quantitative analysis of the subject using government bonds and equities transaction data covering most of the Japanese market. Specifically, we conducted a panel data regression analysis of government bond repo transactions, controlling for factors such as transaction entities and transaction types, and provided a detailed analysis of the haircut-setting mechanism. Accordingly, we determined that explanatory variables affecting credit risk, market risk, and liquidity risk, such as the credit quality of government bonds, the residual maturity of government bonds, and the presence of foreign exchange risk, significantly impact haircut setting. Furthermore, financial institutions closer to the center of the network, which engage in transactions with additional financial institutions, tend to set lower haircut rates through more efficient matching of borrowing and lending needs for cash and securities. Thus, the credit quality of government bonds transacted, exchange rate stability, and the presence of intermediaries important to the trading network significantly impact the degree of market functioning. The results were robust, paving the way for further discussions on trends and risk management of securities financing transactions, which are essential to financial markets.

Keywords: Securities Financing Transactions; Repurchase Agreement; Haircut; Network Analysis

JEL classification: D80, E43, G10, G20, L14

---

1 This paper was prepared for the eleventh IFC conference on Post-pandemic landscape for central bank statistics, 25-26 August 2022, and was originally released in Bank of Japan Working Paper Series. The authors are grateful to Akira Otani, Takayuki Kambara, Teppei Nagano, Takashi Nagahata, Kenji Fujita, Daisuke Miyakawa, and colleagues at the Bank of Japan for comments and discussions. Views expressed in the paper are those of the authors and do not necessarily reflect those of the Bank of Japan.

2 Financial Markets Department (currently, Secretariat of the Policy Board), Bank of Japan, 2-1-1 Nihonbashi-Hongokuchō, Chuo-ku, Tokyo 103-8660, Japan (e-mail: kazuya.suzuki@boj.or.jp)

3 Financial Markets Department, Bank of Japan, 2-1-1 Nihonbashi-Hongokuchō, Chuo-ku, Tokyo 103-8660, Japan (e-mail: kana.sasamoto@boj.or.jp)
Contents

Quantitative Analysis of Haircuts: Evidence from the Japanese Repo and Securities Lending Markets .................................................................................................................. 1

1. Introduction .................................................................................................................. 3

2. Haircuts in Securities Financing Transactions ............................................................. 4

3. Data .................................................................................................................................. 6

4. Panel Data Analysis ........................................................................................................ 20

5. Conclusion ...................................................................................................................... 33

References .......................................................................................................................... 34
1. Introduction

Securities Financing Transactions (SFTs) refer to transactions where cash and securities are exchanged with a counterparty and returned after a certain period. They are of two types: repurchase agreements and lending transactions. SFTs are secured and combine the characteristics of lending and borrowing cash and securities, making them the primary means for financial institutions to exchange cash and securities in the short term (Duffie, 1996; Baba and Inamura, 2004; Kinugasa and Nagano, 2017).

However, SFTs were also key to the risk-taking that induced the 2007–2009 financial crisis, during which the funding environment for financial institutions rapidly deteriorated as the haircut rate (i.e., the multiplier used to discount collateral) was raised through bilateral transactions in the US market (funding was actively raised using high-risk securitized products and other instruments). With asset prices also falling sharply, the sale of financial asset holdings accelerated in response to counterparty demands for additional collateral—margin calls—inducing a spiral of falling asset prices and higher haircut rates (Brunnermeier and Pedersen, 2009; Gorton and Metrick, 2012). Thus, many hedge funds defaulted in the fall of 2008 (Adrian et al., 2014). In tri-party transactions, money market funds (MMFs), the main cash lenders, sharply curtailed transactions with financial institutions with a high potential to default, contributing to the cash crunch at Bear Stearns and Lehman Brothers (Copeland et al., 2014). Moreover, during the European debt crisis, the haircut rate on government bonds was raised in response to a significant drop in prices, inducing a decline in liquidity in the repo market (European Securities and Markets Authority, 2016; Boissel et al., 2017).

Evidently, SFTs significantly impact leverage build-up in the financial system. Specifically, large fluctuations in haircut rates reduce market function by impeding the smooth lending of cash and securities and contribute to business cycle fluctuations through increased or decreased leverage (i.e., procyclicality) (Financial Stability Board [FSB], 2014). Accordingly, studies have examined the theoretical aspects of the mechanism of changes in haircuts. However, many note that underdeveloped data collection of SFTs impedes empirical research. Therefore, prior empirical research employs limited data and focuses mainly on the financial crisis period. In this context, the G20 and FSB initiated discussions on regulation and supervision, including the collection of SFT data. The FSB’s November 2015 report recommended that national authorities collect transaction data on individual trading units on a monthly or more frequent basis (FSB, 2015a) and introduce regulations to establish minimum haircut floors for non-centrally cleared SFTs (FSB, 2015b). Later, the Financial Services Agency (FSA) of Japan and the Bank of Japan established a framework for collecting SFT data from financial institutions in Japan and began collecting in January 2019. Since January 2020, monthly aggregates of data portions have been published on the Bank’s website (Ono et al., 2015; Sasamoto et al., 2020).

---

4 See Adrian et al. (2014), Shimamura et al. (2017), Julliard et al. (2019), and Gorton et al. (2020).
5 Minimum haircut floors are set for each transacted security and residual maturity to redemption period. Transactions using sovereign bonds, such as government bonds, are excluded.
As a result, this study employs such comprehensive transaction data to reveal the market structure and standard mechanisms for transaction haircut settings using government bonds and equities. Specifically, we regressed various explanatory variables that could not be analyzed in previous studies due to data limitations, in conjunction with entity, collateral type, and time fixed effect on the haircut rate. The results are generally robust and of value to financial authorities and practitioners in trading and risk management of SFTs at financial institutions.

The remainder of the paper is structured as follows: Section 2 reviews the definition and role of haircuts in SFTs as presented in prior studies. Section 3 outlines the study data and explains the market structure of transactions using government bonds and equities, respectively. Section 4 presents the regression analysis method using panel data and reports the analysis results for government bond repo transactions. Finally, Section 5 concludes the study and discusses the scope for future research.

2. Haircuts in Securities Financing Transactions

2-1 Definition of Haircut Rate

The haircut rate in SFTs is the multiplier used to calculate the collateral value. We consider the effective haircut rate consistent with the FSB (2021) definition in view of the haircut rate as the multiplier used to calculate the valuation of securities in repo transactions. The study also considers the “ratio for calculating cash collateral,” a multiplier for securities used along with the multiplier to calculate the valuation of cash and other securities in securities lending transactions (see Figure 1).

2-2 Role of Haircuts in Securities Financing Transactions

As noted, SFTs are secured. For instance, regarding secured bank loans, collaterals preserve the rights of the creditor in the event of default and help reduce “information asymmetry.” In other words, even if a borrower’s credit profile appears inadequate from the creditor bank’s perspective, collateral is pledged with a haircut per the collateral liquidity and creditworthiness. Thus, counterparties are screened so that business is conducted with only counterparties that agree to the terms.

---

7 Generally, it refers to securities in the case of repurchase agreements and cash or collateral securities in the case of securities lending transactions.

8 Typical markets in which “information asymmetry” occurs include those for medical insurance (Arrow, 1963) and the used car (Akerlof, 1970) market. Typical methods of addressing information asymmetry include “screening” by information-disadvantaged parties and active disclosure by information-advantaged parties (signaling).

9 Bester (1985) shows that in secured bank loans, borrowers with a low (high) probability of default are more (less) likely to accept higher haircuts to reduce interest rates because they have sufficient liquidity to pledge collateral, and haircuts help banks screen borrowers.
Hence, the role of haircuts in SFTs is analogous to bank loans; however, SFTs are somewhat more complex because they combine the characteristics of lending and borrowing cash and securities. Beyond transactions with a cash lending aspect—General Collateral (GC) transactions—as in general bank loans, transactions with a securities lending aspect—Special Collateral (SC) transactions—to cover short positions and fails are also actively conducted.\(^5\) The consequences of such characteristics are as follows: First, the lender of cash is the real creditor in GC transactions as well as the lender of securities in SC transactions. Therefore, haircuts are set for cash collateral in the event of default by a borrower of cash in GC transactions and for an increase in security prices (cost of repurchasing the securities and rebuilding the position) in the event of default by a borrower of securities in SC transactions. Thus, per the rates defined in Figure 1, the haircut rate for SC transactions is smaller than that for GC transactions and can even be negative in some cases (Bank for International Settlements [BIS], 2010; Baklanova et al., 2019; Gottardi et al., 2019).

---

**Haircut for repurchase agreements and securities lending**

**Definition of Haircut Rate**

Panel A: Haircut rate for repurchase agreements

- **α**: Haircut rate \(\times (1 - \alpha)\)

<table>
<thead>
<tr>
<th>Collateral (Government bonds etc.)</th>
<th>Cash</th>
</tr>
</thead>
<tbody>
<tr>
<td>100</td>
<td>98</td>
</tr>
</tbody>
</table>

Haircut rate \(\alpha\)

\[
\alpha = 1 - \frac{\text{Amount of cash}}{\text{Market value of collateral (government bonds etc.)}}
\]

\[
= 1 - \frac{98}{100} = +2.00\%
\]

Panel B: Haircut rate for securities lending transactions

- **β**: Effective haircut rate \(\times (1 - \beta)\)

<table>
<thead>
<tr>
<th>Securities</th>
<th>Collateral (Cash etc.)</th>
</tr>
</thead>
<tbody>
<tr>
<td>102</td>
<td>98</td>
</tr>
</tbody>
</table>

Effective haircut rate \(\beta\)

\[
\beta = 1 - \frac{\text{Market value of collateral (cash etc.)}}{\text{Market value of securities}}
\]

\[
= 1 - \frac{98}{102} = +3.92\%
\]

\(^5\) Repurchase agreements and securities lending transactions are economically equivalent but are often distinguished because there is a difference in legal ownership. The figure illustrates GC transactions; in SC transactions, the “effective haircut rate” can be negative because the market value of cash exceeds that of the securities.

---

\(^6\) Securities financing transactions may experience temporary delays in settlement (i.e., fails). For more information on the occurrence of fails, please refer to the Bank of Japan’s monthly publication “Basic Figures on Fails” (https://www.boj.or.jp/en/statistics/set/bffail/index.htm/). For more information on practices surrounding fails in Japan’s securities financing transactions, see Kasai et al. (2001).
2-3 Prior Studies

Prior studies focus on theoretical aspects of the factors affecting haircuts in SFTs, holding that they are generally explained by credit and market risk (Martin et al., 2014; Gottardi et al., 2019) and liquidity risk (Brunnermeier and Pedersen, 2009; Martin et al., 2014; Parlatore, 2019). Others consider counterparty risk (Dang et al., 2013; Gottardi et al., 2019) or operational risks regarding the efficiency of the non-defaulting party in margin management and custody (International Capital Market Association, 2012). In the US market, studies examine the differences in haircut-setting mechanisms between bilateral and tri-party transactions.¹¹

However, to the best of the author’s knowledge, most empirical studies were conducted using central counterparty (CCP)-cleared transaction and tri-party transaction data, for which data are comparatively easy to collect, or limited bilateral transaction data.¹² For the US market, Copeland et al. (2014) use tri-party transaction data collected by the Federal Reserve Bank of New York. Additionally, Baklanova et al. (2019) use bilateral transaction data collected on a pilot basis by the Office of Financial Research and the Federal Reserve Board, and Gorton et al. (2020) use transaction data for the Emergency Facility introduced by the Federal Reserve Board during the financial crisis. For the UK market, Julliard et al. (2019) use data from six major financial institutions collected by a financial authority.

3. Data

3-1 Data Sources

Since January 2019, this study has used granular transaction data collected by the FSA and the Bank of Japan from financial institutions in Japan. Detailed information is recorded for each transaction that is outstanding at the end of each month. Our data covers the parties included in the transaction (lender and borrower of securities), type of securities traded, market value of cash and securities traded, transaction maturity, repo rate, haircut rate, whether a transaction is a bilateral or agency-intermediated transaction, and whether a transaction is CCP-cleared or -uncleared. However, there are certain limitations: data on the issues of securities traded in repo and securities

¹¹ Copeland et al. (2014) show that during the financial crisis, haircut rates were lower and more stable in tri-party transactions than bilateral transactions. Hu et al. (2021) indicate that haircut and repo rates are almost unaffected by counterparties in tri-party transactions conducted by US MMFs. However, Auh and Landoni (2015) demonstrate that transaction maturity and collateral quality (credit quality and liquidity) significantly affect the haircut rate in bilateral transactions conducted by hedge funds.

¹² A CCP intermediates between parties to a financial transaction (obtaining and assuming claims and obligations) and the counterparty to the settlement. Beyond netting the claims and obligations assumed in settlement, the system serves as a guarantee for participants by fulfilling the obligations assumed from the relevant participants in the event of default on settlement by a participant.

¹³ “Agency-intermediated transactions” are where a third-party financial institution mediates the parties to a securities financing transaction to provide services regarding the management of the collateral exchanged between the parties to the transaction.
lending transactions is absent, and data coverage of information about securities in securities lending transactions is low (Table 1).

The data is reported by approximately 50 top financial institutions in terms of transaction amount, selected to capture more than 90% of SFTs to which the institutions in Japan (including overseas financial institutions based in Japan) are a party. Thus, this data has high coverage and contains detailed information on transactions, including bilateral and non-cleared transactions, which are typically challenging to ascertain.

### Summary of transaction information on collected data

<table>
<thead>
<tr>
<th>Series</th>
<th>Repo</th>
<th>Securities Lending</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Counterparty</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>- Counterparty name</td>
<td>○</td>
<td>○</td>
</tr>
<tr>
<td>- Counterparty jurisdiction (pure locational approach)</td>
<td>○</td>
<td>○</td>
</tr>
<tr>
<td>- Counterparty sector</td>
<td>○</td>
<td>○</td>
</tr>
<tr>
<td><strong>Security</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>- Security issue (e.g., ISIN code)</td>
<td>×</td>
<td>×</td>
</tr>
<tr>
<td>- Security type (e.g., Government Bonds, Equities, Corporate Bonds)</td>
<td>○</td>
<td>○</td>
</tr>
<tr>
<td>- Market value</td>
<td>○</td>
<td>○</td>
</tr>
<tr>
<td>- Jurisdiction of the issuer of the underlying security</td>
<td>○</td>
<td>×</td>
</tr>
<tr>
<td>- Currency</td>
<td>○</td>
<td>○</td>
</tr>
<tr>
<td>- Residual maturity</td>
<td>○</td>
<td>×</td>
</tr>
<tr>
<td>- Credit rating</td>
<td>○</td>
<td>×</td>
</tr>
<tr>
<td>- Collateral re-use eligibility</td>
<td>○</td>
<td>×</td>
</tr>
<tr>
<td>- On a pure principal-to-principal basis or with the intermediation of an agent</td>
<td>○</td>
<td>×</td>
</tr>
<tr>
<td>- Haircut rate</td>
<td>○</td>
<td>△</td>
</tr>
<tr>
<td><strong>Cash</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>- Currency</td>
<td>○</td>
<td>○</td>
</tr>
<tr>
<td>- Amount</td>
<td>○</td>
<td>○</td>
</tr>
<tr>
<td>- Management party</td>
<td>×</td>
<td>○</td>
</tr>
<tr>
<td><strong>Transaction information</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>- CCP cleared or not</td>
<td>○</td>
<td>○</td>
</tr>
<tr>
<td>- GC or SC</td>
<td>○</td>
<td>○</td>
</tr>
<tr>
<td>- Bilaterally or Agent-intermediated</td>
<td>○</td>
<td>○</td>
</tr>
<tr>
<td>- Transaction Maturity</td>
<td>○</td>
<td>○</td>
</tr>
<tr>
<td>- Repo rate</td>
<td>○</td>
<td>△</td>
</tr>
</tbody>
</table>

1 This table presents the main information of the data by repo and securities lending transaction. "○" represents collected data, "△" represents uncollected data but calculation is possible, and "×" represents uncollected data. In securities lending transactions, non-cash collateral (bonds, equities, etc.) are collected in addition to cash collateral, and the repo rate is calculated as the difference between the collateral interest rate and lending rate.

14 In addition to transactions with corporations in Japan, the data covers transactions with corporations overseas as counterparties, those between head and overseas offices, between different corporations within the same group, and those with international organizations. However, they do not cover internal transactions within the same legal entity in Japan, transactions conducted between overseas offices of the reporting financial institution and overseas offices, those with individuals, those in which the reporting financial institution provides advice only, and those with central banks and the BIS.
3-2 Overview of the Securities Financing Transaction Market

Reviewing average transaction balances by security type (average end of month balances from January 2019 through December 2021, which applies hereafter, unless otherwise noted) based on this data, securities lenders (borrowers) reported 219 (190) trillion yen \(^{15}\) (Figure 2). Transactions using “government bonds” comprise approximately 80% of the total, with 181 (151) trillion yen reported by lenders (borrowers) of securities. Those using “equities,” which have the next largest balance, are reported by lenders (borrowers) of securities at 8 (10) trillion yen. Moreover, agency bonds, corporate bonds, securitized products, collateral swap transactions where securities are exchanged for each other, and basket transactions where multiple issues are traded at once total 30 (28) trillion yen, as reported by lenders (borrowers) of securities.

Outstanding balance of securities financing transactions in Japan by security type\(^{1}\)

![Figure 2]

\(^{1}\) Average outstanding balance has been calculated at the end of every month from January 2019 to December 2021. Foreign currency has been converted into Japanese yen using the exchange rates at the end of the month. The average exchange rate for 1 US dollar is 108.6 Japanese yen. The amount reported “as a securities lender” does not match that reported “as a securities borrower” because of transactions with data non-reporting parties. “Government bonds” include government-guaranteed bonds and other sovereign bonds. “Equities” regard transactions where the only security associated with the transaction is equity, and the only collateral is cash. “Others” include transactions using agency bonds, securitized products, corporate bonds, and supra bonds, as well as collateral swap transactions where securities (e.g., government bonds and equities) are exchanged for each other and basket transactions where several types of securities are traded at once.

Below is a description of the characteristics of the market structure, including haircut and repo rates, for transactions using “government bonds” and “equities,” which are typical in Japan. Using the data makes it possible to ascertain previously unidentified information; for instance, in addition to data on the jurisdiction of

\(^{15}\) The amounts reported by the lenders of securities do not match those of the borrowers of the securities because of transactions with data non-reporting parties.
government bond and cash currency, the characteristics of trading securities are classified based on the level of their combination.

(1) Government bonds

Transactions using government bonds are classified into three categories per type of transaction: (1) standard repurchase agreements, (2) subsequent collateral allocation repurchase agreements, and (3) cash-secured lending transactions (called “Gentan” transactions in Japan) (see Table 2). Historically, cash-secured lending transactions have been the mainstream in Japan (Kanno and Kato, 2001). After that, repo transactions increased following the introduction of subsequent collateral allocation repurchase agreements in conjunction with the shortening of Japanese government bond settlement cycle to T+1 in 2018 (Fujimoto et al., 2019). Repurchase agreements have been increasing moderately since 2019, while cash-secured lending transactions have been declining (Figure 3).

Outstanding balance of transactions using government bonds by transaction type\(^1\)  

<table>
<thead>
<tr>
<th>Panel A: As a securities lender</th>
<th>Panel B: As a securities borrower</th>
</tr>
</thead>
<tbody>
<tr>
<td><img src="image1.png" alt="Graph" /></td>
<td><img src="image2.png" alt="Graph" /></td>
</tr>
</tbody>
</table>

\(^1\) Average outstanding balance has been calculated at the end of every month from January 2019 to December 2021. Foreign currency has been converted into Japanese yen using the exchange rates at the end of the month.

**Standard repurchase agreements**

Standard repurchase agreements are currently the largest type of SFT in Japan. Table 2 shows that the exchange of Japanese government bonds for Japanese yen is top of the list, accounting for more than 80% of standard repurchase agreements. Moreover, transactions occur where US government bonds are exchanged for US dollars or where European government bonds are exchanged for euros. Cross-currency transactions are also undertaken, where Japanese government bonds are exchanged for US dollars.

Taking the US dollar as an example of a currency other than Japanese yen, while the data reporters’ funding of US dollars is approximately 16 trillion yen, their supplying of US dollars is approximately 2 trillion yen. Thus, on a net basis, Japanese
financial institutions procured US dollars equivalent to approximately 14 trillion yen via repurchase agreements using US and Japanese government bonds.

By combining the type of government bonds transacted (jurisdiction, denomination) with the currency of the cash being transacted, the characteristics of haircut and repo rates can be more clearly identified. First, transactions involving the exchange of Japanese government bonds and Japanese yen of the greatest transaction volume were traded at a haircut (repo) rate of almost 0% (-0.10%) at the median and weighted average values. On US dollar transactions, the haircut rate for transactions exchanging US government bonds for US dollars was almost 0%, and a weighted average repo rate was trading at around +0.9%. However, cross-currency transactions exchanging Japanese government bonds for US dollars have a weighted average haircut rate of +5.31% (repo transactions) and +2.85% (reverse repo transactions). This indicates that the haircut rate level in cross-currency transactions differs significantly from that of same-currency transactions. The haircut rate is often set at +2.00% in transactions exchanging European government bonds for euros. A time series of weighted average haircut rates for these representative transactions demonstrates that they have remained stable despite the COVID-19 turmoil upsetting the financial markets (Figure 4).

Overnight transactions account for approximately 40% of the total in terms of residual transaction maturity (Table 2). Only approximately 3% of all transactions exceed three months' maturity. The weighted average of the haircut and repo rates increases as the transaction maturity lengthens, thereby reflecting the increased market risk and term structure associated with a transaction maturity.17 However, margin calls, where additional collateral is delivered in response to changes in the market value of collateral during the transaction period, may reduce the impact of transaction maturity. Section 4 examines the magnitude of the impact using panel regression data analysis.18

Similarly, many transactions use government bonds with a residual maturity of more than one year. In theory, given that the price volatility increases as the residual maturity to redemption increases, the haircut rate is expected to increase accordingly. Nevertheless, the haircut rate in Table 2 indicates that the rate is lower for transactions using government bonds with a residual maturity greater than one year in comparison to transactions with a remaining maturity of less than one year.19 This is likely because many of the transactions exchanging Japanese yen and Japanese government bonds, for which the haircut rate is mostly set at 0%, are conducted using government bonds with a residual maturity of more than one year. Section 4 examines this point in detail via panel data analysis.

16 Practically, “repo transactions” and “reverse repo transactions” refer to transactions where a party (i.e., the data reporter) acts as the lender and borrower of securities, respectively. However, these repurchase agreements are sometimes collectively referred to “repo transactions.”

17 In value-at-risk (a typical risk measurement method), the amount of market risk increases in proportion to the square root of the transaction period. Thus, the panel data analysis in Section 4 regresses the square root of the transaction period.

18 For a detailed description of the margin call mechanism in Japan’s repurchase agreements, see Kanno and Kato (2001).

19 Where the bond price is \( B \), the bond yield is \( y \), and the residual maturity to redemption is \( D \), the relationship between the change in bond price \( \delta B \) and that in yield \( \delta y \) is approximately \( \delta B / B = -D \delta y \). Thus, the longer the residual maturity of the bond, the greater the price volatility in a linear relationship.
Trends in haircut rate

Standard repurchase agreements with government bonds

Panel A: Transactions of Japanese government bonds and Japanese yen

Panel B: Transactions of US treasuries and the US dollar

Panel C: Transactions of Japanese government bonds and the US dollar

Panel D: Transactions of French government bonds and the euro
<table>
<thead>
<tr>
<th>Category</th>
<th>Breakdown items</th>
<th>Outstanding balance (100 mil. yen)</th>
<th>Haircut rate (%)</th>
<th>Repo rate (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>Repo</td>
<td>Reverse Repo</td>
<td>Repo</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Month-end average</td>
<td>Month-end average</td>
<td>Weighted average</td>
</tr>
<tr>
<td>Total average</td>
<td></td>
<td>1,201,575</td>
<td>781,917</td>
<td>0.33</td>
</tr>
<tr>
<td>Jurisdiction of government bond</td>
<td>JP×JPY×JPY</td>
<td>963,893</td>
<td>723,068</td>
<td>0.06</td>
</tr>
<tr>
<td></td>
<td>US×USD×USD</td>
<td>127,463</td>
<td>14,990</td>
<td>0.06</td>
</tr>
<tr>
<td></td>
<td>JP×JPY×USD</td>
<td>30,389</td>
<td>7,791</td>
<td>5.31</td>
</tr>
<tr>
<td></td>
<td>FR×EUR×EUR</td>
<td>29,176</td>
<td>6,661</td>
<td>0.45</td>
</tr>
<tr>
<td></td>
<td>DE×EUR×EUR</td>
<td>9,411</td>
<td>4,547</td>
<td>0.99</td>
</tr>
<tr>
<td></td>
<td>IT×EUR×EUR</td>
<td>8,185</td>
<td>6,930</td>
<td>1.96</td>
</tr>
<tr>
<td></td>
<td>ES×EUR×EUR</td>
<td>6,783</td>
<td>4,261</td>
<td>6.40</td>
</tr>
<tr>
<td></td>
<td>GB×GBP×GBP</td>
<td>5,481</td>
<td>5,494</td>
<td>3.86</td>
</tr>
<tr>
<td></td>
<td>BE×EUR×EUR</td>
<td>3,673</td>
<td>1,656</td>
<td>0.86</td>
</tr>
<tr>
<td></td>
<td>AU×AUD×AUD</td>
<td>3,531</td>
<td>639</td>
<td>4.76</td>
</tr>
<tr>
<td>Omitted below</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Currency of cash</td>
<td>JPY</td>
<td>963,949</td>
<td>723,079</td>
<td>0.06</td>
</tr>
<tr>
<td></td>
<td>USD</td>
<td>160,920</td>
<td>23,033</td>
<td>1.13</td>
</tr>
<tr>
<td></td>
<td>EUR</td>
<td>65,979</td>
<td>29,356</td>
<td>1.60</td>
</tr>
<tr>
<td></td>
<td>GBP</td>
<td>6,553</td>
<td>5,757</td>
<td>4.20</td>
</tr>
<tr>
<td></td>
<td>AUD</td>
<td>3,915</td>
<td>680</td>
<td>4.99</td>
</tr>
<tr>
<td>Omitted below</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

1 This table presents the market structure of standard repurchase agreements with government bonds in Japan. Average outstanding balance, haircut rate, and repo rate have been calculated at the end of every month from January 2019 to December 2021. Currencies other than Japanese yen have been converted into Japanese yen using the exchange rates at the end of the month. The average exchange rate for 1 US dollar is 108.6 Japanese yen. The abbreviations for country and currency names are as follows: JP: Japan, US: United States, FR: France, DE: Germany, IT: Italy, ES: Spain, GB: United Kingdom, BE: Belgium, AU: Australia, JPY: Japanese yen, USD: US dollar, EUR: Euro, GBP: Sterling, AUD: Australian dollar.
## Market structure of standard repurchase agreements with government bonds - Continued

Table 2

<table>
<thead>
<tr>
<th>Category</th>
<th>Breakdown items</th>
<th>Outstanding balance (100 mil. yen)</th>
<th>Haircut rate (%)</th>
<th>Repo rate (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>Repo Month-end average</td>
<td>Reverse Repo Month-end average</td>
<td>Repo Weighted average Median</td>
</tr>
<tr>
<td>Transaction maturity</td>
<td></td>
<td>428,328</td>
<td>293,579</td>
<td>0.19 0.00</td>
</tr>
<tr>
<td></td>
<td>Overnight</td>
<td>485,036</td>
<td>333,823</td>
<td>0.24 0.00</td>
</tr>
<tr>
<td></td>
<td>From 2 days (including) to 1 week (including)</td>
<td>195,182</td>
<td>126,802</td>
<td>0.16 0.00</td>
</tr>
<tr>
<td></td>
<td>From 1 week (not including) to 1 month (including)</td>
<td>344,269</td>
<td>221,312</td>
<td>0.16 0.00</td>
</tr>
<tr>
<td></td>
<td>From 1 month (not including) to 3 months (including)</td>
<td>183,993</td>
<td>110,614</td>
<td>0.17 0.00</td>
</tr>
<tr>
<td></td>
<td>From 3 months (not including) to 6 months (including)</td>
<td>24,664</td>
<td>16,259</td>
<td>0.22 0.00</td>
</tr>
<tr>
<td></td>
<td>From 6 months (not including) to 12 months (including)</td>
<td>3,915</td>
<td>2,281</td>
<td>0.60 0.00</td>
</tr>
<tr>
<td></td>
<td>One year (not including) and more</td>
<td>12,984</td>
<td>3,797</td>
<td>2.66 0.00</td>
</tr>
<tr>
<td></td>
<td>Open or continuing terms contracts</td>
<td>13,662</td>
<td>10,794</td>
<td>5.50 3.93</td>
</tr>
<tr>
<td>Residual maturity of government bond</td>
<td>Below 1 month (including)</td>
<td>14,081</td>
<td>10,468</td>
<td>3.74 2.00</td>
</tr>
<tr>
<td></td>
<td>More than 1 month (not including) and up to 3 months (including)</td>
<td>42,289</td>
<td>29,775</td>
<td>0.28 0.00</td>
</tr>
<tr>
<td></td>
<td>More than 3 months (not including) and up to 6 months (including)</td>
<td>44,421</td>
<td>28,589</td>
<td>0.44 0.00</td>
</tr>
<tr>
<td></td>
<td>More than 6 months (not including) and up to 1 year (including)</td>
<td>61,943</td>
<td>36,159</td>
<td>0.79 0.00</td>
</tr>
<tr>
<td></td>
<td>More than 1 year (not including) and up to 5 years (including)</td>
<td>388,439</td>
<td>247,055</td>
<td>0.86 0.00</td>
</tr>
<tr>
<td></td>
<td>More than 5 years (not including) and up to 10 years (including)</td>
<td>334,701</td>
<td>211,765</td>
<td>0.31 0.00</td>
</tr>
<tr>
<td></td>
<td>More than 10 years (not including)</td>
<td>314,623</td>
<td>217,293</td>
<td>0.25 0.00</td>
</tr>
<tr>
<td>Clearing information</td>
<td>Centrally cleared transactions</td>
<td>581,917</td>
<td>515,661</td>
<td>0.24 0.00</td>
</tr>
<tr>
<td></td>
<td>Not centrally cleared transactions</td>
<td>619,657</td>
<td>266,256</td>
<td>0.00 0.00</td>
</tr>
<tr>
<td>Collateral management</td>
<td>Agency-intermediated transactions</td>
<td>197,129</td>
<td>61,300</td>
<td>0.00 0.00</td>
</tr>
<tr>
<td></td>
<td>Bilateral transactions</td>
<td>1,004,446</td>
<td>720,617</td>
<td>0.00 0.00</td>
</tr>
<tr>
<td>Transaction purpose</td>
<td>General collateral (GC) transactions</td>
<td>847,829</td>
<td>507,106</td>
<td>0.46 0.00</td>
</tr>
<tr>
<td></td>
<td>Special collateral (SC) transactions</td>
<td>353,746</td>
<td>274,811</td>
<td>0.01 0.00</td>
</tr>
</tbody>
</table>

Quantitative Analysis of Haircuts: Evidence from the Japanese Repo and Securities Lending Markets 13
From the data by bilateral or agency-intermediated and CCP-cleared or non-cleared transactions, the haircut rate is often set to 0% for agency-intermediated and cleared transactions. Thus, haircuts are utilized in transactions where risk management is relatively important, such as bilateral and non-centrally cleared transactions.

Finally, regarding GC or SC transactions, the haircut rate for SC transactions is smaller than that for GC transactions, and the repo rate is set lower for SC transactions than for GC transactions. As discussed in Section 2, it can be interpreted as stemming from the nature of SC transactions as securities lending rather than cash lending.

**Subsequent collateral allocation repurchase agreements**

Subsequent collateral allocation repurchase agreements are CCP-cleared transactions cleared by the Japan Securities Clearing Corporation (JSCC) and comprise a new form of transaction introduced in May 2018 (JSCC, 2018). Given that the issue of Japanese government bonds to be traded is unspecified in advance, it is a GC transaction with a cash lending aspect. Moreover, haircuts cannot be set (traded without haircuts) because the risk is managed through the margin and clearing fund by the JSCC.

Overnight transactions account for approximately 60% of all subsequent collateral allocation repurchase agreements (Table 3). The median and weighted average repo rates were -0.08% and were thus transacted at a rate somewhat higher than those for standard repurchase agreements and cash-secured lending transactions. As described by Fujimoto et al. (2019), standard repurchase agreements and cash-secured lending transactions include compensation, such as borrowing fees and administrative costs, because of the pre-designation of Japanese government bond issues. The term structure of the repo rate agrees with the “Tokyo Repo Rate” published by the Japan Securities Dealers Association (Figure 5).20

**Cash-secured lending transactions**

Despite the decline of cash-secured lending transactions with the expansion of repurchase agreements, they continue to comprise a certain proportion of the total balance, with 39 (51) trillion yen reported by the lenders (borrowers) of securities (Table 4). The breakdown indicates that, as with repurchase agreements, the most active transactions are those where yen-denominated Japanese government bonds are exchanged for Japanese yen.21 Moreover, there are foreign currency transactions where US dollar-denominated government bonds are exchanged for US dollars, euro-denominated government bonds for euros, and cross-currency transactions, where yen-denominated government bonds are exchanged, for US dollars. As with standard repurchase agreements, cross-currency transactions have relatively high haircut rates.

---

20 The “Tokyo Repo Rate” does not precisely match the rate for subsequent collateral allocation repurchase agreements because the survey covers GC repo rates for standard repurchase agreements, subsequent collateral allocation repurchase agreements, and cash-secured lending transactions (Japan Securities Dealers Association, 2017).

21 Information on the jurisdiction where the securities are issued is not a data collection item for securities lending transactions (see Table 1).
(2) Equities

Equities, having the second-largest trading value after government bonds, are traded in securities lending, and basket transactions with multiple equities exchanged under a single contract are common. Figures 6 and Table 5 identify and arrange transactions where the only equities are exchanged by cash collateral.22

Most transactions using equities involve the exchange of yen-denominated equities for Japanese yen (Table 5). From the remaining transaction period, open-ended transactions with no predetermined transaction period account for more than 70% of the total. In open-ended transactions, the median haircut rate is -4.76%, equivalent to pledging cash collateral corresponding to 105% of the equities. The weighted average repo rate for open-ended transactions reported by securities lenders (borrowers) is -0.21% (-0.31%),23 suggesting that many transactions have an aspect of securities lending and borrowing by cash collateral (SC transactions).24

22 Furthermore, there are transactions where multiple types of securities (e.g., government bonds and equities) and collateral (e.g., cash in multiple currencies and government bonds) are exchanged. However, they comprise a small portion of the total and are included under “Others” in Figure 2.

23 In a securities lending transaction, the securities borrower (the collateral lender) pays a lending fee for the lent securities, and the securities lender (the collateral borrower) pays interest on the collateral. The repo rate is commonly defined as the difference (collateral interest rate and lending rate), and this definition has been used in this paper.

24 Under the data reporting guidelines, if it is difficult to determine whether a transaction is a GC or SC transaction, the transaction is to be reported as a GC transaction. Given that transactions using
However, considering transactions predetermined outside of open-ended transactions, some show positive haircut and repo rates, suggesting that they include many transactions with an aspect of cash lending and borrowing by equities collateral (GC transactions).

This distinction between transactions based on haircut and repo rates is supported by the distribution of haircut rates by business type, as in Figure 6. In other words, “Trust banks and asset management,” the main investors of equities, trade as equity lenders with haircut rates concentrated within the range of -4% and -5%, whereas “Tanshi companies, banks, and other finance companies” often trade as equity borrowers (cash lenders) and have positive haircut rates. “Securities companies” are involved in a wide range of transactions as intermediaries.

![Distribution of haircut rates for transactions with equities (by business type)](image)

Panel A: Securities companies
Panel B: Trust banks and asset management companies
Panel C: Tanshi companies, banks, other finance companies

1 Transactions have been calculated at the end of every month from January 2019 to December 2021, classified by haircut rate. Balances are average balances.

Equities are often basket transactions, and it is difficult to distinguish between them in many cases, the share of SC transactions is thought to be larger in practice than that presented in Table 5.

25 Tanshi companies mainly act as intermediaries for interbank loans and SFTs in Japan.
Market structure of subsequent collateral allocation repurchase agreements with government bonds\(^1\)

<table>
<thead>
<tr>
<th>Category</th>
<th>Breakdown items</th>
<th>Outstanding balance (100 mil. yen)</th>
<th>Repo rate (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>Repo &amp; Reverse Repo</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>Month-end average &amp; Month-end</td>
<td>Weighted average &amp; Median</td>
</tr>
<tr>
<td>Total average</td>
<td></td>
<td>220,201 &amp; 220,307</td>
<td>-0.08 &amp; -0.08</td>
</tr>
<tr>
<td>Jurisdiction of government bond &amp; Currency of government bond &amp; Currency of cash</td>
<td>JP × JPY × JPY</td>
<td>220,201 &amp; 220,307</td>
<td>-0.08 &amp; -0.08</td>
</tr>
<tr>
<td>Currency of cash</td>
<td>JPY</td>
<td>220,201 &amp; 220,307</td>
<td>-0.08 &amp; -0.08</td>
</tr>
<tr>
<td>Transaction maturity</td>
<td>Overnight</td>
<td>130,378 &amp; 130,689</td>
<td>-0.06 &amp; -0.05</td>
</tr>
<tr>
<td></td>
<td>From 2 days (included) to 1 week (included)</td>
<td>27,703 &amp; 27,687</td>
<td>-0.08 &amp; -0.08</td>
</tr>
<tr>
<td></td>
<td>From 1 week (not included) to 1 month (included)</td>
<td>39,627 &amp; 39,438</td>
<td>-0.09 &amp; -0.09</td>
</tr>
<tr>
<td></td>
<td>From 1 month (not included) to 3 months (included)</td>
<td>22,218 &amp; 22,218</td>
<td>-0.11 &amp; -0.10</td>
</tr>
<tr>
<td></td>
<td>From 3 months (not included) to 6 months (included)</td>
<td>2,891 &amp; 2,891</td>
<td>-0.10 &amp; -0.10</td>
</tr>
<tr>
<td>Collateral management</td>
<td>Agency-intermediated transactions</td>
<td>2,311 &amp; 94,300</td>
<td>-0.10 &amp; -0.10</td>
</tr>
<tr>
<td></td>
<td>Bilateral transactions</td>
<td>217,889 &amp; 126,007</td>
<td>-0.08 &amp; -0.08</td>
</tr>
</tbody>
</table>

\(^1\) This table presents the market structure of subsequent collateral allocation repurchase agreements with government bonds in Japan. The average outstanding balance and repo rate have been calculated at the end of every month from January 2019 to December 2021. The abbreviations for country and currency names are as follows: JP: Japan, JPY: Japanese yen.
## Market structure of cash-secured lending transactions with government bonds

This table presents the market structure of cash-secured securities lending transactions with government bonds in Japan. The average outstanding balance, haircut rate, and repo rate have been calculated at the end of every month from January 2019 to December 2021. Currencies other than Japanese yen have been converted into Japanese yen using the exchange rates at the end of the month. The average exchange rate for 1 US dollar is 108.6 Japanese yen. The abbreviations for currency names are as follows: JPY: Japanese yen, USD: US dollar, EUR: Euro.

<table>
<thead>
<tr>
<th>Category</th>
<th>Breakdown items</th>
<th>Outstanding balance (100 mil. yen)</th>
<th>Haircut rate (%)</th>
<th>Repo rate (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>Securities out</td>
<td>Securities in</td>
<td>Securities out</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Month-end average</td>
<td>Weighted average</td>
<td>Median</td>
</tr>
<tr>
<td>Total average</td>
<td></td>
<td>385,924</td>
<td>510,556</td>
<td>0.17</td>
</tr>
<tr>
<td>Currency of government bond</td>
<td>JPY×JPY</td>
<td>326,119</td>
<td>418,967</td>
<td>0.09</td>
</tr>
<tr>
<td>Currency of cash × Currency of cash</td>
<td>USD×USD</td>
<td>23,867</td>
<td>38,277</td>
<td>0.23</td>
</tr>
<tr>
<td></td>
<td>EUR×EUR</td>
<td>23,156</td>
<td>32,388</td>
<td>0.06</td>
</tr>
<tr>
<td></td>
<td>JPY×USD</td>
<td>2,971</td>
<td>14,568</td>
<td>3.69</td>
</tr>
<tr>
<td>Omitted below</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Currency of cash</td>
<td>JPY</td>
<td>328,969</td>
<td>423,364</td>
<td>0.05</td>
</tr>
<tr>
<td></td>
<td>USD</td>
<td>26,839</td>
<td>52,845</td>
<td>0.61</td>
</tr>
<tr>
<td></td>
<td>EUR</td>
<td>23,159</td>
<td>32,395</td>
<td>0.06</td>
</tr>
<tr>
<td>Omitted below</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Transaction maturity</td>
<td>Overnight</td>
<td>134,654</td>
<td>138,175</td>
<td>0.01</td>
</tr>
<tr>
<td></td>
<td>From 2 days (included) to 1 week (included)</td>
<td>49,217</td>
<td>73,887</td>
<td>0.26</td>
</tr>
<tr>
<td></td>
<td>From 1 week (not included) to 1 month (included)</td>
<td>112,899</td>
<td>179,387</td>
<td>0.16</td>
</tr>
<tr>
<td></td>
<td>From 1 month (not included) to 3 months (included)</td>
<td>68,998</td>
<td>101,887</td>
<td>0.01</td>
</tr>
<tr>
<td></td>
<td>From 3 months (not included) to 6 months (included)</td>
<td>11,345</td>
<td>15,796</td>
<td>-0.53</td>
</tr>
<tr>
<td></td>
<td>From 6 months (not included) to 12 months (included)</td>
<td>1,619</td>
<td>2,411</td>
<td>0.28</td>
</tr>
<tr>
<td></td>
<td>One year (not included) and more</td>
<td>0</td>
<td>1,067</td>
<td>NaN</td>
</tr>
<tr>
<td></td>
<td>Open or continuing terms contracts</td>
<td>535</td>
<td>0</td>
<td>0.00</td>
</tr>
<tr>
<td>Clearing information</td>
<td>Centrally cleared transactions</td>
<td>231,237</td>
<td>231,491</td>
<td>0.00</td>
</tr>
<tr>
<td></td>
<td>Not centrally cleared transactions</td>
<td>154,687</td>
<td>279,065</td>
<td>0.43</td>
</tr>
<tr>
<td>Collateral management</td>
<td>Agency-intermediated transactions</td>
<td>70,411</td>
<td>56,430</td>
<td>0.61</td>
</tr>
<tr>
<td></td>
<td>Bilateral transactions</td>
<td>315,513</td>
<td>454,126</td>
<td>0.07</td>
</tr>
<tr>
<td>Transaction purpose</td>
<td>General collateral (GC) transactions</td>
<td>259,380</td>
<td>353,387</td>
<td>0.19</td>
</tr>
<tr>
<td></td>
<td>Special collateral (SC) transactions</td>
<td>126,543</td>
<td>157,169</td>
<td>0.13</td>
</tr>
</tbody>
</table>

1 This table presents the market structure of cash-secured securities lending transactions with government bonds in Japan. The average outstanding balance, haircut rate, and repo rate have been calculated at the end of every month from January 2019 to December 2021. Currencies other than Japanese yen have been converted into Japanese yen using the exchange rates at the end of the month. The average exchange rate for 1 US dollar is 108.6 Japanese yen. The abbreviations for currency names are as follows: JPY: Japanese yen, USD: US dollar, EUR: Euro.
### Market structure of equities lending transactions

<table>
<thead>
<tr>
<th>Category</th>
<th>Breakdown items</th>
<th>Outstanding balance (100 mil. yen)</th>
<th>Haircut rate (%)</th>
<th>Repo rate (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>Securities out</td>
<td>Securities in</td>
<td>Securities out</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Month-end average</td>
<td>Month-end average</td>
<td>Weighted average</td>
</tr>
<tr>
<td>Total average</td>
<td></td>
<td>76,496</td>
<td>98,183</td>
<td>-3.12</td>
</tr>
<tr>
<td>Currency of equities</td>
<td>JPY×JPY</td>
<td>75,014</td>
<td>96,534</td>
<td>-3.27</td>
</tr>
<tr>
<td>×Currency of cash</td>
<td></td>
<td>Omitted below</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Currency of cash</td>
<td>JPY</td>
<td>75,015</td>
<td>97,756</td>
<td>-3.27</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Omitted below</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Transaction maturity</td>
<td>Overnight</td>
<td>700</td>
<td>1,149</td>
<td>-1.72</td>
</tr>
<tr>
<td></td>
<td>From 2 days (included) to 1 week (included)</td>
<td>2,351</td>
<td>2,644</td>
<td>-0.25</td>
</tr>
<tr>
<td></td>
<td>From 1 week (not included) to 1 month (included)</td>
<td>3,613</td>
<td>6,749</td>
<td>3.18</td>
</tr>
<tr>
<td></td>
<td>From 1 month (not included) to 3 months (included)</td>
<td>1,466</td>
<td>2,188</td>
<td>4.36</td>
</tr>
<tr>
<td></td>
<td>From 3 months (not included) to 6 months (included)</td>
<td>2,342</td>
<td>1,832</td>
<td>4.18</td>
</tr>
<tr>
<td></td>
<td>From 6 months (not included) to 12 months (included)</td>
<td>741</td>
<td>2,155</td>
<td>4.27</td>
</tr>
<tr>
<td></td>
<td>One year (not included) and more</td>
<td>137</td>
<td>722</td>
<td>5.90</td>
</tr>
<tr>
<td></td>
<td>Open or continuing terms contracts</td>
<td>56,972</td>
<td>74,418</td>
<td>-4.36</td>
</tr>
<tr>
<td>Clearing information</td>
<td>Centrally cleared transactions</td>
<td>53,584</td>
<td>67,983</td>
<td>-4.59</td>
</tr>
<tr>
<td></td>
<td>Not centrally cleared transactions</td>
<td>22,911</td>
<td>29,316</td>
<td>0.31</td>
</tr>
<tr>
<td>Collateral management</td>
<td>Agency-intermediated transactions</td>
<td>22,461</td>
<td>6,493</td>
<td>-5.13</td>
</tr>
<tr>
<td></td>
<td>Bilateral transactions</td>
<td>54,035</td>
<td>91,690</td>
<td>-2.28</td>
</tr>
<tr>
<td>Transaction purpose</td>
<td>General collateral (GC) transactions</td>
<td>33,916</td>
<td>47,469</td>
<td>-2.19</td>
</tr>
<tr>
<td></td>
<td>Special collateral (SC) transactions</td>
<td>42,577</td>
<td>50,615</td>
<td>-3.86</td>
</tr>
</tbody>
</table>

---

1 This table presents the market structure of equities lending transactions in Japan. Transactions are extracted where the only securities linked to the transaction are equities, and the only collateral is cash. The average outstanding balance, haircut rate, and repo rate have been calculated at the end of every month from January 2019 to December 2021. Currencies other than Japanese yen have been converted into Japanese yen using the exchange rates at the end of the month. The average exchange rate for 1 US dollar is 108.6 Japanese yen. The abbreviations for country and currency names are as follows: JP: Japan, JPY: Japanese yen.
4. Panel Data Analysis

4-1 Estimation Methodology

This section furnishes a more detailed analysis of standard repurchase agreements for government bonds, the largest balance of SFTs in Japan. It employs the least squares dummy variable model, using panel data to estimate quantitatively which variables likely affect haircuts and to what extent. Specifically, we consider the following regression equation, with the haircut rate as the explained variable.

\[ \text{Haircut}_j = \alpha_0 + \sum_k \alpha_{1,k} X_{j,k} + \sum_l \delta_{l,1} (d_{j,l} = 1) + \sum_m \rho_{m,1} (p_{j,m} = m) + \sum_n \theta_{n,1} (s_{j,n} = n) + \varepsilon_j, \]

where \( j \) is a subscript representing a specific transaction, and \( 1(x = y) \) denotes a dummy variable with a function that takes 1 when \( x = y \) and 0 otherwise. Further, \( k \) in the continuous variable \( X_{j,k} \) characterizes a transaction as a subscript that distinguishes between repo rate, transaction maturity, transaction amount, and network centrality (degree centrality). Moreover, \( l \) in \( d_{j,l} \) is a subscript that distinguishes between a government bond credit rating dummy, a government bond residual maturity dummy, an open-end transaction dummy, a cross-currency transaction dummy, an agency-intermediated transaction dummy, an SC transaction dummy, and a CCP-cleared transaction dummy. Table 6 lists the explanatory variables, and Tables 7-1 and 7-2 summarize the descriptive statistics for continuous variable \( X_{j,k} \) and the explained variable, the haircut rate. This analysis mainly focused on the values and statistical significance of the regression coefficients \( \alpha_{1,k} and \delta_l \) for the explanatory variables.

The explanatory variables are expected to influence the setting of the haircut rate via the following pathways. First, the credit and market risk of government bonds depends on the transaction maturity and government bond credit rating, open-ended transaction, residual maturity of the government bond, and cross-currency transaction dummies. In addition to trading volume, the government bond residual maturity dummy is expected to affect liquidity risk through differences in trading volume by maturity in the bond market. The agency-intermediated transaction and CCP-cleared transaction dummies could potentially impact operational risk.

Furthermore, \( p_{j,m} \) and \( s_{j,n} \) are dummy variables treated as fixed effects; \( m \) in \( p_{j,m} \) is a subscript that distinguishes the combination of data reporter, counterparty, and transaction reporting date, therefore \( p_{j,m} \) is a time-fixed effect. Additionally, \( n \) in \( s_{j,n} \) is a subscript that distinguishes between the combination of the jurisdiction of the bond, the denomination of the bond, and cash currency. Thus, by capturing combinations of data reporter, counterparty, and transaction reporting date in \( p_{j,m} \) and jurisdiction of the bond, denomination of the bond, and cash currency in \( s_{j,n} \) as fixed effects, we control for the effects of transacting entities and transaction types and measure, to the extent possible, the pure effects of each explanatory variable.
### Description of the explanatory variables

<table>
<thead>
<tr>
<th>Variable 1</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Continuous variable</strong></td>
<td></td>
</tr>
<tr>
<td>Transaction maturity</td>
<td>Maturity of the transaction in square root of days; substitute 0 if the transaction maturity is open-ended</td>
</tr>
<tr>
<td>Transaction amount</td>
<td>Log principal amount of the transaction in 100 million JPY</td>
</tr>
<tr>
<td>Repo rate</td>
<td>Percentage of repo rate</td>
</tr>
<tr>
<td>Network centrality</td>
<td>Data reporter’s degree centrality, as of the transaction reporting month</td>
</tr>
<tr>
<td><strong>Dummy variable</strong></td>
<td></td>
</tr>
<tr>
<td>Collateral quality</td>
<td></td>
</tr>
<tr>
<td>Investment grade</td>
<td>Dummy variable = 1 if the credit rating of government bond is investment grade</td>
</tr>
<tr>
<td>Non-investment grade</td>
<td>Dummy variable = 1 if the credit rating of government bond is non-investment grade</td>
</tr>
<tr>
<td>Residual maturity of government bond</td>
<td></td>
</tr>
<tr>
<td>Below 1 month</td>
<td>Dummy variable = 1 if residual maturity of government bond is below 1 month (included)</td>
</tr>
<tr>
<td>More than 1 month and up to 3 months</td>
<td>Dummy variable = 1 if residual maturity of government bond is more than 1 month (not included) and up to 3 months (included)</td>
</tr>
<tr>
<td>More than 3 months and up to 6 months</td>
<td>Dummy variable = 1 if residual maturity of government bond is more than 3 months (not included) and up to 6 months (included)</td>
</tr>
<tr>
<td>More than 6 months and up to 1 year</td>
<td>Dummy variable = 1 if residual maturity of government bond is more than 6 months (not included) and up to 1 year (included)</td>
</tr>
<tr>
<td>More than 1 year and up to 5 years</td>
<td>Dummy variable = 1 if residual maturity of government bond is more than 1 year (not included) and up to 5 years (included)</td>
</tr>
<tr>
<td>More than 5 years and up to 10 years</td>
<td>Dummy variable = 1 if residual maturity of government bond is more than 5 years (not included) and up to 10 years (included)</td>
</tr>
<tr>
<td>More than 10 years</td>
<td>Dummy variable = 1 if residual maturity of government bond is more than 10 years (not included)</td>
</tr>
<tr>
<td>Transaction maturity</td>
<td></td>
</tr>
<tr>
<td>Open-end transactions</td>
<td>Dummy variable = 1 if the transaction maturity is open-ended</td>
</tr>
<tr>
<td>Transaction type</td>
<td></td>
</tr>
<tr>
<td>Cross currency</td>
<td>Dummy variable = 1 if the currency of government bond and cash are different</td>
</tr>
<tr>
<td>Special collateral (SC)</td>
<td>Dummy variable = 1 if reported as special collateral (SC) transaction</td>
</tr>
<tr>
<td>Agency-intermediated</td>
<td>Dummy variable = 1 if reported as agency-intermediated transaction</td>
</tr>
<tr>
<td>Centrally cleared</td>
<td>Dummy variable = 1 if reported as centrally cleared transaction</td>
</tr>
</tbody>
</table>

1 In addition to the above, the “Collateral quality” reporting category includes “no rating,” and the “Residual maturity of government bond” reporting category also includes “no residual maturity.”
## Descriptive statistics for haircut rates and explanatory variables

**Standard repurchase agreements with government bonds**

Table 7.1

Panel A: All samples

<table>
<thead>
<tr>
<th></th>
<th>Haircut rate</th>
<th>Repo rate</th>
<th>Transaction maturity</th>
<th>Transaction amount</th>
<th>Network centrality (Degree centrality)</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Repo</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Obs</td>
<td>595,392</td>
<td>595,392</td>
<td>550,102</td>
<td>595,392</td>
<td>595,392</td>
</tr>
<tr>
<td>Mean</td>
<td>0.80</td>
<td>0.04</td>
<td>3.32</td>
<td>1.30</td>
<td>0.250</td>
</tr>
<tr>
<td>Median</td>
<td>0.00</td>
<td>-0.10</td>
<td>2.45</td>
<td>1.65</td>
<td>0.233</td>
</tr>
<tr>
<td>Std dev</td>
<td>2.87</td>
<td>0.69</td>
<td>2.64</td>
<td>0.99</td>
<td>0.170</td>
</tr>
<tr>
<td>Min</td>
<td>-5.00</td>
<td>-3.10</td>
<td>1.00</td>
<td>-5.99</td>
<td>0.010</td>
</tr>
<tr>
<td>Max</td>
<td>20.00</td>
<td>23.35</td>
<td>43.34</td>
<td>4.06</td>
<td>0.645</td>
</tr>
<tr>
<td><strong>Reverse Repo</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Obs</td>
<td>392,779</td>
<td>392,779</td>
<td>362,209</td>
<td>392,779</td>
<td>392,779</td>
</tr>
<tr>
<td>Mean</td>
<td>0.62</td>
<td>-0.05</td>
<td>2.95</td>
<td>1.21</td>
<td>0.249</td>
</tr>
<tr>
<td>Median</td>
<td>0.00</td>
<td>-0.10</td>
<td>2.24</td>
<td>1.48</td>
<td>0.200</td>
</tr>
<tr>
<td>Std dev</td>
<td>2.62</td>
<td>0.45</td>
<td>2.31</td>
<td>1.06</td>
<td>0.186</td>
</tr>
<tr>
<td>Min</td>
<td>-2.00</td>
<td>-3.10</td>
<td>1.00</td>
<td>-5.99</td>
<td>0.010</td>
</tr>
<tr>
<td>Max</td>
<td>20.00</td>
<td>23.35</td>
<td>42.26</td>
<td>4.00</td>
<td>0.645</td>
</tr>
</tbody>
</table>

Panel B: Excluding zero-haircut samples

<table>
<thead>
<tr>
<th></th>
<th>Haircut rate</th>
<th>Repo rate</th>
<th>Transaction maturity</th>
<th>Transaction amount</th>
<th>Network centrality (Degree centrality)</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Repo</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Obs</td>
<td>88,784</td>
<td>88,784</td>
<td>44,873</td>
<td>88,784</td>
<td>88,784</td>
</tr>
<tr>
<td>Mean</td>
<td>5.39</td>
<td>0.21</td>
<td>4.03</td>
<td>0.40</td>
<td>0.079</td>
</tr>
<tr>
<td>Median</td>
<td>2.00</td>
<td>-0.16</td>
<td>2.83</td>
<td>0.99</td>
<td>0.043</td>
</tr>
<tr>
<td>Std dev</td>
<td>5.53</td>
<td>1.28</td>
<td>3.99</td>
<td>1.65</td>
<td>0.065</td>
</tr>
<tr>
<td>Min</td>
<td>-5.00</td>
<td>-1.38</td>
<td>1.00</td>
<td>-5.99</td>
<td>0.010</td>
</tr>
<tr>
<td>Max</td>
<td>20.00</td>
<td>23.35</td>
<td>42.71</td>
<td>3.16</td>
<td>0.355</td>
</tr>
<tr>
<td><strong>Reverse Repo</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Obs</td>
<td>45,994</td>
<td>45,994</td>
<td>18,939</td>
<td>45,994</td>
<td>45,994</td>
</tr>
<tr>
<td>Mean</td>
<td>5.29</td>
<td>0.10</td>
<td>2.75</td>
<td>0.10</td>
<td>0.058</td>
</tr>
<tr>
<td>Median</td>
<td>2.00</td>
<td>-0.43</td>
<td>1.41</td>
<td>0.54</td>
<td>0.042</td>
</tr>
<tr>
<td>Std dev</td>
<td>5.81</td>
<td>1.10</td>
<td>4.13</td>
<td>1.78</td>
<td>0.051</td>
</tr>
<tr>
<td>Min</td>
<td>-2.00</td>
<td>-1.49</td>
<td>1.00</td>
<td>-5.99</td>
<td>0.010</td>
</tr>
<tr>
<td>Max</td>
<td>20.00</td>
<td>23.35</td>
<td>33.09</td>
<td>3.40</td>
<td>0.355</td>
</tr>
</tbody>
</table>

---

1 All transactions were calculated at the end of every month from January 2019 to December 2021. “Haircut rate” and “Repo rate” are in percentages (%). “Transaction maturity” is the square root of the number of days remaining (excluding open-ended transactions). “Transaction amount” is the common log of the principal amount (100 million yen).
Distribution of haircut rates

<table>
<thead>
<tr>
<th>Haircut rate</th>
<th>&lt; 0%</th>
<th>[0%, 2%)</th>
<th>[2%, 4%)</th>
<th>[4%, 6%)</th>
<th>[6%, 8%)</th>
<th>[8%, 10%)</th>
<th>[10%, 12%)</th>
<th>≥ 12%</th>
<th>Sum</th>
</tr>
</thead>
<tbody>
<tr>
<td>Repo Obs</td>
<td>1,535</td>
<td>518,984</td>
<td>40,849</td>
<td>7,373</td>
<td>578</td>
<td>48</td>
<td>18,463</td>
<td>7,562</td>
<td>595,392</td>
</tr>
<tr>
<td>Repo Share</td>
<td>0.3%</td>
<td>87.2%</td>
<td>6.9%</td>
<td>1.2%</td>
<td>0.1%</td>
<td>0.0%</td>
<td>3.1%</td>
<td>1.3%</td>
<td>100.0%</td>
</tr>
<tr>
<td>Reverse Obs</td>
<td>611</td>
<td>348,679</td>
<td>28,702</td>
<td>1,590</td>
<td>606</td>
<td>77</td>
<td>7,847</td>
<td>4,667</td>
<td>392,779</td>
</tr>
<tr>
<td>Reverse Share</td>
<td>0.2%</td>
<td>88.8%</td>
<td>7.3%</td>
<td>0.4%</td>
<td>0.2%</td>
<td>0.0%</td>
<td>2.0%</td>
<td>1.2%</td>
<td>100.0%</td>
</tr>
</tbody>
</table>

Panel B: Excluding zero-haircut samples

<table>
<thead>
<tr>
<th>Haircut rate</th>
<th>&lt; 0%</th>
<th>[0%, 2%)</th>
<th>[2%, 4%)</th>
<th>[4%, 6%)</th>
<th>[6%, 8%)</th>
<th>[8%, 10%)</th>
<th>[10%, 12%)</th>
<th>≥ 12%</th>
<th>Sum</th>
</tr>
</thead>
<tbody>
<tr>
<td>Repo Obs</td>
<td>1,535</td>
<td>12,376</td>
<td>40,849</td>
<td>7,373</td>
<td>578</td>
<td>48</td>
<td>18,463</td>
<td>7,562</td>
<td>88,784</td>
</tr>
<tr>
<td>Repo Share</td>
<td>1.7%</td>
<td>13.9%</td>
<td>46.0%</td>
<td>8.3%</td>
<td>0.7%</td>
<td>0.1%</td>
<td>20.8%</td>
<td>8.5%</td>
<td>100.0%</td>
</tr>
<tr>
<td>Reverse Obs</td>
<td>611</td>
<td>1,894</td>
<td>28,702</td>
<td>1,590</td>
<td>606</td>
<td>77</td>
<td>7,847</td>
<td>4,667</td>
<td>45,994</td>
</tr>
<tr>
<td>Reverse Share</td>
<td>1.3%</td>
<td>4.1%</td>
<td>62.4%</td>
<td>3.5%</td>
<td>1.3%</td>
<td>0.2%</td>
<td>17.1%</td>
<td>10.1%</td>
<td>100.0%</td>
</tr>
</tbody>
</table>

1^ All transactions have been calculated at the end of every month from January 2019 to December 2021.

4.2 Estimate Results and Discussion

Tables 8 and 9 present a summary of the regression analysis results. Table 8 shows the estimated results for all samples, whereas Table 9 conveys the estimated results, excluding transaction data samples with zero-haircut, separated into repo and reverse repo transactions. The robustness of the results is ensured by comparing the estimated results for all samples, including transactions with zero-haircut and excluding transaction data samples with zero-haircut.\(^{26}\) The explained variables are haircut rates, with the regression coefficients for the various explanatory variables displayed. Model (1) in Tables 8 and 9 is the baseline model in this study, and the estimate results are based on the following explanatory variables: government bond characteristics (government bond credit rating [residual maturity] dummy); transaction period (residual transaction maturity and open-ended trading dummy); transaction terms (principal amount); type of transaction (SC trading, agency-intermediated trading, and CCP-cleared trading dummies). Model (2) adds transaction terms (repo rate) to the explanatory variables in Model (1), while Model (3) incorporates transaction type (cross-currency transaction dummy) to the explanatory variables in Model (1); only the government bond-issuing jurisdiction is considered in fixed effect \(s_{j,n}\). Model (4) adds a network centrality (degree centrality) to the explanatory variables in Model (1); only the combination of counterparty and transaction reporting date is considered in the time-fixed effects \(p_{j,m}\).

Further, to examine quantitatively how and to what extent each explanatory variable has an impact, based on the descriptive statistics of the explanatory variables (Table 7-1) and the results of regression analysis (Tables 8 and 9), Table 10

\(^{26}\) Baklanova et al. (2019) estimate regression analysis on a sample that excludes transactions through the Fixed Income Clearing Corporation, which does not set haircuts, and transactions with zero haircuts to avoid sample bias in their analysis of bilateral transactions using US treasuries.
summarizes the absolute value of the regression coefficient for each explanatory variable multiplied by the standard deviation (for dummy variables, the absolute value of the regression coefficient). From Table 10, it is possible to compare the extent to which each explanatory variable influences the haircut setting.

As noted below, the explanatory variables that affect credit, market, and liquidity risk, such as the credit quality of government bonds traded (government bond credit rating dummy), residual maturity of government bonds dummy, and foreign exchange risk (cross-currency transaction dummy), significantly impact the haircut setting.

**Characteristics of government bonds**

The credit rating of the government bonds traded has a significant impact on haircut setting. Table 10 indicates that the difference in haircut rates between investment-grade and non-investment-grade bonds is 1.72% and 1.59% for all samples, 2.90% and 3.20% for the samples without zero-haircuts (repo and reverse repo transactions, respectively, which is the same hereinafter, unless otherwise noted), confirming that the higher the credit rating, the lower the haircut rate.

The government bond residual maturity dummy should reflect the impact on price volatility and the liquidity of the government bond market by residual maturity. Table 10 shows that lengthening the residual maturity of government bonds boosts the haircut rate by 0.40% and 0.10% for all samples and 1.02% and 0.95% for samples without zero-haircuts, with a commensurate boost to the haircut rate. Furthermore, to examine the contribution by residual maturity more closely, Figure 7.1 illustrates that the longer the residual maturity of government bonds, the more the haircut rate increases. The price volatility of Japanese government bonds has remained low under the Bank of Japan’s monetary policy. Therefore, we examined whether similar results could be obtained when examining Japanese government bonds alone. Figure 7.2 illustrates that the effect of the government bond residual maturity dummy for Japanese government bonds is smaller than that of other government bonds, partly because the price volatility of Japanese government bonds has remained lower than that of other government bonds.

**Transaction maturity**

Transaction maturity can affect haircut rates, primarily through an increase or decrease in market risk. From the estimated results, lengthening the remaining duration of the transaction contributes to increasing the haircut rate. Of course, the magnitude of the effect is 0.04% and 0.03% for all samples and 0.10% and 0.06% for samples without zero-haircuts; thus, it does not significantly affect the haircut rate (Table 10). Notably, this result may be due to margin calls, where collateral is delivered in response to changes in the price of government bonds during the term of the transaction, which may mitigate the effect of the residual maturity of the transaction. However, the impact of the open-ended transaction dummy was 0.38% and 0.25% for all samples and 0.49% and 0.31% for samples without zero-haircuts. This indicates that open-ended transactions with no predetermined transaction period increased the haircut rate correspondingly.

---

27 Transactions using non-investment-grade bonds comprise only a small portion of all standard repurchase agreements with government bonds—approximately 1% regarding the number of transactions and 0.05%, outstanding transactions.
Regression results for all transactions

<table>
<thead>
<tr>
<th>Category</th>
<th>Variable</th>
<th>Repo</th>
<th>Reverse repo</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>(1)</td>
<td>(2)</td>
</tr>
<tr>
<td>Collateral quality</td>
<td>Investment grade</td>
<td>-0.5958***</td>
<td>-0.6006***</td>
</tr>
<tr>
<td>Non-investment grade</td>
<td>1.1247***</td>
<td>1.2262***</td>
<td>4.9440***</td>
</tr>
<tr>
<td>Residual maturity of government bond</td>
<td>Below 1 month (included)</td>
<td>-0.9352***</td>
<td>-0.2925</td>
</tr>
<tr>
<td></td>
<td>More than 1 month (not included) and up to 3 months (included)</td>
<td>-0.7852***</td>
<td>-0.1437</td>
</tr>
<tr>
<td></td>
<td>More than 3 months (not included) and up to 6 months (included)</td>
<td>-0.5973***</td>
<td>0.0392</td>
</tr>
<tr>
<td></td>
<td>More than 6 months (not included) and up to 1 year (included)</td>
<td>-0.5383***</td>
<td>0.0948</td>
</tr>
<tr>
<td></td>
<td>More than 1 year (not included) and up to 5 years (included)</td>
<td>-0.5227***</td>
<td>0.1174</td>
</tr>
<tr>
<td></td>
<td>More than 5 years (not included) and up to 10 years (included)</td>
<td>-0.4949***</td>
<td>0.1443</td>
</tr>
<tr>
<td></td>
<td>More than 10 years (not included)</td>
<td>-0.5347***</td>
<td>0.0970</td>
</tr>
<tr>
<td>Transaction maturity</td>
<td>Transaction maturity days</td>
<td>0.0151***</td>
<td>0.0162***</td>
</tr>
<tr>
<td></td>
<td>Open-end transactions</td>
<td>0.3830***</td>
<td>0.3589***</td>
</tr>
<tr>
<td>Transaction terms</td>
<td>Transaction amount</td>
<td>0.0282***</td>
<td>0.0238***</td>
</tr>
<tr>
<td></td>
<td>Repo rate</td>
<td>0.5367***</td>
<td>3.6807***</td>
</tr>
<tr>
<td>Transaction type</td>
<td>Cross currency</td>
<td>-0.1021***</td>
<td>-0.0901***</td>
</tr>
<tr>
<td></td>
<td>Special collateral</td>
<td>-0.2320***</td>
<td>-0.2861***</td>
</tr>
<tr>
<td></td>
<td>Agency-intermediated</td>
<td>0.1019***</td>
<td>0.1309***</td>
</tr>
<tr>
<td></td>
<td>Centrally cleared</td>
<td>0.0161***</td>
<td>0.0127***</td>
</tr>
<tr>
<td>Network centrality</td>
<td>Degree centrality</td>
<td>0.0359***</td>
<td>0.0328***</td>
</tr>
<tr>
<td>Fixed effect</td>
<td>Data Reporter × Counterparty × Transaction reporting date</td>
<td>YES</td>
<td>YES</td>
</tr>
<tr>
<td></td>
<td>Counterparty × Transaction reporting date</td>
<td>NO</td>
<td>NO</td>
</tr>
<tr>
<td></td>
<td>Jurisdiction of government bond × Currency of government bond</td>
<td>YES</td>
<td>YES</td>
</tr>
<tr>
<td></td>
<td>Jurisdiction of government bond × Currency of cash</td>
<td>NO</td>
<td>NO</td>
</tr>
<tr>
<td>Number of transactions</td>
<td>595,392</td>
<td>595,392</td>
<td>595,392</td>
</tr>
<tr>
<td>R²(Within)</td>
<td>0.660</td>
<td>0.664</td>
<td>0.614</td>
</tr>
</tbody>
</table>

1 This table presents estimate coefficients from fixed-effect panel OLS regressions of the haircut rate for all transactions. Table 6 presents the explanatory variables. Fixed effects are the combination of data reporter, counterparty, transaction reporting date, and the combination of jurisdiction of government bond, currency of government bond, and currency of cash. From the F-test, the fixed effects are supported at the 1% significance level in all models. Column (1) reports the baseline result. Column (2) adds a repo rate to column (1) to confirm the relationship between haircut and repo rates. Column (3) adds a dummy variable that identifies cross-currency transactions to column (1). Column (4) adds network centrality (degree centrality) to column (1) to confirm the network effect. The degrees of freedom for repo transaction are column (1): 384,672, column (2): 384,671, column (3): 384,727, and column (4): 391,213. ***, **, and * represent statistical significance at the 1%, 5%, and 10% levels, respectively.
Regression results for excluding zero haircuts samples

<table>
<thead>
<tr>
<th>Category</th>
<th>Variable</th>
<th>Repo</th>
<th>Reverse repo</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>(1)</td>
<td>(2)</td>
</tr>
<tr>
<td>Collateral quality</td>
<td>Investment grade</td>
<td>-1.9139***</td>
<td>-1.9197***</td>
</tr>
<tr>
<td></td>
<td>Non-investment grade</td>
<td>0.9816***</td>
<td>1.1714***</td>
</tr>
<tr>
<td>Residual maturity</td>
<td>Below 1 month (included)</td>
<td>-8.1009***</td>
<td>-6.6695***</td>
</tr>
<tr>
<td>of government bond</td>
<td>More than 1 month (not included) and up to 3 months (included)</td>
<td>-7.9593***</td>
<td>-6.5316***</td>
</tr>
<tr>
<td></td>
<td>More than 3 months (not included) and up to 6 months (included)</td>
<td>-7.5967***</td>
<td>-6.1703***</td>
</tr>
<tr>
<td></td>
<td>More than 6 months (not included) and up to 1 year (included)</td>
<td>-7.3122***</td>
<td>-5.9033***</td>
</tr>
<tr>
<td></td>
<td>More than 1 year (not included) and up to 5 years (included)</td>
<td>-7.0446***</td>
<td>-5.6481***</td>
</tr>
<tr>
<td></td>
<td>More than 5 years (not included) and up to 10 years (included)</td>
<td>-6.9140***</td>
<td>-5.5284***</td>
</tr>
<tr>
<td></td>
<td>More than 10 years (not included)</td>
<td>-7.0795***</td>
<td>-5.7103***</td>
</tr>
<tr>
<td>Transaction maturity</td>
<td>Transaction maturity days</td>
<td>0.0241***</td>
<td>0.0232***</td>
</tr>
<tr>
<td></td>
<td>Open-end transactions</td>
<td>0.4897***</td>
<td>0.4572***</td>
</tr>
<tr>
<td>Transaction terms</td>
<td>Transaction amount</td>
<td>0.0155**</td>
<td>0.0140*</td>
</tr>
<tr>
<td></td>
<td>Repo rate</td>
<td>0.8049***</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Cross currency</td>
<td></td>
<td>0.1617</td>
</tr>
<tr>
<td></td>
<td>Special collateral</td>
<td>-1.6883***</td>
<td>-1.7374***</td>
</tr>
<tr>
<td></td>
<td>Agency-intermediated</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Centrally cleared</td>
<td>-1.1390***</td>
<td>-1.2253***</td>
</tr>
<tr>
<td>Network centrality</td>
<td>Degree centrality</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Fixed effect</td>
<td>Data Reporter × Counterparty × Transaction reporting date</td>
<td>YES</td>
<td>YES</td>
</tr>
<tr>
<td></td>
<td>Counterparty × Transaction reporting date</td>
<td>NO</td>
<td>NO</td>
</tr>
<tr>
<td></td>
<td>Jurisdiction of government bond</td>
<td>YES</td>
<td>YES</td>
</tr>
<tr>
<td></td>
<td>× Currency of government bond × Currency of cash</td>
<td>NO</td>
<td>NO</td>
</tr>
<tr>
<td></td>
<td>Jurisdiction of government bond</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Number of transactions</td>
<td></td>
<td>88,784</td>
<td>88,784</td>
</tr>
<tr>
<td>R²(Within)</td>
<td></td>
<td>0.726</td>
<td>0.733</td>
</tr>
</tbody>
</table>

1 This table presents estimate coefficients from fixed-effect panel OLS regressions of haircut rate for excluding zero haircuts sample. Table 6 presents the explanatory variables. Fixed effects are the combination of data reporter, counterparty, transaction reporting date, and the combination of jurisdiction of government bond, currency of government bond, and currency of cash. From the F-test, the fixed effects are supported at the 1% significance level in all models. Column (1) reports the baseline result. Column (2) adds a repo rate to column (1) to confirm the relationship between haircut and repo rates. Column (3) adds a dummy variable that identifies cross-currency transactions to column (1). Column (4) adds network centrality (degree centrality) to column (1) to confirm the network effect. The degrees of freedom for repo transaction are column (1): 45,393, column (2): 45,392, column (3): 45,441, and column (4): 45,630. ***, **, and * represent statistical significance at the 1%, 5%, and 10% levels, respectively.
### Deviation of explanatory variables on haircut rate

<table>
<thead>
<tr>
<th>Category/Variable</th>
<th>Regression model</th>
<th>All samples</th>
<th>Excluding zero haircut</th>
<th>Interpretation of Estimation Results</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Characteristics of Government Bond</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Collateral quality (Dummy variable)</td>
<td>(1)</td>
<td>1.72***</td>
<td>1.59***</td>
<td>The higher the credit rating, the lower the haircut rate.</td>
</tr>
<tr>
<td>Residual maturity (Dummy variable)</td>
<td>(1)</td>
<td>0.40**</td>
<td>0.10***</td>
<td>The haircut rate of government bonds with a long residual maturity tends to be higher than that of government bonds with a short residual maturity.</td>
</tr>
<tr>
<td><strong>Transaction Maturity</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Transaction maturity days</td>
<td>(1)</td>
<td>0.04***</td>
<td>0.03***</td>
<td>Haircut rates increase with long transaction maturity, but the impact is not significant.</td>
</tr>
<tr>
<td>Open-end transactions (Dummy variable)</td>
<td>(1)</td>
<td>0.38***</td>
<td>0.25***</td>
<td>Open-end transactions increase haircut rates.</td>
</tr>
<tr>
<td><strong>Transaction Terms</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Repo rate</td>
<td>(2)</td>
<td>0.37***</td>
<td>0.29***</td>
<td>There is a positive correlation between haircut rate and repo rate.</td>
</tr>
<tr>
<td>Transaction amount</td>
<td>(1)</td>
<td>0.03***</td>
<td>0.01***</td>
<td>The effect of transaction amount does not significantly affect the haircut rate.</td>
</tr>
<tr>
<td><strong>Network Centrality</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Degree centrality</td>
<td>(4)</td>
<td>0.22***</td>
<td>0.03***</td>
<td>Low haircut rate for financial institutions near the center of the network.</td>
</tr>
<tr>
<td><strong>Transaction Type</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Cross currency (Dummy variable)</td>
<td>(3)</td>
<td>3.68***</td>
<td>0.96***</td>
<td>Cross-currency transactions increase haircut rates.</td>
</tr>
<tr>
<td>Special collateral (Dummy variable)</td>
<td>(1)</td>
<td>0.10***</td>
<td>0.01***</td>
<td>Haircut rate for GC transactions is higher than that for SC transactions.</td>
</tr>
<tr>
<td>Agency-intermediated (Dummy variable)</td>
<td>(1)</td>
<td>—</td>
<td>0.41***</td>
<td>Haircut rate for bilateral transactions is higher than that for agency-intermediated transactions.</td>
</tr>
<tr>
<td>Centrally cleared (Dummy variable)</td>
<td>(1)</td>
<td>0.23***</td>
<td>0.47***</td>
<td>In repo transactions, haircut rate for centrally cleared transactions is lower than that for non-centrally cleared transactions.</td>
</tr>
</tbody>
</table>

---

1 The deviation of explanatory variables on haircut rate is the product of the absolute value of the regression coefficient and the standard deviation based on Tables 7 through 9. Regarding a dummy variable, it is simply the absolute value of the regression coefficient. "Collateral quality" is the absolute value of the difference between the regression coefficients of "Investment grade" and "Non-investment grade." If the significance level of "Non-investment grade" is less than 10%, it is the absolute value of the regression coefficient of "Investment grade." "Residual maturity" is the absolute value of the difference between the regression coefficients of "Below 1 month (included)" and "More than 10 years (not included)" or "More than 5 years (not included) and up to 10 years (included)." ***, **, and * represent statistical significance at the 1%, 5%, and 10% levels, respectively.
Term structure of residual maturity of government bonds

Comparison of all samples and samples excluding zero haircut

Panel A: All samples

Panel B: Excluding zero haircut

Term structure of residual maturity of government bonds

Comparison of jurisdictions for government bonds

Panel A: All samples

Panel B: Excluding zero haircut

1 ***, **, and * represent statistical significance at the 1%, 5%, and 10% levels, respectively.

This figure presents the result of the jurisdiction government bond divided into “Japan” and “other than Japan,” estimated by the same regression equation as model (1). ***, **, and * represent statistical significance at the 1%, 5%, and 10% levels, respectively.
Transaction conditions

Several prior studies have discussed the relationship between haircut and repo rates. Prior theoretical studies commonly hold that if haircut and repo rates can be determined simultaneously, then there is a complementary relationship and a negative correlation (e.g., Auh and Landoni, 2015). However, Baklanova et al. (2019), who analyze the US market, note that, in practice, haircut and repo rates are not always simultaneously determined. Haircut rates are predetermined by elements, including risk management departments, independent of front office traders, while repo rates are often determined independently by traders when creating contracts. Thus, they found no evidence of a negative correlation, even after controlling for trading entities and time effects. The results in this study also confirm a positive correlation between haircut and repo rates, likely influenced by haircut rates generally being set independent of repo rates in Japan.

The transaction amount may affect haircuts, as per the balance with overall market liquidity. The estimates demonstrate that the impact of the transaction amount is minimal, at 0.03% and 0.01% for all samples and 0.03% and 0.03% for samples without zero-haircuts (Table 10). Thus, the principal amount of each transaction is not sufficiently large enough to affect haircuts.28

Network effects

Referring to prior studies that analyzed the network structure of the interbank and Japanese government bond repo markets in Japan (Imakubo and Soejima, 2010; Horikawa et al., 2021), this study measures the importance of each financial institution on the trading network using network centrality indicators, such as “Degree centrality” and “PageRank,” to analyze how network effects affect haircut setting.29 Figure 8 illustrates the trading network for government bond standard repurchase agreements and the degree centrality of each trading entity. Financial institutions closer to the center of the network, with larger degree centrality values, exert a greater impact on the transaction network.

The estimated results indicate that, similar to prior studies in the UK market (Julliard et al., 2019), haircut rates are set lower at financial institutions closer to the center of the network.30 The magnitude of the effect of degree centrality is 0.22% and 0.03% for all samples and 0.59% and 0.78% for samples without zero-haircuts (Table 10). Financial institutions’ proximity to the center of the network indicates that they

28 In Japan, when transactions have settlement values exceeding five billion yen, guidelines recommend that the settlement should be divided into smaller blocks to facilitate the settlement using the BOJ-net, the main settlement system (Japan Securities Dealers Association, 2016). In fact, these data also confirm that settlement values are often around five billion yen. Considering the bias such market practices may cause in the estimation, we conducted the same estimation by focusing on transactions with a settlement value of 10 billion yen or more. Even when doing so, the effect of the transaction amount on the haircut rate was small.

29 Degree centrality is the simplest network centrality indicator, and when applied to repo transactions, the higher the number of counterparties for each financial institution, the higher the value. However, “PageRank” adopts a higher value for the size of each financial institution’s transactions and the large transactions of the parties to which each financial institution is connected, thus measuring to what extent each financial institution affects the entire network.

30 Similar estimates using “PageRank” instead of degree centrality indicate that financial institutions closer to the network center have lower haircuts. The impact magnitude is also comparable to that of degree centrality, though this is not presented in this article.
face numerous counterparties, effectively matching their funding and supply needs, thereby inducing lower haircut settings.

**Transaction network in standard repurchase agreements¹**

1 This figure illustrates Japan’s transaction network in standard repurchase agreements using government bonds. Each node indicates a legal entity. Each legal entity is identified by the name of financial institution, business type, and location jurisdiction. Therefore, even the same financial group is divided into legal entities as much as possible. However, the names of financial institutions other than those provided by the data reporter may be classified as “domestic resident” or “domestic non-resident.” In this case, multiple legal entities are counted in the same node. The layout of the nodes is based on the force-directed algorithm of Kamada and Kawai (1989), similar to that of Imakubo and Soejima (2010). The degree centrality in this figure is calculated based on all samples from January 2019 to December 2021, and the edge thickness is the sum of repo and reverse repo transactions from January 2019 to December 2021.
In this regard, Horikawa et al. (2021) demonstrated that in the Japanese government bond repo market, financial institutions close to the center of the network serve as transaction intermediaries, and ongoing business relationships are established around such actors. Such transaction relationships support the causal relationship assumed by this study, where financial institutions closer to the center of the network set lower haircuts. However, it is logically possible that the inverse is the case (i.e., the stance toward haircut setting changes the position of the financial institution in the network). Moreover, the influence of unobservable data reporter characteristics may introduce bias into the estimated results for the network centrality. Hence, to address such endogeneity issues, we followed Temizsoy et al. (2017), who analyzed network effects in the European interbank market and conducted a regression analysis using a lag term for the network centrality indicator as an instrumental variable (IV) as a robustness check (Table 11). The results indicate that the effect of degree centrality is robust, and the magnitude and statistical significance of the regression coefficients are almost the same relative to the case where no instrumental variable is used (Tables 8 and 9). Viewed in detail, the absolute value of the regression coefficient is slightly smaller than that without the instrumental variable for all samples. Further, the absolute value of the regression coefficient is slightly larger than when the instrumental variable is not used when excluding zero-haircut samples. Thus, the unobservable data reporter characteristics affect degree centrality and haircut setting, and the use of the instrumental variable can be considered to have removed the biases.31

### Transaction type

Cross-currency transactions, where the government bond issue and cash currency are different, have higher haircut rates than transactions where the currencies are the same. This is likely because foreign exchange risk from currency mismatches is considered. The magnitude of the effect of the cross-currency transaction dummy is 3.68% and 0.96% for all samples and 0.16% and 6.99% for samples without zero-haircut samples, which, together with government bond credit ratings, significantly impacts haircut settings (Table 10).

The haircut rate is higher for GC transactions than SC transactions, consistent with the role of haircuts in SFTs discussed in Section 2. The magnitude of the effect of the SC transaction dummy is 0.10% and 0.01% for all samples and 1.69% for repo transactions excluding zero-haircut samples (Table 10).

---

31 The Wu-Hausman test for explanatory variable endogeneity rejected the null hypothesis that degree centrality is exogenous at the 1% and 10% levels (Table 11). Therefore, it is an endogenous variable.
### Instrumental variable estimates of network effect

<table>
<thead>
<tr>
<th>Category</th>
<th>Variable</th>
<th>All samples</th>
<th>Excluding zero haircut</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>Repo (4)</td>
<td>Reverse repo (4)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Excluding zero haircut (4)</td>
<td>Reverse repo (4)</td>
</tr>
<tr>
<td>Credit rating of government bond</td>
<td>Investment grade</td>
<td>-0.5661***</td>
<td>-0.6233***</td>
</tr>
<tr>
<td></td>
<td>Non-investment grade</td>
<td>0.8861***</td>
<td>0.9455***</td>
</tr>
<tr>
<td>Residual maturity of government bond</td>
<td>Below 1 month (included)</td>
<td>0.6949***</td>
<td>0.3990***</td>
</tr>
<tr>
<td></td>
<td>More than 1 month (not included) and up to 3 months (included)</td>
<td>0.7310***</td>
<td>0.3847***</td>
</tr>
<tr>
<td></td>
<td>More than 3 months (not included) and up to 6 months (included)</td>
<td>0.7909***</td>
<td>0.5108***</td>
</tr>
<tr>
<td></td>
<td>More than 6 months (not included) and up to 1 year (included)</td>
<td>0.8351***</td>
<td>0.5740***</td>
</tr>
<tr>
<td></td>
<td>More than 1 year (not included) and up to 5 years (included)</td>
<td>0.7511***</td>
<td>0.5060***</td>
</tr>
<tr>
<td></td>
<td>More than 5 years (not included) and up to 10 years (included)</td>
<td>0.6520***</td>
<td>0.4980***</td>
</tr>
<tr>
<td></td>
<td>More than 10 years (not included)</td>
<td>0.6463***</td>
<td>0.4952***</td>
</tr>
<tr>
<td>Transaction maturity</td>
<td>Transaction maturity days</td>
<td>0.0096***</td>
<td>0.0030***</td>
</tr>
<tr>
<td></td>
<td>Open-end transactions</td>
<td>2.1580***</td>
<td>0.5630***</td>
</tr>
<tr>
<td>Transaction terms</td>
<td>Transaction amount</td>
<td>-0.0383***</td>
<td>0.0146***</td>
</tr>
<tr>
<td>Transaction type</td>
<td>Agency-intermediated</td>
<td>-0.1226***</td>
<td>-0.0533***</td>
</tr>
<tr>
<td></td>
<td>Centrally cleared</td>
<td>0.0141***</td>
<td>0.0469***</td>
</tr>
<tr>
<td></td>
<td>Degree centrality</td>
<td>0.0337***</td>
<td>-0.0793***</td>
</tr>
<tr>
<td>Network centrality</td>
<td>Degree centrality</td>
<td>-1.2879***</td>
<td>-0.1727***</td>
</tr>
<tr>
<td></td>
<td></td>
<td>-9.3998***</td>
<td>-17.819***</td>
</tr>
<tr>
<td>Fixed effect</td>
<td>Counterparty × Transaction reporting date</td>
<td>YES</td>
<td>YES</td>
</tr>
<tr>
<td></td>
<td>Jurisdiction of government bond × Transaction reporting date</td>
<td>YES</td>
<td>YES</td>
</tr>
<tr>
<td></td>
<td>Jurisdiction of government bond × Currency of government bond</td>
<td>YES</td>
<td>YES</td>
</tr>
<tr>
<td></td>
<td>Jurisdiction of government bond × Currency of cash</td>
<td>YES</td>
<td>YES</td>
</tr>
<tr>
<td></td>
<td>Number of transactions</td>
<td>595,217</td>
<td>392,748</td>
</tr>
<tr>
<td></td>
<td>Adj. $R^2$</td>
<td>0.786</td>
<td>0.862</td>
</tr>
<tr>
<td></td>
<td>Wu-Hausman endogeneity test (F-statistic)</td>
<td>39.011***</td>
<td>3.431*</td>
</tr>
</tbody>
</table>

1 This table reports estimate coefficients from fixed-effect panel OLS regressions of haircut rate using an instrument variable. Specifically, we use degree centrality in $t - 1$ as an instrumental variable for degree centrality in $t$. Tables 6 and 7, respectively, present a list of explanatory variables and descriptive statistics. Tables 8 and 9 present the results from the fixed-effect panel OLS regressions without the instrumental variable. ***, **, and * represent statistical significance at the 1%, 5%, and 10% levels, respectively.
For repo and reverse repo transactions, the effect of CCP-cleared transactions is mixed. For repo transactions, the haircut rate for cleared transactions is lower than that for non-cleared transactions, with the size of the CCP-cleared transaction dummy effect being 0.23% for all samples and 1.14% for samples without zero-haircuts (Table 10). Conversely, for reverse repo transactions, the haircut rate for cleared transactions is higher than that for the non-cleared. The haircut rate for agency-intermediated transactions is 0.41% lower than that for bilateral transactions for reverse repo transactions in all samples only; the other results are not statistically significant. Thus, the CCP-cleared transaction and agency-intermediated transaction dummies, which could potentially affect operational risk, did not produce stable results in the haircut setting, with insignificant effects. Notably, prior studies in the US and European markets (Copeland et al., 2014; European Securities and Markets Authority, 2016; Boissel et al., 2017; Nguyen, 2020) analyzed the repo market when it was under intense stress during the financial crisis. While the Japanese repo market was generally stable from January 2019 to December 2021, the COVID-19 pandemic exerted an impact (Bank of Japan Financial Markets Department, 2020). The impact of the CCP-cleared and agency-intermediated transaction dummies could also vary per financial environment.

5. Conclusion

This study is the first to analyze the market structure and haircut-setting mechanism of securities financing transactions in-depth using government bonds and equities transaction data from financial institutions located in Japan, collected by the FSA and the Bank of Japan.

From the panel data regression analysis, we determined that explanatory variables affecting credit, market, and liquidity risk, such as government bonds’ credit quality, the residual maturity of government bonds, and the presence of foreign exchange risk, significantly impact haircut setting in government bond repo transactions. The results indicate that financial institutions closer to the center of the network, which engage in transactions with additional financial institutions, tend to set lower haircut rates through more efficient matching of borrowing and lending needs for cash and securities. Moreover, the credit quality of government bonds transacted, exchange rate stability, and the presence of intermediaries important to the trading network significantly impact the degree of functioning of the government bond repo market.

These findings can further discussions on trends and risk management of SFTs, including haircuts, as appropriate monitoring of SFTs, which are essential to financial markets and are conducted via ongoing data analyses. However, the study has some limitations that pave the way for future studies. First, this study failed to adequately probe the accumulation of knowledge concerning transactions using securities beyond government bonds and equities. Second, the low data coverage for securities lending transactions, including equities, prevents a detailed analysis. Thus, further studies can aim to bridge the gap for a better understanding of the issues to clarify the big picture of SFTs in Japan. Moreover, the accumulation of time-series data would also allow for empirical analysis of market stress.
References


European Securities and Markets Authority (2016). Report on securities financing transactions and leverage in the EU.


Japan Securities Dealers Association (2017). Terms and Conditions for Calculation and Publication of Tokyo Repo Rate (reference institutions average).


Quantitative Analysis of Haircuts: Evidence from the Japanese Repo and Securities Lending Markets

Kazuya Suzuki and Kana Sasamoto
Bank of Japan

Eleventh IFC Conference on “Post-pandemic landscape for central bank statistics”
BIS Basel, 25 and 26 August 2022

Views expressed here are those of the authors and do not necessarily reflect those of the Bank of Japan.
Introduction & Motivation

- Securities Financing Transactions (SFTs) were key to the risk-taking that induced the 2007–2009 global financial crisis. Funding environment rapidly deteriorated as the haircut rate was raised...

- Japan has a large SFT market, mainly in Japanese government bond transactions. The month-end average outstanding balance is 220 trillion JPY, or 2 trillion USD.

  - The Federal Reserve Board estimates the total repo outstanding in the US market as of September 2020 to be approximately 4 trillion USD.

- This paper reveals:
  - the market structure of Japan’s SFT market
  - standard haircut setting mechanisms for government bond transactions.

Outstanding balance of SFTs in Japan by security and transaction type

Note: Average outstanding balance has been calculated at the end of every month from January 2019 to December 2021.
Prior Studies & Our Data

- Prior Studies includes:
  
  - **Theoretical studies**: credit and market risk (Martin et al., 2014; Gottardi et al., 2019) and liquidity risk (Brunnermeier and Pedersen, 2009; Martin et al., 2014; Parlatore, 2019).
  
  
  - **Empirical study for the UK market**: Julliard et al. (2019) use data from six major financial institutions collected by a financial authority.

- **Our data captures over 90% of Japan’s SFT market, including bilateral and CCP-uncleared transactions, which are typically challenging to ascertain.** The data is reported by approximately 50 top financial institutions, including overseas financial institutions based in Japan.
Transactions with government bonds in Japan

**Transaction type**

- Repurchase agreements increased following the shortening of the Japanese government bond settlement cycle to T+1 in 2018.

- Historically, cash-secured lending transactions (called “Gentan” transactions) have been the mainstream in Japan. Despite the decline with the expansion of repurchase agreements, they continue to comprise a certain proportion.

**Combination of government bond and cash**

- The exchange of Japanese government bonds for yen accounts for 80% of standard repurchase agreements.

- Moreover, US government bonds are exchanged for US dollars or European government bonds are exchanged for euros.

- Japanese government bonds are also exchanged for US dollars.

---

**Outstanding balance of transactions using government bonds by transaction type**

<table>
<thead>
<tr>
<th>Date</th>
<th>Cash-secured lending transactions (Gentan transactions)</th>
<th>Subsequent collateral allocation repurchase agreements</th>
<th>Standard repurchase agreements</th>
</tr>
</thead>
<tbody>
<tr>
<td>Jan. 19</td>
<td>200</td>
<td>180</td>
<td>120</td>
</tr>
<tr>
<td>July 19</td>
<td>180</td>
<td>160</td>
<td>100</td>
</tr>
<tr>
<td>Jan. 20</td>
<td>160</td>
<td>140</td>
<td>90</td>
</tr>
<tr>
<td>July 20</td>
<td>140</td>
<td>120</td>
<td>70</td>
</tr>
<tr>
<td>Jan. 21</td>
<td>120</td>
<td>100</td>
<td>50</td>
</tr>
<tr>
<td>July 21</td>
<td>100</td>
<td>80</td>
<td>30</td>
</tr>
</tbody>
</table>

**Combination by jurisdiction of government bond, currency of government bond, and currency of cash**

- JP × JPY × JPY: 80%
- FR × EUR × EUR: 2%
- JP × JPY × USD: 3%
- US × USD × USD: 11%
- Other: 4%

Haircut rate by transaction type

- Exchanging Japanese government bonds for Japanese yen and US government bonds for US dollars are traded at a haircut rate of almost 0%.
- The haircut rate level in cross-currency transactions to exchange Japanese government bonds for US dollars differs significantly from that of same-currency transactions due to foreign exchange risk.
- A time series of haircut rates have remained stable despite COVID-19 turmoil upsetting the financial markets.

Note: Latest data as at December 2021.
Regression Model

Analysis is conducted with an OLS dummy variable model.

\[ Haircut_j = \alpha_0 + \sum_k \alpha_{1,k} X_{j,k} + \sum_l \delta_l 1(d_{j,l} = l) + \sum_m \rho_m 1(p_{j,m} = m) + \sum_n \theta_n 1(s_{j,n} = n) + \varepsilon_j, \]

Where:

- **Explanatory Variables**
  - \( X_{j,k} \) are continuous variables (Transaction maturity days, Transaction amount, Repo rate, Network centrality)
  - \( d_{j,l} \) are dummy variables
    - Characteristics of Government Bond (Collateral quality, Residual maturity)
    - Transaction maturity (Open-end transactions)
    - Transaction type (Cross currency, Special collateral, etc.)
  - \( p_{j,m} \) are time fixed-effects on the transacting entities (Data reporter × Counterparty × Transaction reporting date)
  - \( s_{j,n} \) are fixed-effects on the transaction types (Jurisdiction of the bond × currency of the bond × Cash currency)

- **Fixed Effects**
Main Results

Characteristics of Government Bond

- **Collateral quality**: The higher the credit rating, the lower the haircut rate.
- **Residual maturity**: The haircut rate of government bond with a long residual maturity tends to be higher.

Transaction Maturity

- **Transaction maturity days**: Haircut rates increase with long transaction maturity, but the impact is not significant.
- **Open-end transactions**: Open-end transactions increase haircut rates.

Transaction Terms

- **Repo rate**: There is a positive correlation between haircut rate and repo rate.
- **Transaction amount**: The effect of transaction amount does not significantly affect the haircut rate.

Network Effect

- **Network centrality**: Low haircut rate for financial institutions near the center of the network.

Transaction Type

- **Cross currency**: Cross-currency transactions increase haircut rates.
- **Special collateral**: Haircut rate for GC transactions is higher than that for SC transactions.
Main Results

- Financial institutions near the center of the network can match borrowing and lending needs more efficiently.

- As the residual maturity increases, the haircut rate is pushed up along with the higher price volatility.

Transaction network in standard repurchase agreements using government bonds

Term structure of residual maturity of government bonds (excluding zero haircuts sample)
Conclusions

- First analysis of the market structure and haircut setting mechanism of SFTs in-depth using transaction data collected by the Financial Services Agency of Japan and the Bank of Japan.

- We determined that explanatory variables affecting credit, market, and liquidity risk, such as the collateral quality of government bonds, the residual maturity of government bonds, and the presence of foreign exchange risk, significantly impact haircut setting.

- Financial institutions closer to the center of the network tend to set lower haircut rates.

- The results are generally robust and of value to financial authorities and practitioners in trading and risk management of SFTs at financial institutions.

Thank you for your attention
How can big data improve the quality of tourism statistics? 
The Bank of Italy’s experience in compiling the “travel” item of the Balance of Payments

Andrea Carboni, Costanza Catalano and Claudio Doria, 
Bank of Italy

---

1 This presentation was prepared for the conference. The views expressed are those of the authors and do not necessarily reflect the views of the BIS, the IFC or the central banks and other institutions represented at the event.
How can big data improve the quality of tourism statistics?

The Bank of Italy’s experience in compiling the “travel” item of the Balance of Payments

Andrea Carboni, Costanza Catalano, Claudio Doria
Department of Economics, Statistics and Research, Bank of Italy

Abstract

In tourism statistics, the search for timelier and cheaper data sources than the traditional ones, like surveys, is becoming more and more important. In this paper, we investigate how mobile phone data (MPD), electronic payments data and internet search data (Google Trends) can improve the compilation of tourism statistics and the “travel” item of the Balance of Payments (BoP). We find that MPD have the characteristics to improve the estimates on the number of international travelers and to be integrated with the survey, although a constant interaction with the data supplier is required to define the phenomena to be caught. We highlight the limitations and the issues related to the use of electronic payment data for estimating expenditures in tourism statistics and we propose a model for producing timelier preliminary estimates for BoP purposes. Finally, we point out that Google Trends data can be used to complement the sample estimates of international travelers and to improve the quality of provisional data.

Keywords: Big data, tourism statistics, balance of payments, mobile phone data, payment data, Google Trends.

JEL classification: C20, C55, C80

Contents

How can big data improve the quality of tourism statistics? ................................................. 1
The Bank of Italy’s experience in compiling the “travel” item of the Balance of Payments ............................................................................................................................................. 1
Introduction ............................................................................................................................................... 2
The estimate of the “travel” item in the Italian balance of Payments ................................ 2
The mobile phone data experiment................................................................................................. 4
The payment statistics analysis.......................................................................................................... 8
The Google Trends experiment ....................................................................................................... 13
Concluding Remarks ............................................................................................................................ 18
1. Introduction

The use of big data is rapidly spreading in several fields as economics and statistics. National Central Banks play a role in this growing exploitation, as in general, official statistics follow a pressing and strictly defined calendar. Therefore, timely information, as that coming from big data sources, is very attractive and potentially useful for compilers. Moreover, big data can be of great help as a supplementary source whenever information from traditional sources is difficult to obtain, time demanding and burdensome to acquire. In 2014, the UN Statistical Commission, recognizing the relevance of these new data sources, established the Global Working Group on Big Data for Official Statistics to promote the use of big data for compiling official statistics.

Against this scenario, the Bank of Italy has carried out an in-depth analysis to understand whether, and how, big data can enhance the data production of the BoP “travel” item. This paper illustrates the results of this research, focusing on mobile phone data, electronic payment statistics (credit/debit cards) and web research information (Google trends), and discussing how the traditional approach for the compilation of the “travel” item can be improved by their use.

The paper is structured as follows. The next section describes the Bank of Italy’s traditional methodology for the compilation of the “travel” item. Sections 3, 4 and 5 respectively illustrate the three research paths – based on the use of mobile phone data, electronic payments statistics and Google Trends data - developed for improving and validating the compilation approach. The final section summarizes the main findings and conclusions.

2. The estimate of the “travel” item in the Italian Balance of Payments

The Balance of Payments (BoP) is a statement that summarizes the economic transactions between residents and non-residents during a specific period. The “travel” item of the BoP covers the monetary value of goods and services acquired in a country by non-resident travelers, in relation to visits to that country (BPM6 10.86),

---

1 E.g. EU Member States disseminate provisional statistics of Balance of Payment on a monthly basis at t+45 days.
4 Only transactions between residents and non-residents are relevant. The definition of “residence” is economic and not administrative: the country of residence of an international traveller is where the centre of her economic interest is located. To ease the exposition, we use for travellers the improper terms of “Italian” and “foreigner” when referring to their residence.
with the exception of the expenses for transport incurred to reach it, which are instead recorded under the “passenger transport” item. The BoP compilation standards of “travel” require a breakdown by counterpart countries and by purpose of the visit.

Countries can adopt different methodological approaches for compiling this item, based on the relevance of tourism in their economy, the characteristics of the border points, the administrative controls on incoming and outgoing flows and, of course, the budget constraints.

Since 1996, the Bank of Italy has been collecting the relevant information (number of international travelers, expenditures, length of the trips) through a sample survey carried out at border points; the data collected are then integrated, when available, with administrative sources.

From a methodological point of view, the survey consists of two operations, carried out at each of the selected border points: counting and interviewing.

Counting aims at estimating the reference population, i.e. the total number of travelers entering or leaving Italy, broken down by country of residence or destination, on a monthly basis. In a selected interval of time, all the travelers crossing the border are counted and their residence is registered. Since having permanent counting operations on all the borders is not feasible, a grossing up algorithm is necessary for estimating the total amount of the international travelers crossing Italian borders during the reference period. Where administrative data are available, as for airports, they integrate the sample survey.

The second main operation consists in interviewing a sample of the travelers passing through each selected border point. The interviews primarily collect data on the expenditures and other relevant aspects for BoP purposes (e.g. reason, counterpart country), but also gather information that allows a broader analysis on tourism related topics, such as the means of payment and the type of accommodation. The interviews are carried out at the end of the stay, which is when the memories of the traveler about the trip are the most recent and reliable and all the expenses have already been determined so no guess by the traveler is necessary. Interviews are realized through questionnaires: each questionnaire refers to the group

---

5 E.g. flight tickets, international train tickets, tolls etc. The expenses for transports within the visited economy are instead included in the travel item.

6 Personal vs. business travels.

7 In Europe, the main data sources for compiling the “travel” item are (frontier and households) sample surveys; some countries integrate these information with payment statistics as an additional source or for control purposes.

8 E.g. the number of international travellers published by airports, ports authorities and railroads companies.
of people, if any, that shares the expenses of the trip with the interviewed traveler (e.g. a family).  

At each border point, interviewing and counting are carried out, as much as possible, at the same time, so that the characteristics of the interviewed traveler are coherent with those of the counted sample. The information acquired with the questionnaires is then grossed up to the reference population, by taking into account the stratification variables listed in Table 1. Annually about 100,000 interviews and 1,000,000 counting operations are realized in more than 60 frontier points. This size ensures that the sampling error of the total international travel expenditure estimates is small and the statistics for the main partner countries are reliable.

<table>
<thead>
<tr>
<th>VARIABLE</th>
<th>LEVELS</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. Direction</td>
<td>2 (inbound, outbound)</td>
</tr>
<tr>
<td>2. Type of carrier</td>
<td>4 (road, rail, airport and seaport)</td>
</tr>
<tr>
<td>3. Frontier point</td>
<td>62 (22 roads, 4 rails, 25 airports, 11 seaports)</td>
</tr>
<tr>
<td>4. Day of data collection</td>
<td>number of days in the month (e.g. 31)</td>
</tr>
<tr>
<td>5. Time of the day</td>
<td>3 (first shift, second shift, third shift)</td>
</tr>
</tbody>
</table>

3. The mobile phone data experiment

Mobile phone data (MPD) are one of the most promising big data sources for the study of many social and economic phenomena and behaviors. Several pilot studies in the literature analyze the potentiality of MPD, e.g. for computing the population of an area (Deville et al., 2014), for estimating the population density (Ricciato et al., 2015), for traffic statistics (Janecek et al., 2015) for transport and urban planning (Lokanathan et al., 2016) and, finally, for travel statistics (Ahas et al. 2007; Ahas et al. 2008; Ahas et al., 2014). In this regard, the contributions of the Estonia Central Bank and the Banque de France have also to be mentioned.

While MPD data can provide a great amount of information (number of international travelers, proxy of the country of residence, locations visited, length of stay, etc.) they say nothing about the expenses, the main variable to be estimated in

---

9 For example, if the interviewed traveler shares the expenses with another traveler and they spend a total of 100 euros, we count two trips, each with a total expense of 50 euros.

10 An interviewed traveler is not necessarily part of the counted ones. For example, on road borders it is very difficult to interview the passengers of the vehicles counted while passing through the frontier; it becomes possible only when there are checkpoints and a collaboration with frontier authorities is arranged.

11 https://statistika.eestipank.ee/failid/mbo/valisreisid_eng.html
   https://statistika.eestipank.ee/failid/mbo/ky_mb2_eng.html

How can big data improve the quality of tourism statistics?

The BoP “travel” item. MPD can thus only be considered as a complementary source of information, useful to estimate the dimension and some characteristics of the reference population.

In 2018, the Bank of Italy started a test phase with the purpose of integrating MPD into the international frontier survey, in order to gradually replace the counting operations. Counting is in fact a costly and demanding activity, and this is particularly true for road borders, given the high number of this type of frontier points in Italy and the scarcity of administrative data, and for seaports, due to restricted access zones as the ones reserved to cruise ships. These problems might affect the quality of the grossing-up factors and hence of the estimated values.

MPD may represent an alternative, efficient and less costly data source to count travelers crossing the frontiers. The arrival of a foreign traveler at the Italian border is signaled by the connection of a mobile phone, with a SIM card\(^{13}\) issued by a non-resident phone operator, to the cells controlled by an Italian phone-operator. Likewise, the disappearance for some period of time of the signal of an Italian SIM card near the border would indicate that this traveler has gone abroad.

The Bank of Italy collaborated with one of the major Italian Mobile Network Operator\(^{14}\) (MNO) to develop an algorithm for the estimate of travelers’ inflows and outflows through each border point by exploiting the MPD. These data are not “ready to use” for BoP purposes and a close, constant cooperation between the Bank of Italy and the MNO has been necessary to define the best metrics to elaborate the raw data and achieve measures compatible with the BoP standards. For example, it is necessary to define the minimum docking time of a foreign SIM card to a cell located in Italy for it to be considered associated to an international traveler present in Italy. This problem is very relevant near the road borders due to handover effects.

Since each frontier point has specific features that should be incorporated in the final algorithm, a test phase was developed for two important Italian border points: the main airport of Rome (Fiumicino), which is the largest in Italy in terms of international traffic, and the highway frontier point of Tarvisio, one of the most relevant in the North-East of Italy.

For the Fiumicino airport, the traditional survey is supported by data provided by the company that manages the airport, Aeroporti di Roma (ADR). This source is used for correcting, by means of calibrated estimators (see Deville and Särndal 1992), the estimate of the total international flows derived from the counting operations, although it does not provide information on the residence of the passengers.

Tables 2 and 3 compare the MPD, the ADR\(^{15}\) statistics and the Bank of Italy’s official statistics on the Fiumicino airport for the period August 2018 - June 2019. The MPD

---

\(^{13}\) Subscriber Identity Module.

\(^{14}\) 31% of the market share in 2018.

\(^{15}\) Differences between the official data grand total and ADR statistics are due to minor adjustment in the Bank of Italy estimation process.
and the ADR grand totals are broadly aligned, with the MPD always larger than the official data (BI). As for the breakdown residents/non-residents, the number of Italian travelers estimated by the MPD is in line with the one estimated by the Bank of Italy (BI). On the other side, the number of foreign travelers estimated by the MNO is always greater than the one estimated by the BI.

Table 2 – Fiumicino airport: comparison between MPD and ADR statistics on the number of international passengers

<table>
<thead>
<tr>
<th></th>
<th>MPD (1)</th>
<th>ADR (2)</th>
<th>MPD/ADR %</th>
</tr>
</thead>
<tbody>
<tr>
<td>Aug-18</td>
<td>1,802,051</td>
<td>1,679,511</td>
<td>7.3</td>
</tr>
<tr>
<td>Sep-18</td>
<td>1,723,145</td>
<td>1,521,956</td>
<td>13.2</td>
</tr>
<tr>
<td>Oct-18</td>
<td>1,590,179</td>
<td>1,437,316</td>
<td>10.6</td>
</tr>
<tr>
<td>Nov-18</td>
<td>1,220,903</td>
<td>1,083,621</td>
<td>12.7</td>
</tr>
<tr>
<td>Dec-18</td>
<td>1,045,675</td>
<td>1,066,898</td>
<td>-2.0</td>
</tr>
<tr>
<td>Jan-19</td>
<td>1,113,629</td>
<td>989,903</td>
<td>12.5</td>
</tr>
<tr>
<td>Total</td>
<td>8,495,582</td>
<td>7,779,205</td>
<td>9.2</td>
</tr>
</tbody>
</table>

(1) Estimates based on mobile phone data.
(2) Aeroporti di Roma administrative data on passenger transits at Fiumicino airport.

Table 3 – Fiumicino airport: comparison between MPD and BI statistics on the number of international passengers

<table>
<thead>
<tr>
<th></th>
<th>BI (1)</th>
<th>MPD (2)</th>
<th>MPD/BP%</th>
<th>BI (1)</th>
<th>MPD (2)</th>
<th>MPD/BP%</th>
</tr>
</thead>
<tbody>
<tr>
<td>Aug-18</td>
<td>1,717,076</td>
<td>1,802,051</td>
<td>4.9</td>
<td>640,288</td>
<td>621,419</td>
<td>-2.9</td>
</tr>
<tr>
<td>Sep-18</td>
<td>1,574,571</td>
<td>1,723,145</td>
<td>9.4</td>
<td>445,444</td>
<td>516,638</td>
<td>15.5</td>
</tr>
<tr>
<td>Oct-18</td>
<td>1,380,639</td>
<td>1,590,179</td>
<td>15.2</td>
<td>423,402</td>
<td>449,204</td>
<td>6.4</td>
</tr>
<tr>
<td>Nov-18</td>
<td>1,053,956</td>
<td>1,220,903</td>
<td>15.8</td>
<td>392,909</td>
<td>466,087</td>
<td>18.6</td>
</tr>
<tr>
<td>Dec-18</td>
<td>1,017,503</td>
<td>1,045,675</td>
<td>0.8</td>
<td>506,510</td>
<td>417,920</td>
<td>-17.5</td>
</tr>
<tr>
<td>Jan-19</td>
<td>811,120</td>
<td>1,113,629</td>
<td>34.0</td>
<td>344,529</td>
<td>457,947</td>
<td>32.9</td>
</tr>
<tr>
<td>Total</td>
<td>7,594,865</td>
<td>8,495,582</td>
<td>11.9</td>
<td>2,754,512</td>
<td>2,929,115</td>
<td>6.3</td>
</tr>
</tbody>
</table>

(1) Bank of Italy official statistics.
(2) Estimates based on mobile phone data.

The estimate of the number of international travelers crossing Tarvisio border only relies on counting operations, due to the lack of complementary administrative sources.

Table 4 compares the Bank of Italy’s and the MPD statistics in this road border point on the same period: the differences are very large, and they are sharper for Italian travelers than for foreigners.
Further interactions with the mobile network operator led to a shortening (from four hours to 30 minutes) of the minimum time a foreign/Italian SIM card has to spend on the national/foreign territory in order to be considered an international traveler, and thus to a recalibration of the algorithm. This resulted in a new test, only involving the months of August and September 2020: the result showed a good convergence on the grand total between the two data sources, with a great improvement compared to the first release (Figure 1). On the other hand, the distribution between resident and non-resident travelers was still quite different, suggesting the need to continue investigating the causes underneath different estimates.

<table>
<thead>
<tr>
<th></th>
<th>TOTAL</th>
<th>ITALIANS</th>
<th>FOREIGNERS</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>BI</td>
<td>MPD</td>
<td>BI</td>
</tr>
<tr>
<td>Aug-18</td>
<td>2,005,595</td>
<td>980,066</td>
<td>-51.1</td>
</tr>
<tr>
<td>Sep-18</td>
<td>1,544,727</td>
<td>785,843</td>
<td>-49.1</td>
</tr>
<tr>
<td>Oct-18</td>
<td>1,026,265</td>
<td>416,988</td>
<td>-59.4</td>
</tr>
<tr>
<td>Nov-18</td>
<td>691,340</td>
<td>340,325</td>
<td>-50.8</td>
</tr>
<tr>
<td>Dec-18</td>
<td>600,309</td>
<td>285,953</td>
<td>-52.4</td>
</tr>
<tr>
<td>Jan-19</td>
<td>686,784</td>
<td>366,037</td>
<td>-46.7</td>
</tr>
<tr>
<td>Total</td>
<td>6,555,020</td>
<td>3,175,212</td>
<td>-51.6</td>
</tr>
</tbody>
</table>

4. The payment statistics analysis

Similarly to mobile phone data, electronic payment data are a promising source for the measurement and study of social and economic phenomena, including the production of statistics on national and international expenditure (Dubreuil 2017, Demunter 2017). Recently, they have started to be used for tourism statistics (Li et.
Al. 2018), in particular by international institutes and national central banks such as the Banco de Portugal (Coelho et. Al. 2011), the Banque de France16 and the Central Bank of Armenia (Yezekyan. 2018). Moreover, the European Central Bank recently approved a regulation17 on payment statistics also with the purpose of gathering data that the Eurozone countries could use for the compilation of their external statistics.18

Electronic payment data are attractive because of their timeliness, relative ease in collection and processing and moderate costs; moreover, their availability is not subject to high-impact perturbative phenomena like the Covid-19 pandemic. The steady increase of the share of electronic payments on total expenditure (Ardizzi et. Al. 2021) will keep strengthening the informative power of this source, although all the other possible means of payment such as transactions made by cash, bank transfers, etc. have to be estimated with other sources.

Against this background, the Bank of Italy conducted an explorative analysis in order to assess if and to what extent electronic payment data can contribute to the production of the “travel” item of the BoP and/or can be used for checking the consistency with the tourism statistics.

For this purpose, two databases were considered, provided by one of the main paytech companies operating in Italy, with data spanning from May 2014 to August 2021. The market share of this company was unknown, making impossible the grossing-up of raw data. One database contains all the electronic payments made by credit and debits cards on POS19 (physical database), while the other includes the online (e-commerce) transactions. Both databases are divided into acquiring and issuing: the first one contains the transactions made by foreigners cards on Italian POS and websites (potentially contributing to estimate the foreigners’ expenditures in Italy), while the latter includes the transactions on foreign POS and websites made by Italian cards (potentially contributing to estimate the Italians’ expenditures abroad).

Every record of the databases is made up of five variables: the date (day-month-year) of the payment, the Merchant Category Code (MCC) identifying the type of purchase20, the nationality of the bank emitting the payment card, the country of the POS/website where the payment has been made and the transactions amount in euro.21

There are major limitations in electronic payments data for the compilation of official statistics on travel:

1. The nationality of the bank issuing the card is just a proxy of the residence of the traveler;

17 Regulation ECB/2020/59.
19 Point Of Sale.
20 The available MCC are the following: clothing, hotels and restaurants, groceries, home, cash advance, work, retail, services, mobile web, travels and transports.
21 The amount is the aggregation of all the transactions sharing the same values of the first four variables.
2. Confidentiality issues allow the use of only aggregated data, which may increase the difficulties in discerning the transactions that are related to tourism from the ones that are not;

3. There is no information on the reason of the trip (business/personal), which is a mandatory BoP requirement;

4. There are difficulties in registering and correctly classifying the Digital International Platforms (DIP) transactions, in terms of misallocation issues for the counterpart country and of failure in recording some transactions. Three main examples could help understand the matter:

   I. The payment of a stay in Paris made by an Italian tourist on the Booking.com platform\(^{22}\) is recorded as a transaction from Italy to The Netherlands and not to France, as it should be recorded in the BoP;

   II. The payment on Airbnb\(^ {23}\) of an accommodation in Rome by a French traveler is recorded as a transaction from France to Ireland, thus not appearing in our database, although it should be recorded in the BoP;\(^ {24}\)

   III. The payment on Airbnb of an accommodation in Rome by an Italian traveler is recorded as a transaction from Italy to Ireland, although it refers to a domestic trip and thus should not be recorded in the BoP.

Figure 2 compares the official BoP data on foreigner travelers’ expenditure in Italy (only by means of electronic cards) with the grand totals of, respectively, the electronic payments recorded in the physical acquiring database, in the e-commerce acquiring database, and the sum of the two. The level of the e-commerce transactions is much lower than the other time series. Indeed, it does not cover the transactions of item 4-II: the large use of these platforms, which mostly have foreign headquarters, can explain its negligible levels.

Figure 3 compares the official BoP data on Italian travelers’ expenditure abroad (only by means of electronic cards) with the grand totals of, respectively, the electronic payments from the physical issuing database, the e-commerce issuing database and the sum of the two. The e-commerce time series shows higher levels than for the acquiring side, although it does not show the typical seasonality of the tourism phenomena, which has peaks in the summer. This is probably because such database contains large amounts of on-line transactions that are not connected to tourism, as the purchases of goods on Amazon, Ebay, etc. The low granularity of the database does not always allow to distinguish them, as some categories contain both transactions that are related to travel and transactions that are not.

---

\(^{22}\) Whose legal headquarter is in The Netherlands.

\(^{23}\) Whose legal headquarter is in Ireland.

\(^{24}\) The digital platform can carry out a further transaction with an Italian counterpart, but not necessarily using a credit card.
In an attempt to push further the exploratory analysis, the following Machine Learning models have been tested: Ridge and Lasso models\textsuperscript{25}, regression trees and boosted regression trees. Due to the relative short length of the whole payment data time series\textsuperscript{26}, the years 2015-2017 have been used as training set, the year 2018 as validation set\textsuperscript{27} and the year 2019 as test set, in order to verify the model out-of-

\textsuperscript{25} With and without enforcing positive coefficients.

\textsuperscript{26} 60 months from January 2015 to December 2019, plus 20 months (January 2020-August 2021) during the Covid-19 pandemic.

\textsuperscript{27} It was used to select the best parameter λ for the Lasso and Ridge models minimizing the MSE on the validation set.
sample using the MSE index. Moreover, the Covid-19 years 2020-2021 were used as supplementary test set for verifying the robustness of the model to external shocks. In each model, the dependent variable is the BoP travel item total, while the independent variables are all the physical MCC data (at lag 0) plus all the e-commerce MCC data for all lags from 0 to -4.

Table 5 reports the performance of such models in terms of the MSE index on the validation, test and Covid set. On the acquiring side, almost all the models show a quite good performance; in particular, the LASSO model with positive coefficients and the regression tree shows the smallest MSE on the test set and perform quite well in the Covid one. On the issuing side, the performance of each model is worse than in the acquiring case, as expected. The best results are obtained by the LASSO models, which have an unexpected good performance in the Covid set.

Table 5- Model performance in predicting the BoP ‘travel’ item grand total on the different sets

<table>
<thead>
<tr>
<th></th>
<th>ACQUIRING</th>
<th>ISUING</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>MSE val</td>
<td>MSE test</td>
</tr>
<tr>
<td>Ridge</td>
<td>0,09</td>
<td>0,76</td>
</tr>
<tr>
<td>Lasso</td>
<td>0,04</td>
<td>0,29</td>
</tr>
<tr>
<td>Lasso pos. coeff</td>
<td>0,04</td>
<td>0,3</td>
</tr>
<tr>
<td>Regression tree</td>
<td>0,2</td>
<td>0,28</td>
</tr>
<tr>
<td>Boosted tree</td>
<td>0,1</td>
<td>0,32</td>
</tr>
</tbody>
</table>


Figure 4 shows the plots of the forecasts in the test and Covid sets compared with the official BoP figures. The graphs confirm what was already pointed out: in forecasting foreigners’ travel expenses in Italy, we obtain a good performance on the test set, while the forecast behaves quite poorly in the subsequent years affected by the Covid pandemic. On the other hand, the forecast of the Italians’ travel expenses abroad is worse in the test set, but surprisingly good in the Covid one, as the trend is fully captured.

28 For the regression tree model, both 2018 and 2019 are used as test sets, fixing the maximum high of the tree to 4.
29 We will call it the ‘Covid set’.
30 Mean Squared Error. The smaller the MSE is, the better the model prediction is.
Figure 4 – Forecasts on test set and Covid set for selected models

Note: Test set on the left of the black dashed vertical line, Covid set on the right.
5. The Google Trends experiments

The third experimentation relies on the use of Google Trends as a complementary source to estimate the “travel” item in the BoP compilation process, in particular to assess the number of international travelers in Italy.

Google Trends (GT) is a website provided by Google that reports the popularity of search queries in the Google search engine over time and across various regions of the world. The popularity of a given query is measured by an index between 0 and 100 (the maximum frequency). The data are collected and aggregated continuously on a daily, weekly or monthly basis. One can visualize the popularity of the selected query by specifying the state or region, the category they belong to, and the time frame of interest. Timeliness is one of the main advantages of this website, as data are updated almost in real time.

In order to understand if this kind of data can be usefully employed, a specific predictive exercise was developed with the aim of forecasting the number of travelers visiting Italy in the period from January 2006 to May 2019. In particular, the tourist flows from the most important counterpart countries in terms of arrivals - France, Germany, United Kingdom, United States and Spain - were considered.

The predictive variable of GT was defined by considering the frequency of the search queries performed in the aforementioned countries that contain the word “Italy” in the category “Travel”. For each of these selected countries, a seasonal AR(1) process was used for modelling the number of travelers \( N_{c,t} \) arrived in Italy during month \( t \) from country \( c \) according to the Bank of Italy’s tourism survey, where the \( l \)-period lagged Google Trends index \( GT_{c,t-l} \) is included as an exogenous regressor:

\[
N_{c,t} = \phi_0 + \phi_1 N_{c,t-1} + \phi_{12} N_{c,t-12} + \beta GT_{c,t-l} + \varepsilon_{c,t} \tag{4}
\]

The most suitable lag of the GT index is chosen by minimizing errors of the out-of-sample forecasting performance, measured in terms of Mean Squared Error (MSE) reduction. In particular, the months between September 2012 and May 2019 have been considered to compare the observed value and the one-step-ahead forecasts, with an expanding windows approach. In all cases, except for France where the coefficient \( \beta \) is not statistically different from zero, the GT index increased the performance of the predictive model.

Figure 5 shows how the ratio between the MSE of specification (4) and the one obtained using the model without \( GT_{c,t-l} (\beta = 0) \) depends on the lag for the different countries considered.

---

31 https://trends.google.com/trends/?geo=IT
32 More than 20 categories of search, which helps avoiding multiple meanings for the chosen query.
33 Adding an observation at each step.
The contemporaneous variable $\text{GT}_t$ ($l=0$) is the best predictor for Germany and Spain, while the variable at lag $l=4$ and $l=6$ minimizes the MSE for UK and US respectively. These last results seem only partially reasonable: US travelers may have to organize their trips towards Italy more in advance than German and Spanish travellers; moreover, it is possible that Google Trends classifies in the Travel category web searches performed by tourists during their travel; this should increase the weight for lag $l=0$ in the model; less clear is the situation for UK, where there is not an intuitive explanation for a better performance of lag $l=4$ in comparison to smaller lags.

The estimates of specification (4) for each country involved in the exercise are shown in Table 6: the GT index is always highly significant and the model indicates a good in-sample fit, measured by a high value of the $R^2$. However, each time series presents a strong auto-regressive component and the marginal contribution of the GT index is significant only for Spain. The negative sign of GT coefficient in the UK and US regressions means that the variable is not robust enough for these two countries, confirming the doubts in the interpretation of the optimal lag.

The number of travellers at lag 1 and 12 are significant at 95% for all the considered countries.

For Spain, the $R^2$ adjusted is 0.67 in the model without GT, and 0.77 in the one with the variable included. For the other countries, the $R^2$ is near to 0.9 in the model with only the AR component and the addition of the GT index only increases it of around 0.01.
Table 6 – Estimates of the model for different countries

<table>
<thead>
<tr>
<th></th>
<th>DE</th>
<th>ES</th>
<th>UK</th>
<th>US</th>
</tr>
</thead>
<tbody>
<tr>
<td>$N_{t-1}$</td>
<td>0.20***</td>
<td>0.34***</td>
<td>0.33***</td>
<td>0.13***</td>
</tr>
<tr>
<td></td>
<td>(0.04)</td>
<td>(0.05)</td>
<td>(0.05)</td>
<td>(0.04)</td>
</tr>
<tr>
<td>$N_{t-12}$</td>
<td>0.73***</td>
<td>0.46***</td>
<td>0.76***</td>
<td>0.89***</td>
</tr>
<tr>
<td></td>
<td>(0.05)</td>
<td>(0.05)</td>
<td>(0.05)</td>
<td>(0.04)</td>
</tr>
<tr>
<td>$GT_{t-1}$</td>
<td>6.01***</td>
<td>2.11***</td>
<td>-1.50***</td>
<td>-0.91***</td>
</tr>
<tr>
<td></td>
<td>(1.17)</td>
<td>(0.29)</td>
<td>(0.44)</td>
<td>(0.24)</td>
</tr>
<tr>
<td>Const</td>
<td>-182.49***</td>
<td>-22.78*</td>
<td>29.87*</td>
<td>38.84***</td>
</tr>
<tr>
<td></td>
<td>(43.34)</td>
<td>(12.07)</td>
<td>(15.52)</td>
<td>(13.93)</td>
</tr>
<tr>
<td>$R^2$</td>
<td>0.92</td>
<td>0.77</td>
<td>0.89</td>
<td>0.92</td>
</tr>
</tbody>
</table>

* p<0.10, ** p<0.05, *** p<0.01

In order to examine the predictive performance of the model Figure 6 compares, for each selected country, the observed value of the number of travellers to Italy and the one-step-ahead predicted levels in the out-of-sample period September 2012 - May 2019. The model seems to capture well the fluctuations of the phenomenon and the main turning points.

All in all, although this data source proved to be interesting during the analysed period, the Covid-19 pandemic pointed out its limits. Indeed, in March 2020 we witness a peak of search queries including the word “Italy” (see Figure 7 for selected countries), while in that month Italy was blocking the tourist inflow because of the pandemic; such peaks very likely reflect the interest by the users in understanding the developing of the pandemic or in checking whether travelling to Italy was still doable or safe, even if were not necessarily followed by actual travels. Indeed, in presence of extraordinary events, the Google classification seems to be less effective and the risk of outliers, given by false positive searches not related to tourism, increases significantly.

Moreover, since the use of Google Trends strictly depends on the keywords included in the analysis, the use of other words as search queries, for example referring to specific Italian locations, could generate more accurate results.

Figure 6 – Observed (solid line) and predicted (dashed line) number of travelers to Italy from Germany, Spain, United Kingdom and United States

---

36 This might explain why the peak in search queries appears also by considering only the GT category “Travel”.

How can big data improve the quality of tourism statistics? 15
How can big data improve the quality of tourism statistics?
How can big data improve the quality of tourism statistics?

Figure 7 - Search queries on Google including the word “Italy” from German, Spanish, UK and USA users
6. Concluding Remarks

In the recent years, the Bank of Italy has carried out several experimental analyses in order to explore the possibility of integrating big data in the production of official statistics, in particular for compiling the “travel” item of the Balance of Payment.

The data that have been tested are appealing for their extraordinary timeliness and the amount of information offered, although they are very far away from being ready to use. Indeed, they need adjustments in order to define metrics that are coherent with the standards and the official definitions. The experiments often required adopting a trial and error approach to align these metrics to the prefixed standards, and making strong assumptions that could potentially affect the results.

According to our tests, mobile phone data seem to be the most suitable ones to be integrated with the frontier survey, as they are able to produce a broadly reliable estimate of the number of international travelers crossing the Italian borders, thus potentially replacing, at least partially, the counting procedures in the Bank of Italy frontier survey. The Bank of Italy is already moving in this direction.

The other big data sources analysed, electronic payments data and Google Trends data, showed more limitations and drawbacks.

Electronic payment data proved useful for achieving a preliminary estimate of total expenditures related to travel, as they are timelier than survey data. However, at this stage they can be used, at least for BoP, only for checking purposes. On the other hand, considering the informative potential of this source, we will continue exploring how to overcome the main problems by identifying the features that the data should have to be fully usable.

The Google Trends index proved to be useful for estimating the number of international travelers. But the sensitivity of the index to extraordinary circumstances like the Covid-19 pandemic needs to be further investigated before considering the integration of such index in the compilation process.
References


How can big data improve the quality of tourism statistics?

The Bank of Italy’s experience in compiling the “travel” item of the Balance of Payments

Andrea Carboni, Costanza Catalano, Claudio Doria
Department of Economics, Statistics and Research – Bank of Italy
Big data for tourism statistics. The Italian case

Tourism statistics: number, expenditures and nights spent of
- Foreigners travelers visiting Italy (the reporting country)
- Italian travelers visiting abroad

BoP standards: expenditures by counterpart countries, business vs. personal trips, border/seasonal workers, international transports...

Sources: sample survey at border points (since 1996)

Drawbacks: costly, time-demanding, subjected to external factors (e.g. the covid-19 pandemic)

Big data: timelier, cheaper, less impacted by external shocks

Experiments on:
- Mobile Phone data
- Electronic payment data
- Internet search queries (Google Trends)
Mobile Phone data

May represent an alternative data source to count travelers crossing the border points.

Only complementary source, no info on expenditures

- Arrival of a foreign traveler: signaled by the connection of foreign SIM cards to the cells controlled by an Italian network operator;
- Departure of an Italian traveler abroad: disappearance of the signal of an Italian SIM card near the border.

Nationality of the company issuing the SIM card = proxy for the traveler’s country of origin

Collaboration with one of the major Italian Mobile Network Operator:
- Algorithm for estimating travelers inflows and outflows
- Constant cooperation necessary to achieve BoP standards (ex. minimum docking time due to handover effect)
- Tests on two main Italian border points (Fiumicino airport and Tarvisio highway)
Tarvisio highway: huge discrepancies (order of 50%), needed a second test where the docking time was shortened.

Nowadays: The Bank of Italy has partially replaced the counting procedures in the frontier survey by MPD.
Electronic payment data

**Database:** aggregated by date, nationality of bank emitting the card, type of purchase (10 categories), country of the POS/website

- Foreign card & Italian POS/website ➔ Foreigners expenditure in Italy
- Italian card & foreign POS/website ➔ Italian expenditure abroad

**Main drawbacks in using payment data for BoP statistics:**

- The nationality of the card is a proxy of the traveler’s residence
- Confidentiality issues allow only aggregated data
- No info on the reason of the trip (business/personal)
- Difficult to correctly classify the Digital Platform transactions:
  - Payment of a stay in Paris by an Italian on Airbnb is recorded as from Italy to Ireland and not to France;
  - Payment of a stay in Rome by a French on Airbnb is recorded as from France to Ireland, thus not appearing in the database;
  - Payment of a stay in Rome by an Italian on Airbnb is recorded as from Italy to Ireland, despite it is a domestic trip.
Electronic payment data

Market-share unknown ➞ Impossible the grossing-up of raw data

Tested some forecast models:
Ridge, Lasso, regression trees, boosted regression trees


Best performance in terms of MSE on the test set:

![Graphs showing travel expenditures in Italy and abroad with LASSO predictions.](image-url)
Internet search data: Google Trends

**Google Trends index:** reports the popularity of a given query in a given time period, country and category. It spans from 0 to 100.

Can the GT index be used to improve the provisional estimates on the number of travelers?

**The model:** seasonal AR(1)

\[ N_{c,t} = \phi_0 + \phi_1 N_{c,t-1} + \phi_{12} N_{c,t-12} + \beta GT_{c,t-l} + \varepsilon_{c,t} \]

- queries including the word 'Italy', category= 'Travel'
- \( N_{c,t} \) = number of travelers from country c in month t
- one-step ahead forecast with expanding windows approach
- lag for GT index chosen by minimizing the out-of-sample MSE
- five countries: France, Germany, Spain, UK and USA

**Results:** In all cases the GT index increased the performance of the predictive model, except for France where \( \beta \) was not statistically different from zero.
Limits: peak of search queries in March 2020 while Italy was blocking the tourist inflow. In presence of extraordinary events the Google classification seems to be less effective with high risk of outliers.
Take-away

All the data sources needed adjustments in order to define metrics that are coherent with the BoP standards.

**Mobile phone data:**
- the most suitable ones to be integrated with the frontier survey in the estimate of the number of international travelers
- Bank of Italy uses MPD for tourism statistics since the end of 2020

**Electronic payment data:**
- useful to achieve a preliminary timelier estimation of the total expenditure of the “travel” item
- For now, the mentioned relevant issues make it usable only for checking purposes

**Google trends:**
- useful as explanatory variable for estimating the number of international travelers
- possible noise of the index could be misleading. Use of other/more words as search queries could generate more accurate results.
Thank you for your attention!

...Questions?
Update on the work of the International Network for Exchanging Experiences on Statistical Handling of Granular Data (INEXDA)

Stefan Bender, Jannick Blaschke and Christian Hirsch,
Deutsche Bundesbank
Update on the work of the International Network for Exchanging Experiences on Statistical Handling of Granular Data (INEXDA)

Prepared by the chair of INEXDA

Abstract

INEXDA provides a platform for administrative data producers to exchange practical experiences on the accessibility of granular data, metadata, and data protection techniques. In this paper, we provide an update on the progress that INEXDA has made since 2018. This includes creating a set of machine-readable metadata items describing access procedures and conducting a survey of INEXDA members to take stock of technical, administrative, and organisational features of procedures for accessing granular data for research purposes. The results provide a comprehensive picture of existing procedures and may serve to provide insight for institutions on the cusp of setting up similar facilities.

Keywords: INEXDA, data sharing, research data centre, data access, research, granular data, exchange

Content

1. Introduction ..........................................................................................................................................................2
2. Working Group on Data Access ..........................................................................................................................3
   2.1 Workshop and the need for unified terminology .........................................................................................3
   2.2 The resulting metadata schema – the annodata schema .............................................................................4
3. INEXDA Survey on Access Modes ...................................................................................................................5
   3.1 INEXDA Working Group on Modes of Data Provision ...............................................................................5
   3.2 Main results on access modes .....................................................................................................................6
   3.3 Main results on provided data .....................................................................................................................7
   3.4 Main results on organisational and technical access modalities .................................................................7
   3.5 Main results on output checking ..................................................................................................................8

1 As of July 2022 Stefan Bender of Bundesbank is the current chair of INEXDA. Jannick Blaschke and Christian Hirsch, both Bundesbank, contributed to this report. The authors would like to thank our colleagues of the INEXDA network, particularly those who participated in the working groups, for their valuable suggestions and feedback. All views expressed in this report are personal views of the authors and do not necessarily reflect the views of INEXDA and its members, Deutsche Bundesbank or the Eurosystem.
1. Introduction

This paper provides an update on the International Network for Exchanging Experiences on Statistical Handling of Granular Data (INEXDA). We presented the network in a similar paper in 2018. In this paper, we would like to describe the work that has been accomplished since 2018. The network has seen its membership increase in both number and diversity, as it now includes members from outside of Europe as well as beyond national statistical institutes.

INEXDA was founded to provide a platform for interested parties to exchange their experiences in the statistical handling of granular data. The idea originated from the global financial crisis of 2008-09, which showed that certain problems may only be detected using very granular data on current economic developments. The network was originally set up to promote the G20 Data Gaps Initiative II, particularly Recommendation 20, which addressed the accessibility of granular data. Moreover, INEXDA acknowledges and supports the work on data sharing by the Irving Fisher Committee on Central Bank Statistics.

The idea of the network has always been to provide a platform for exchange. However, we also wanted to provide hands-on practical solutions intended for use in institutions that provide access to granular data or are planning to do so. One example of these solutions already existing in 2018 was the INEXDA metadata schema for describing granular data (Bender, Hausstein, and Hirsch, 2019). This schema builds on an existing metadata standard data documentation initiative (DDI) and has been modified to better accommodate the particularities of granular data.

However, are there any differences in describing macro data and aggregated data? While there are indeed differences, it is possible to take metadata designed for macradata as a reasonable starting point. One difference is that some metadata items become more important when moving from granular to (aggregated) macro data. For granular data, for example, metadata items describing identifiers are very important, as granular data usage typically requires multiple datasets to be combined for analysis.

Due to its increasing member size, INEXDA uses specialised working groups to advance its agenda. We will therefore describe recent developments largely on the basis of the output of these working groups. First is the Working Group on Access.
Modes. Access modes define the ways in which access to data is provided, which could, for example, be achieved via download or on-site access at the premises of the institution providing the data.

Providing access to granular data also involves facing significant legal and technical challenges related to, amongst other aspects, safeguarding statistical confidentiality. As a result, INEXDA decided to organise a workshop bringing together senior experts from central banks, national statistical institutes and international organisations to share their experiences on real-world approaches to facilitating access to granular data for analytical purposes. This workshop served as a starting point for the group’s work.

Building on the outcome of the workshop and the output of the working group, INEXDA released a second metadata schema aimed at describing access procedures in a machine-readable way (INEXDA Working Group on Data Access, 2020). This schema, known as the “annodata schema”, works in much the same way as metadata schemas that describe data, but has instead been designed to describe the procedures for accessing the data. The annodata schema is thus both compatible with and complementary to the INEXDA metadata schema described above.

Building on the annodata schema a working group conducted a survey among INEXDA members on currently available modes of data provision for external scientific researchers. The focus of the survey was to generate insights on the practical approaches employed by each participating institution. While generating useful information on the many particularities this very detailed approach also hampered the comparison of the responses of individual members.

This paper is organised as follows: Section 2 describes the annodata schema in more detail. We put special emphasis on access modes, which represent the smallest set of information needed to describe access to granular data. Section 3 describes the results of a survey of INEXDA members on the access modes used by member organisations to provide access to granular data for scientific research purposes. Section 4 presents the plans for the coming years and concludes.

2. Working Group on Data Access

2.1 Workshop and the need for unified terminology

Providing access to granular data for scientific research and analytical purposes requires developing a set of procedures that define how authorised personnel ought to use the data. Spanning a project’s life cycle from request to publication, these procedures must ensure ongoing compliance with the legal framework and also satisfy the technical, organisational, and administrative requirements set out by the data provider. The INEXDA Working Group on Data Access focused on (i) taking stock and (ii) identifying common features and differences among these procedures.

As one of its first activities after being established, the group organised a workshop aimed at bringing together experts from central banks, national statistical institutes and international organisations to share their experiences on real-world approaches to providing access to granular data. Presentations revolved around the topics of (i) “data discovery centres” (i.e. information on where to find the data), (ii)
“data access centres” (i.e. how to grant access to the data), and (iii) “data hubs” (i.e. access to multi-source remote data).

As a result of the workshop, it became clear that the working group could not immediately start tackling its original task of taking stock of and examining existing access procedures, as it still had to complete some preliminary work. This included, first, overcoming the challenges posed by differences in legal, organisational and technical conditions as well as different terminology that complicated the discussion on data access procedures. The working group therefore started by defining unambiguous terminology, including definitions for “access protocol”, “access mode” and “access regime”.

Second, the working group’s tasks required documenting and processing a wealth of institutional and data-specific requirements immanent in procedures for accessing granular data. Drawing on existing metadata schemas for this task proved insufficient, as they typically focus on the question of “How are the data produced?” rather than help to answer the question of “How is access to the data managed?”. The working group therefore used its newly established common taxonomy to describe the set of legal, technical, and organisational rules that make up the procedures for accessing granular data.

To our knowledge, this is the first machine-readable description of data access procedures and it stands in clear contrast to the predominantly free-text descriptions used in existing metadata schemas. This new metadata schema is therefore not intended to replace any existing metadata schemas, but rather to complement them with well-structured information on data access procedures.

2.2 The resulting metadata schema – the annodata schema

This metadata schema, which INEXDA has named the “annodata schema” (derived from “annotation to data”), comprises three interrelated dimensions. The first dimension covers the dataset aspect of access procedures. This dimension collects all information needed to unambiguously identify a dataset and its origins, and connects datasets to access modes (e.g. on-site access, remote execution). The second dimension contains all information regarding the aggregation of granular data, which is especially important, as access to granular data generally involves combining different datasets. Finally, the third dimension comprises information on users and projects.

A pivotal concept in the annodata schema is the “access mode”. An access mode connects inputs to outputs. In the example below, a function with three inputs produces one output. The inputs are:

1. “Who?” – Information on the user (e.g. internal or external to the organisation)
2. “What?” – Information on the dataset (e.g. anonymised or perturbed)

---

5 See Bender et al. (2021) for a more thorough discussion on this point.

6 For more information on the motivation and rational behind the annodata schema, see the final report of the Working Group (INEXDA Working Group on Data Access, 2020).

7 A complete version of the annodata schema, including a description of all metadata items, can be found in the final report of the Working Group (INEXDA Working Group on Data Access, 2020).
3. “Why?” – Information on the purpose of use (e.g. scientific research)

As the output, these three inputs determine “how” a specific user requesting access to a specific type of dataset for a specific type of use can access the data. Examples of “how” could include “secure on-site access” in a research data centre or “remote access”. Of course, the inputs could also produce a nil return as their output if no corresponding access mode exists.

The decision regarding which inputs lead to which output comes directly from the legal framework or its interpretation by the institution providing the data. For example, the legal framework could state that external researchers may use anonymised data for scientific research if they access the data within a secure environment. Here, the contribution of the annodata schema is to convert these rules from text form into machine-readable metadata.

With the common terminology set out, the Working Group on Data Access provided the basis for future discussions within INEXDA and facilitated the survey on access modes that was conducted by the subsequent working group. The next section presents the main results of this survey.

3. INEXDA Survey on Access Modes

3.1 INEXDA Working Group on Modes of Data Provision

In 2020, INEXDA launched its third working group, which focused on “Modes of Data Provision”. The group was chaired by the Banque de France and its main objective was to conduct a survey of INEXDA members on the existing modes of data provision for external scientific researchers. The survey results would then allow common features and differences with regard to access procedures amongst INEXDA members to be identified. The survey was conducted in 2020 and 2021 with the participation of eight countries as well as both the European Central Bank and Eurostat.

The survey builds on the INEXDA annodata schema and the terminology established therein. For this reason, the survey itself was conducted at the level of access modes. The overarching goal of the survey was to obtain insights into the practical approaches employed by each participating institution. The focus on access modes provided the detailed information needed to draw conclusions from the responses of individual members.

This very detailed approach reflected all particularities of the practical approaches to access to granular data stemming from the legal, organisational and technical frameworks. However, while useful for obtaining insights, these particularities also make it more difficult to compare the responses of individual members. For example, differences in the legal framework might lead to different interpretations of an access mode. This section therefore focuses on presenting the main results and lessons learned from the survey. In the future, INEXDA will conduct an updated survey to capture the effects of the COVID-19 pandemic.

---

8 The Deutsche Bundesbank (Germany), Banco de España (Spain), Banque de France (France), Banca d’Italia (Italy), Banco de México (Mexico), Banco de Portugal (Portugal), Bank of Russia (Russia), and Central Bank of the Republic of Turkey (Turkey).
3.2 Main results on access modes

The survey found that most respondents provided access to data in several ways, often combining the provision of Public Use Files (PUFs) with on-site solutions. Among the ten participating institutions, eight offered multiple ways to access data. Access modes differed with respect to the available dataset or degree of anonymisation, with formal anonymisation being the most commonly used method. However, all institutions granted access to their data for scientific research in some form.

### Distribution of access modes

<table>
<thead>
<tr>
<th>Access Mode Type</th>
<th>Number of Institutions</th>
<th>Number of Access Modes</th>
</tr>
</thead>
<tbody>
<tr>
<td>(Secure) on-site access</td>
<td>9</td>
<td>13</td>
</tr>
<tr>
<td>Remote execution</td>
<td>3</td>
<td>4</td>
</tr>
<tr>
<td>Remote access</td>
<td>3</td>
<td>3</td>
</tr>
<tr>
<td>Scientific Use Files</td>
<td>3</td>
<td>3</td>
</tr>
</tbody>
</table>

* Ten INEXDA members participated with 1-n possible access modes each. Public Use Files were excluded from this figure.


As Figure 1 shows, on-site access and secure on-site access, where researchers work in a secure environment on the premises of the institution providing the data, were the most frequent types of access. Three institutions offered remote execution, remote access and Scientific Use Files.

The number of researchers per access mode varied between institutions, but, in general, more researchers accessed data through the two download modes, although the actual usage of Public Use Files is hard to determine. Across institutions and access modes, the average number of researchers per project was between one and two. Access was granted primarily to researchers affiliated to universities or research institutions, while financial and commercial firms were able to access data solely through PUFs.

---

9 Secure on-site access: Provision on the premises of the institution in a dedicated secure environment (data room). Remote execution: Researcher sends a code and the RDC sends back results. Remote access: Researcher can access data stored in another location remotely from their own institution in a controlled environment. Download: Files are available for download and copy and are used according to the terms of use (e.g. Scientific Use Files or Public Use Files).
3.3 Main results on provided data

As nine of the ten respondents of the survey were central banks, most of the data made available to researchers were on the topic of “Economy and finance”, followed by “Industry, trade and services”. This is depicted in Figure 2. Within the “Economy and finance” topic area, access was most often granted to data on “Firms and banks”.

The majority of these datasets were provided in standardised form, so they were identical across research projects. The two biggest advantages of standardised datasets as compared to customised datasets are the economy of resources when preparing the data and the significantly easier reproducibility of research results. The latter is achieved, in particular, by the options for tailoring metadata and access rules perfectly to the dataset and uniquely identifying and citing the datasets with a Digital Object Identifier (DOI). However, some datasets were also provided in customised form, for example due to legal constraints.

When asked about the handling of external data brought in by the researchers themselves, a predominantly uniform approach could be observed. External researchers were usually allowed to bring their own data, but these were checked by the institution providing the data before being made available in the project. One respondent even offered automatic import of external data.

3.4 Main results on organisational and technical access modalities

The survey showed that there were different approaches to the governance of data for research access. Depending on the institution, different decision bodies were involved in the process and there were a variety of administrative measures to ensure
that no confidential data was ever disclosed. All respondents declared that, for all data other than Public Use Files, researchers needed to sign some form of confidentiality agreement.

Since the technical environment for data access is very much related to the respective access mode, the working group decided to limit the scope at this point to on-site access. Currently, data access is granted mainly via file systems, although some institutions have already switched to more advanced systems, such as dedicated servers or Hadoop platforms. Researchers can usually access the data via specific workstations or virtual machines. Regarding the available software, researchers can mostly choose between Stata, Python, R, and, in some cases, also Matlab and SAS.

3.5 Main results on output checking

All respondents of the survey agreed that, for data that is not accessed via download, output control is mandatory. This is usually performed by the researchers themselves and then checked by the institution providing the data. Only results that are compliant with the output rules can be taken out of the institution’s environment. The output check is usually carried out manually with some semi-automated tools providing assistance. These tools may, for example, calculate whether the release of a specific result, such as a summary table or regression, would lead to a disclosure problem.

4. Conclusion and future plans

This paper provides an update on the work that INEXDA has accomplished. Since 2018, INEXDA has experienced growth in its number of members, which has led to the composition of its membership becoming more diverse. The network devised the annodata schema, a machine-readable set of items for describing data access procedures intended to complement existing metadata schemas that predominantly utilise free-text descriptions. Finally, a survey conducted among members provided information on the many particularities in practical approaches to granular data access.

For the future, INEXDA plans to use the outcomes of the annodata schema and the survey to discuss possible potential for harmonisation with regard to access procedures within INEXDA. In addition, the results facilitate the development of common best practices for data sharing and the handling of granular data. This would be extremely helpful for any institution planning to establish a new access mode, for example, and looking for reference. Of course, these considerations need to acknowledge the limits set by the prevailing legal, technical and procedural frameworks in which INEXDA members operate.

Finally, INEXDA will continue supporting international initiatives aimed at strengthening micro data access. For instance, INEXDA will support the new G20 Data Gaps Initiative, for which the IMF, the Financial Stability Board and the Inter-agency Group on Economic and Financial Statistics are currently developing a workplan.
References


Update on the work of the International Network for Exchanging Experiences on Statistical Handling of Granular Data (INEXDA)

Stefan Bender
INEXDA chair, Deutsche Bundesbank

Christian Hirsch
INEXDA secretary, Deutsche Bundesbank

The views expressed here do not necessarily reflect the opinion of the Deutsche Bundesbank, the INEXDA network, or the Eurosystem.
INEXDA and its members

International Network of Exchanging Experiences on Statistical Handling of Granular Data

General mission

- General mission is to promote data sharing and data access
- Promoting the G20 Data Gaps Initiative II, in particular recommendation 20, addressing the accessibility of granular data. INEXDA is mentioned in a G20 paper
- Acknowledging and supporting the work on data sharing of the Irving Fisher Committee on Central Bank Statistics

Organisation

Current chair of INEXDA is Stefan Bender of the Deutsche Bundesbank.

Current members

- Banco Central de Chile
- Deutsche Bundesbank
- Bank of England
- Office for National Statistics
- Banco de España
- European Central Bank
- Eurostat
- Banque de France
- Banca D’Italia
- Banco de México
- Banco de Portugal
- Banco Poštovní Česka
- Türkei Cumhuriyet Merkez Bankası

2022 - Update on the work of INEXDA
INEXDA’s working arrangement

Micro data access is a multifaceted field. INEXDA’s focus is on developing, applying and disseminating new ideas.

For this it uses dedicated working groups to cover diverse topics in the field of micro data access:
Eight types of annodata

1. Access regime
2. Database
3. Dataset family
4. Record linkage
5. Combining restrictions
6. Global rules
7. Researchers
8. Research projects

1-3 at dataset family level
4-6 at global level (i.e. irrespective of researcher affiliation, research field, and access mode)
7-8 at project level
How micro data access could be automated

**Research project**

1. Researchers
   - ![Researchers icon]

2. Research field
   - ![Light bulb icon]
   - Scientific research

3. Datasets
   - ![Metadata icon]

**Annodata: Researcher and projects**
- Contact information
- Researcher type: internal / external
- Accreditation of research institute
- ...

**Annodata: Access regime**
- Access modes
- Access protocols
  - “Sign contract A”

**Annodata: Access regime**
- Access modes
- Access protocols
  - “Sign contract B”

**Dataset level**

**Global level**

**Annodata: Data linkage**
- Record linkage
- Combining restrictions
- Decision rules
Working Group on “Modes of Data Provision”

The working group, which was chaired by Banque de France, built on the annodata schema and conducted a survey on the very granular level of access modes.

### Distribution of Access Modes

- **(Secure) on-site access**: 9 institutions, 13 access modes
- **Remote execution**: 3 institutions, 4 access modes
- **Remote access**: 3 institutions, 3 access modes
- **Scientific Use Files**: 3 institutions, 3 access modes

* 10 INEXDA members participated with 1-n possible access modes each. Public Use Files were excluded from this figure.

### Data source by topic

- **Economy and finance**: 23 access modes
- **Industry, trade and services**: 7 access modes
- **General and regional statistics**: 6 access modes
- **Population and social conditions**: 4 access modes
- **International trade in goods**: 4 access modes
- **Agriculture, forestry and fisheries**: 2 access modes
- **Environment and energy**: 1 access mode
- **Transport**: 1 access mode
- **Science, technology and digital...**: 1 access mode

* 10 INEXDA members participated with 1-n possible access modes each.
INEXDA’s work

Conclusion and future plans

In the past, INEXDA worked on data access procedures (e.g. annodata schema, survey).

For the future, INEXDA plans to

- launch new working groups on anonymization and output control.
- use the outcomes of the working groups to develop common best practices for data sharing and the handling of granular data. This would be extremely helpful for any institution, which plans to establish e.g. a new access mode and is looking for reference.

These considerations obviously need to acknowledge the limits set by extant legal, technical or procedural frameworks in which INEXDA members operate.
Thank you for your attention!

Contact at INEXDA:
Chair  Stefan Bender, Deutsche Bundesbank
Secretary  Inexda@Bundesbank.de

Website:
inexda.org/
How to become an INEXDA member

Other central banks, national statistical institutes and international organizations are encouraged to join INEXDA.

The following procedure has been established to admit new members:

<table>
<thead>
<tr>
<th>Step</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>A representative of the applying institution has to attend at least one INEXDA meeting in person before the formal application process is initiated.</td>
</tr>
<tr>
<td>2</td>
<td>Send an application letter signed by the head of the statistical department of the respective institution (or, in case of national statistical institutes, by the head of the responsible statistical department) to the chair of INEXDA.</td>
</tr>
<tr>
<td>3</td>
<td>All INEXDA members must agree to any application by a new institution.</td>
</tr>
<tr>
<td>4</td>
<td>Sign the MoU*. Congratulations, you are now a member of INEXDA.</td>
</tr>
</tbody>
</table>

* The INEXDA Memorandum of Understanding (MoU) can be accessed here.
Sharing researcher-generated code and value-added documentation in a trusted research environment

Louise Corti and Hannah Hodge-Waller,
Office for National Statistics, UK

---

1 This presentation was prepared for the conference. The views expressed are those of the authors and do not necessarily reflect the views of the BIS, the IFC or the central banks and other institutions represented at the event.
Sharing researcher-generated code and value-added documentation in a Trusted Research Environment

Prepared by Louise Corti and Hannah Hodge-Waller (Office for National Statistics) 1

Abstract

This paper sets out work that the ONS Secure Research Service is progressing to enable sharing of documentation and code created by researchers accessing controlled data. We have seen an increasing demand from researchers, funders and some data owners to enable wider access to this value-added material. This is less about reproducibility concerns, and more about recognising and building upon the significant amount of labour that goes into preparing research-ready data, new derived variables, and measures and histories. The paper sets out some of the key issues we have been addressing with researchers and data owners and sets out options for sharing of code in our Trusted Research Environment.

Keywords: data sharing, code sharing, reproducibility, trusted research environment, INEXDA, data access

Content

1. Introduction .......................................................................................................................................... 2

2. About the ONS Secure Research Service .................................................................................. 2

3. Why share code? ................................................................................................................................... 6

4. Good practice around coding across UK government ........................................................................... 7

5. Use cases for SRS code sharing pilot .......................................................................................... 8

5.1 ADR funded projects and Fellowships .................................................................................. 9

5.2 Longitudinal Educational Outcomes (LEO) data preparation code .......................................... 10

5.3 Covid infection study (CIS) code repository ............................................................................. 10

5.4 ‘Discuss and Collaborate’ space for the Integrated Data Programme TRE ........................................ 11

6. Workflow and protocols for code sharing in the SRS ...................................................................... 11

7. SRS user engagements on code sharing.......................................................................................... 14

8. Conclusion and future plans ............................................................................................................. 17

References ................................................................................................................................................ 17

1 Louise is Head of Analytical Insights and Impact, and Hannah the Statistical Code Sharing Manager for the Integrated Data Service Strategy Division, at the Office for National Statistics. The authors would like to thank colleagues at the Secure Research Service at ONS who are contributing to our SRS code sharing pilot work: Ian Banda, Beth Hopkins, Lorraine Ireland, Lauren Kerry, Alice McTiernan, Beth Routley and Sabrina Corps.
1. Introduction

This paper provides an overview on pilot work to progress code sharing within the Office for National Statistics’ (ONS) Secure Research Service (SRS), a Trusted Research Environment (TRE). The work builds on early ideas discussed at a workshop on researcher code sharing challenges, hosted by the Office for National Statistics and the UK Data Service in February 2022. At the meeting representatives of INEXDA joined colleagues from UK data access organisations and researchers, to consider the pros and cons of reproducible working in Trusted Research Environments (TRE) or RDC (Corti and Engeli, 2020). This paper is intended to share plans and experiences gathered so far regarding sharing of code based on research use of granular data.

It is now generally appreciated that that in addition to transparency around reported analysis, and the increasing use of reproducible code and processes, there is also value of making better use of the research work that goes on behind the scenes when getting data into shape for analysis. Recognising the significant legal and technical challenges involved in providing access to granular data, any wider sharing of code outside the closed researcher environment requires safeguarding statistical confidentiality, validation of its quality, and disclaimers to manage data owner and publishers’ reputation.

Working to identify strong use cases, we have begun building polices, protocols, templates, and guidance for writing, submitting, quality assuring and publishing ‘code’ created by researchers. In the first phase, we have been focussing on ‘value-added code’ rather than purely analytical code; so, the preparatory work undertaken by researchers to prepare data, such as making their data ‘research-ready’ for modelling, creating new variables, histories and so on. There is no reason why ‘analytic code’ cannot be included in a TRE code repository, but typically this is made available outside of the TRE, for example, when a package of code is requested by a journal to confirm that published results can be reproduced on the same data.

The paper introduces the SRS, discusses code sharing benefits and efforts around reproducibility within UK government. It goes on to introduce our use cases for the code sharing pilot in the SRS and the protocols being developed. It finishes with feedback from researchers around sharing code and plans for capacity building in this space.

A final note is that the work is still very much a work in progress. Having recently been successful in recruiting a dedicated Statistical Code Sharing Manager role, we feel that the detailed work can now accelerate, and we can look to share completed protocols and case studies over the next few months.

2. About the ONS Secure Research Service

The Office for National Statistics (ONS) Secure Research Service (SRS) is a Trusted Research Environment (TRE). It gives accredited or approved researchers secure
access to a wealth of de-identified, unpublished data (microdata) to work on research projects for the public good. The SRS is accredited as a Processor by the UK Statistics Authority (UKSA) under the UK legislation, the Digital Economy Act (DEA) 2017 for the provision of data for research purposes. (Office for National Statistics, 2022)

To ensure safe use of these data sources, the Secure Research Service (SRS) makes use of the Five Safes Framework; the set of principles adopted by secure labs, which researchers and their organisations must adhere to when undertaking research. The Five Safes protocols provide complete assurance for data owners:

- **Safe People**: trained and accredited researchers trusted to use data appropriately
- **Safe Projects**: data that are only used for valuable, ethical research that delivers clear public benefits
- **Safe Settings**: settings in which access to data is only possible using our secure technology systems
- **Safe Data**: data that have been de-identified
- **Safe Outputs**: all research outputs that are checked to ensure they cannot identify data subjects

Researchers must become an Accredited Researcher and submit a project proposal for it be Accredited, via an online submission system.

Figure 1 sets out the typical journey for a researcher using the service. Platform access is typically via remote access to the Windows-based cloud environment platform, where a unique project space is set up for Accredited Researchers to work on their Project. Data sets that have been approved for the work are mapped to the project space and requests can be made to add additional data. A range of software is available with the most popular being Stata and R.

The SRS researcher journey

![The researcher journey](https://www.ons.gov.uk/aboutus/whatwedo/statistics/requestingstatistics/approvedresearcherscheme#accessing-the-secure-research-service-srs)

1. **Safe People**: trained and accredited researchers trusted to use data appropriately
2. **Safe Projects**: data that are only used for valuable, ethical research that delivers clear public benefits
3. **Safe Settings**: settings in which access to data is only possible using our secure technology systems
4. **Safe Data**: data that have been de-identified
5. **Safe Outputs**: all research outputs that are checked to ensure they cannot identify data subjects

Researchers must become an Accredited Researcher and submit a project proposal for it be Accredited, via an online submission system.

Figure 1 sets out the typical journey for a researcher using the service. Platform access is typically via remote access to the Windows-based cloud environment platform, where a unique project space is set up for Accredited Researchers to work on their Project. Data sets that have been approved for the work are mapped to the project space and requests can be made to add additional data. A range of software is available with the most popular being Stata and R.

The SRS researcher journey

![The researcher journey](https://www.ons.gov.uk/aboutus/whatwedo/statistics/requestingstatistics/approvedresearcherscheme#accessing-the-secure-research-service-srs)

Figure 1

**Office for National Statistics**

1. **Safe People**: trained and accredited researchers trusted to use data appropriately
2. **Safe Projects**: data that are only used for valuable, ethical research that delivers clear public benefits
3. **Safe Settings**: settings in which access to data is only possible using our secure technology systems
4. **Safe Data**: data that have been de-identified
5. **Safe Outputs**: all research outputs that are checked to ensure they cannot identify data subjects

Researchers must become an Accredited Researcher and submit a project proposal for it be Accredited, via an online submission system.
Most datasets are available to access through remote access to the SRS, though in very few instances, data can only be accessed from an approved safe setting. The online metadata catalogue lists available datasets and any associated access restrictions. There are around 125 datasets listed that can be requested for use in Accredited projects (Figure 2). These span a number of themes, with business data being amongst the most requested (Figure 3).

In August 2022 the SRS catalogue plans to roll out Digital Object Identifiers (DOI) 2220 for the SRS metadata catalogue, to uniquely identify and cite the datasets. This is the first DOI pilot within UK government digital publishing, and the use case will feed into wider guidelines for implementation in data catalogues and other government published material with research value.

Figure 4 gives a summary profile of the volume of activity in the SRS over the past year. Overall, the service is supporting around 600 live projects, although research projects are not actively worked on all the time. The most popular research themes for this last 12 months have been education and ‘health, primarily due to some important linked data assets around the theme of education being ingested, and data to support analysis during the pandemic.

Top data sources being used requested

May 2022

<table>
<thead>
<tr>
<th>Dataset</th>
<th>Full Dataset Name</th>
</tr>
</thead>
<tbody>
<tr>
<td>NPD Bespoke Extracts</td>
<td>NPD bespoke data extracts</td>
</tr>
<tr>
<td>BSD</td>
<td>Business Structure Database - UK</td>
</tr>
<tr>
<td>ABS</td>
<td>Annual Business Survey - UK</td>
</tr>
<tr>
<td>ASHE</td>
<td>Annual Survey of Hours and Earnings - UK</td>
</tr>
<tr>
<td>LS</td>
<td>Longitudinal Study of England and Wales</td>
</tr>
<tr>
<td>APS (Population)</td>
<td>Annual Population Survey - UK</td>
</tr>
<tr>
<td>LFS person</td>
<td>Labour Force Survey Person - UK</td>
</tr>
<tr>
<td>LFS Longitudinal</td>
<td>Labour Force Survey Longitudinal - UK</td>
</tr>
<tr>
<td>LFS Household</td>
<td>Labour Force Survey Household - UK</td>
</tr>
<tr>
<td>BERD - GB</td>
<td>Business Enterprise Research and Development - England, Wales and Scotland</td>
</tr>
<tr>
<td>ARDx</td>
<td>Annual Respondents Database x - UK</td>
</tr>
<tr>
<td>UKIS</td>
<td>UK Innovation Survey</td>
</tr>
<tr>
<td>ARD2</td>
<td>Annual Respondents Database 2 - UK</td>
</tr>
<tr>
<td>BRES</td>
<td>Business Register Employment Survey - UK</td>
</tr>
<tr>
<td>Mortality</td>
<td>Death Registration Finalised Extracts - England and Wales</td>
</tr>
<tr>
<td>BICS</td>
<td>Business Insights and Conditions Survey - UK</td>
</tr>
</tbody>
</table>

* NPD is National Pupil Database (Education administrative data)

Source: SRS Data Owner Dashboard

Research Activity in the ONS SRS

Cumulative number of accredited projects and researchers in the SRS, since July 2021

Metrics are shown for the past 12 months, to provide an example of volume of activity.

Source: SRS Data Owner Dashboard
The Integrated Data Service (IDS) is a new and upcoming UK cross-government service that aims to build on the success of the Secure Research Service (SRS). The IDS has the Office for National Statistics as the lead delivery partner and will enable co-ordinated secure access to a range of high-quality data, critical to informing policy decisions and improving public services.

3. Why share code?

The drive towards transparency and accountability in research has seen community and government-led policies being developed for opening access to research resources. An emerging 'reproducibility crisis' in some disciplines demands that more data and code should be released for substantiating published results. Indeed, code sharing is increasingly being viewed as best practice in empirical scientific research.

As a relatively high bar, research reproducibility for a published research article is where the authors provide all the data, code, and processing instructions necessary to rerun exactly the same analysis and obtain identical results. Contributing one’s own code voluntarily is also valuable for the future progress of science. It demonstrates a willingness to follow open science principles, promotes a positive and collaborative approach and ensures that researchers are able to return to their own work later and remember what they have done.

Despite some disciplines leading the way with making underlying code available, sharing code is still not widely done across all in quantitative social and economic science. Economists, psychologists and political scientists have pushed forward with reproducibility mandates, keen to demonstrate research accountability. For example, the American Economic Association journals employ data editors to rerun code and validate results. As a tightknit community, demographers have also moved as a discipline towards R for analysis, and code sharing for complex routines is quite prevalent. Other disciplines, such as sociology or gerontology, appear more wary of sharing code/syntax, worrying more about not having been taught to write ‘good code’, and exposing this aspect of their work. This is both a cultural and a capability issue but is likely declining due to new research cohorts trending towards using open-source software and better code management practices.

Our interactions with researchers have shown that many feel anxious about exposing code they have written, considering that they may not have the skills to write ‘good code’. Others do it routinely, submitting analytic code with their publications and publishing it openly in GitHub.

Gold standard code can be submitted to an external service for validation and receipt of a certificate of reproducibility, for example, the cascad or CODECHECK.

---

5 The cascad service is a certification agency for scientific code and data. The certification allows researchers to signal the reproducibility nature of their research and to their peers. https://www.cascad.tech/
6 The CODECHECK service supports code checkers with a workflow, guidelines and tools to evaluate computer programs underlying scientific papers. The independent time-stamped runs conducted by code checkers awards a “certificate of executable computation” and increase availability, discovery and reproducibility of crucial artefacts for computational science. https://codecheck.org.uk/
services. And code can be published in an open data repository such as or Zenodo\textsuperscript{7}, aiding citation and visibility (for example, with a DOI). Intellectual property for code can stay with the researcher but working collaboratively with data owners likely merits joint ownership.

As users of data, researchers can build upon existing code to avoid recreating basic recoding routines or derivation of complex histories. Access to syntax for derived variables created by data owners also supports the derivation of new variables for analysis.

There are challenges of reproducing work that's undertaken in a trusted research environment (TRE). Data access is restricted and unpublished material that has not been disclosure checked and approved cannot be taken out of the TRE. Any code shared outside the secure environment must first be reviewed for disclosure risk. Code tracking and versioning tools can also be used inside a TRE such as R Markdown, Jupyter Notebook and GitLab \textsuperscript{8} to manage and document code.

The current ‘As is’ process for code sharing in the SRS does not include the use of a dedicated code repository but relies on a bespoke request from a researcher to be allowed to use code from another of their own or colleague’s projects, using the existing statistical disclosure control (SDC) output’ clearance mechanism. This process does not currently check code for good practice, and it almost impossible for researchers to know what useful code might be available, thereby preventing opportunities to experiment with useful data preparation work and derivations. The draft ‘As is’ and proposed workflows are set out in Section 6.

4. Good practice around coding across UK government

In its National Data Strategy, ONS signals its commitment to good practice and a commitment to transparency. Transparency, in its broadest sense, can be represented in many ways, but the ability to track back from ‘fork to field’ is important when it comes to analysis and modelling based on public data assets.

The UK response to the pandemic powerfully illustrated the benefits of responsible and effective use and sharing of data to understand COVID –19, to support people, and cooperate across borders. Stemming from this we are seeing a clearer understanding around good practice and how we enable our commitment to transparency. Making code available that underpins both data preparation and analysis promotes openness and provides value-added resources on which new analysis can build. Indeed, the IDS service signals its commitment to code sharing by providing improved toolsets and track-back functionality.

The Government Digital Service (GDS) has been a leader in promoting reproducible ways of working, setting high standards and providing capacity building resources for the government analysis community. Help on learning to code at GDS

\textsuperscript{7} Zenodo is an open science repository for European scientists to submit research materials. https://zenodo.org/

\textsuperscript{8} R Markdown (https://rmarkdown.rstudio.com/, Jupyter Notebook (https://jupyter.org/) and GitLab (https://about.gitlab.com/)

Sharing researcher-generated code and value-added documentation in a TRE
is included in their training programme.\textsuperscript{9} The cross-government Civil Service network, the Government Analysis Function, also helps support good practice, and covers around 17,000 people involved in the generation and dissemination of analysis. It provides a solid learning curriculum covering methods, analysis and reproducible coding.\textsuperscript{10}

At the ONS, the data engineering teams have been building capacity and capability by developing skills and reuse around processes and code, especially for statistical production. Reproducible Analytical Pipelines (RAP) experts work alongside with business areas to create sustainable data pipelines.\textsuperscript{11} The ONS Data Science Campus routinely use data science methods and provide guidance on creating and sharing high quality sustainable code and pipelines. ONS also offers a host of great learning resources and opportunities, some of which are open to all. The ONS Data Access Platform Capability and Training Support (DAP CATS) team create and promote useful learning materials and coding resources, including a number of excellent Jupyter Executable Books\textsuperscript{12} (Turrell, 2022; UK Government Analytical Community, 2020; Data Access Platform capability team, 2020).

The IDS is building on these forward-looking high standards that will meet reproducibility needs and avoid reinvention of wheels along the statistical production journey. Data can be easily updated through systematic engineering, and blocks of code for derived variables/measures and analytic outputs can be constructed and rerun by analysts. Not only is the IDP aiming for maintainable outcomes that can be updated and future proofed, shared methods libraries and code repositories will also be available.

5. Use cases for SRS code sharing pilot

While discussions in the SRS around enabling code sharing have been underway for some time, a number of timely and varied use cases have presented themselves; which we have adopted as exploratory pilots.


\textsuperscript{11} Infrastructure for Reproducible Analytical Pipeline (RAP) by the Government statistical Service (GSS): https://gss.civilservice.gov.uk/reproducible-analytical-pipelines/infrastructure-for-rap/

\textsuperscript{12} Jupyter Book (https://jupyterbook.org/intro.html) is an open-source project for building publication-quality books and documents from computational material. It builds on the Jupyter Notebook web-based interactive development environment for notebooks, code, and data. See the ‘Executable Books’ website: https://executablebooks.org/en/latest/gallery.html
5.1 ADR funded projects and Fellowships

The first opportunity concerns a group of funded Research Fellows working on SRS projects, supported by the UK’s Administrative Data Research (UDRUK) Fellowship programme. ADR UK have been supportive of enabling researcher code sharing in the SRS and worked with ONS to (i) add a clause around delivery of code into the Fellow’s contracts and (ii) fund a dedicated post within the ONS to help progress the work.13 The Contract Terms and Conditions state that: “Enhanced or derived data, code, products or tools for reuse created during this grant will be deposited in the ONS Secure Research Service as set out in ‘Intellectual Property’ below”. This requirement certainly sets a bar.

In January 2022, the ONS impact team organised and hosted a workshop to discuss putting this code sharing clause into practice with the Fellows. Following a short introduction on good practice, the workshop used breakout groups to discuss challenges and opportunities. Attendees contributed their experiences and expectations when working with code – either supporting delivery of services or as part of research and intelligence activities, within a trusted research environment (TRE) or as an active contributor to public platforms such as GitHub/Lab.

The exercise gave a useful view on ‘readiness’ or sharing code and to plan where to start, as well as identifying what level of guidance would be needed to help support capacity building in this area.

Following from that one of the ADR funded projects, the Wealth and Employment Dynamics in Britain (WED) project, agreed to prepare code for wider sharing and work with us on a template and good practice guide for documenting Stata code.14 The WED project is expected to increase understanding of how people’s wages progress through their career, factoring in key demographic characteristics, as well as the particular dynamics of low-pay labour markets. The ONS helped support the linkage of three key data sources on: employee earnings, a snapshot in time of all firms in the UK registered for tax, the 2011 Census data (for England and Wales), and tax data relating to employment spells and earnings, as well as income from occupation pensions and benefits.

This project has the potential to transform understanding of wage and employment issues in Britain, from labour market entry, through job mobility and career progression to retirement decisions. The project has worked to create a new linked dataset and to fully document the Stata code detailing the data manipulation operations to create this new data. The code is currently being reviewed, using our emerging standards. In Section 6 we set out areas for review of code, including QA such as understandability and comprehensiveness of annotation, and sdc.

---


14 The ‘Wage and employment dynamics in Britain’ project is an ADR UK-funded data linkage project aiming to provide important new insights into the dynamics of earnings and employment in Britain. https://www.adruk.org/our-work/browse-all-projects/wage-and-employment-dynamics-in-britain-143/
5.2 Longitudinal Educational Outcomes (LEO) data preparation code

LEO is a de-identified, person level administrative dataset that brings together data on individual’s education, employment, earnings data and benefits claims. The asset links data provided by five separate government departments. The dataset is widely recognised as having the potential to provide transformative insight and evidence on the longer-term labour market outcomes and educational pathways of around 38 million English learners, supporting government decision making in order to improve services. The first iteration of this unique linked data source was acquired and delivered by ONS through the Secure Research Service.

There are currently around 50 researchers using, or applying to use, the data in the SRS across nine separate projects. Given the very large size and complexity of the data, researchers have to request (and justify the need for) a number of variables. The permitted variables are accessed via bespoke SQL views. Some projects actively worked with primary data owner, the Department for Education, in the early stages to produce derived variables and so on. Some of the code has been shared across ONS-led projects, and these are a target for wider sharing.

The code is primarily written in R and covers data manipulation activity, rather than pure analysis. One of the ONS projects was the first to externally publish policy relevant findings based on the data (Tolland, Tierney and Bathgate, 2021).

5.3 Covid infection study (CIS) code repository

Following the onset of the Covid-19 outbreak, the UK government agreed it was crucial to understand how Covid-19 was spreading across the population in order to control the pandemic and its effects. To assist the government’s response, ONS, in partnership with the University of Oxford, the University of Manchester, and Public Health England, set up the Coronavirus Infection Survey (CIS). Launched in April 2020 the study tested people for the virus whether they have symptoms or not, with data routinely collected from each individual in a household, via nose and throat swabs (measure of infection rate) and blood samples measure of antibodies). Socio-demographic and employment characteristics were also collected, along with symptoms experienced, self-isolating or shielding, and exposure to a suspected carrier of Covid-19.

The CIS was made available for analysis in the SRS. A large Accredited Project with around 80 researchers on it, began to analyse data in the SRS. Given the limitations of the secure platform, and the significant modelling ask, some analyses, using R, moved to the ONS Google Cloud Platform (GCP) where less granular CIS data were available.

GCP already uses GitHub to share code, using source repositories and Jupyter notebooks, and set up around ten separate repositories for each key analytical

---

15 Accessing the Longitudinal Education Outcomes (LEO) dataset.
https://www.gov.uk/guidance/apply-to-access-the-longitudinal-education-outcomes-leo-dataset

https://www.ons.gov.uk/peoplepopulationandcommunity/healthandsocialcare/conditionsanddiseases/bulletins/coronaviruscovid19infectionsurveypilot/previousReleases
pipeline. Work in the SRS is more challenging to version control without access to an external location for remote repos, and associated tools for merging branches. In lieu, CISA built version-controlled code using a master folder as a remote location and cloned repos from there. Most pipelines did not use GIT but had only basic change control indicated (a combination of comments within the code and a commented changelog header at the top of the file).

This use case lends itself well to the pilot, as the multiple strands of code, and pipelines could be reviewed, combined and documented into key reference and actionable code for future use. This is very much a typical scenario, and, in hindsight, it is easy to question why analytic code wasn’t meticulously managed or curated. However, the urgency of producing analysis for senior officials and policy makers at the time took precedence over code management. The experience of working across sectors also unearthed different approaches to documenting code.

The use case also lend itself well to a ‘lessons learned’ exercise around having gold standard protocols for the future. While retrospective tidy up can be a lot of work, there is an opportunity to take a review. This example also provides a good example to see how the data producer feels about analysts’ code; for example, would they wish to add disclaimers on any researcher-generated code, so as to distinguish added-value work from a formally published dataset?

5.4 ‘Discuss and Collaborate’ space for the Integrated Data Programme TRE

The IDS has aspirations to build on the great work of the ONS SRS. The platform aims to facilitate cross-government working and discussions are underway to plan a collaboration space, where Accredited Researchers could formulate joint ideas for analysis and share methods, analysis and code outside of their own project spaces. The pilot work being undertaken in our SRS pilots is helping to discover both service, researcher, data owner and future user needs.

6. Workflow and protocols for code sharing in the SRS

The following diagrams show the process maps for the current and draft proposed journeys for code sharing in the TRE. Figure 5 shows the As-is process.
Wider code sharing models beyond 1-1 transfer between research projects has been scoped out. The SRS Code Repository or Code library must meet the skills of two sets of researchers, demarked by their existing code management practices.

These are:

- **Group 1.** Those used to managing and versioning code using a Git type environment; primarily R and Python users;
- **Group 2:** Those used to storing and managing code or syntax files using traditional folder structures, possibly (and hopefully) with file naming conventions (primarily Stata and SPSS users).

The plan is for the SRS Code Repository to be an SRS fully managed repository that is accessible to all those entering the environment. Thus, code should be findable, searchable and well documented, and have a workflow for submission, review and publishing. All code that is offered for deposit must be sdc reviewed and quality assessed. Figure 6 sets out the possible planned workflow.
In addition to assessing the most suitable structure for either Git or folder structure libraries, Figure 7 sets out the mandatory documentation and user files that will be needed.
### Code submission documentation

<table>
<thead>
<tr>
<th>Stage</th>
<th>SRS Code Sharing Policy</th>
<th>Internal policies</th>
</tr>
</thead>
<tbody>
<tr>
<td>Administration/Internal</td>
<td>Data owner agreement Include data owner options for review or disclaimer. Option to add to Data Sharing Agreement</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Admin Check list Check list to track governance, submission and code publishing steps</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Code Repository conventions Set out repository conventions and is an Annex to the Code Deposit Guide, Includes: file naming and labelling, User Guide, Disclaimers, Citation statement</td>
<td></td>
</tr>
<tr>
<td>Before Project Starts</td>
<td>Code Templates Sets out recommended styles, headers, types of annotation. For different software languages</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Code Deposit Guide Guide to Good Coding and requirements (SRS oriented)</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Training Documents/Sessions Optional; but intended to help assist new researchers or those new to code sharing with promoting best practices</td>
<td></td>
</tr>
<tr>
<td>During Project</td>
<td>Code Advice/Drop in Sessions Focussed on best practice. Embedding sharing and reproducibility</td>
<td></td>
</tr>
<tr>
<td>Code submission point</td>
<td>SRS Code Sharing Policy External policy</td>
<td></td>
</tr>
<tr>
<td></td>
<td>SRS review criteria Criteria used to review code that is submitted. Covers sdc and QA and required documentations e.g., ReadMe</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Code clearance request form Simple form for submitting code with basic details</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Code clearance confirmation Email to confirm clearance</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Code Update Status Agreed with researcher, how often is code reviewed; process map if changes are needed to code</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Read Me file, per deposit Describe the code file(s) contents and any key user information. Includes the citation for the material</td>
<td></td>
</tr>
</tbody>
</table>

Source: IDS Analytical Insights Team, Office for National Statistics

### 7. SRS user engagements on code sharing

Over the past year, the Insights team have been seeking the views of SRS users on code sharing, in tandem with the pilot work to progress the sharing of researcher-generated code and additional user-focussed documentation. Overall, the engagement exercises showed that that both analysts and data owners see the value of sharing code.

Every year the SRS runs an Annual User Experience Survey in early spring. The survey asks researchers about their satisfaction with various aspects of the service, about data needs and on suggested improvements.
A question on code sharing was specifically included this year. Figure 8 suggests that code sharing is indeed happening, via various means inside and outside of the TRE including via the SRS internal libraries and externally via other organisations and websites. More government departments and private sector researchers had shared code with colleagues compared to within academia.

**Code sharing through the SRS**

Shared or published code undertaken in the ONS SRS* Figure 8

![Bar chart showing code sharing stats](image)

* Annual User Experience Survey (Secure Research Service). Delivered online February-March 2022. N = 57

Source: ONS Analysis Insights, Integrated Data Service, Office for National Statistics

When asked more generally, in an open question, about how ONS might better support analytical needs, Figure 9 shows code sharing as the second most volunteered suggestion. Comments volunteered by survey respondents around code sharing ability and software packages included:

- Improve functions when adding to the environment
- Enable shared working areas
- Ability to share code
- Greater flexibility in the R libraries installed
- Enable Reproducible Analytical Pipelines in SRS, for instance RMarkdown
- Promote standard cleaning code across the government

Users were also invited to note whether using the SRS had enhanced their capability. Of the 95 responding, 40 mentioned coding skills. Of interest, 65 mentioned sdc awareness, and 45 noted data management.
How might ONS better support analytical needs?

Open responses coded by theme to the question, “How might ONS better support analytical needs?” in the Office for National Statistics (SRS)*

Following the survey, in a similar vein, in Spring 2022, the IDS Analytical Insights team undertook a series of focus groups with SRS users from government departments. While the main focus was on seeking a better understanding of the types of analysis being undertaken in the environment, the kinds of public research outputs being created and to hear about actual or expected outcomes, we used the opportunity to ask about software needs and code sharing. Specifically, questions asked about analytic and software needs were:

- What software are you using to perform the majority of your analyses at present?
- Does it meet your needs?
- Does your team you have any skills gaps for the analyses?
- What tools do you use to manage your code within a secure environment? e.g., for collaboration or QA.
Comments volunteered on code sharing included:

“It would be great if there was a standard cleaning code across the government”

“Would be good to get better resources to enable Reproducible Analytical Pipelines in SRS, for instance RMarkdown.”

We found out that many researchers did not know that there was a Git GUI already available to them for managing and versioning code for their projects. It is part of a standard set of software that users can access but has never been actively promoted.

8. Conclusion and future plans

This paper has set out an update on code sharing pilot work being undertaken for the ONS Secure Research Service. The timing is excellent, given that the challenges about how to go about demonstrating transparency and reduce repetition and redundancy in creating and running analysis pipelines come to the fore. Further, the topic is being raised in many forums, by researchers, funders and sponsors and engineers. The SRS Code Sharing group is continuing to connect in with other areas with ONS and promote its work, and will start to publish case studies, protocols, guidance and templates once they are ready.

Going forward, we have two strands of engagement activities being planned. The first is a call through our researcher networks for early adopters, so that we get a pipeline of keen researchers who want to have their code assessed and shared. The second is a series of Drop-In sessions and webinars on Writing and Documenting Good Code. There is still the need to highlight benefits and openly examine barriers and show how they can be overcome. Sharing tips and templates for how newcomers can get started is a good first step.

References


Sharing researcher-generated code and value-added documentation in a Trusted Research Environment

August 2022

Louise Corti, Head Analytical Insights and Impact, ONS
louise.corti@ons.gov.uk
Overview

• Introduce researcher code sharing in the UK Secure Research Service (SRS)
• Highlight proposed processes, policies, documentation and support & training activities
• Highlight use cases

Who has shared code with analysts other than with close colleagues?
Why share code?

✓ Viewed as best practice for demonstrating transparency and accountability in empirical scientific research
✓ Enables building upon existing code to support the derivation of new variables, avoid recreating complex recoding routines
✓ Exposes code for peer review /validation/ promotion
✓ Some journals require underlying code submission

❖ Users can feel anxious about exposing their code; consider they don’t have skills or time to write ‘good code’
❖ Users ask if data owners can supply code for derived variables
Options for code sharing

PROJECT: Peer Reviewer added to project to QA-reviewed code, feedback provided

PROJECT: Contribute QA-reviewed code to project area

INTERNAL: Contribute sdc & QA-reviewed code to global SRS folder/Git

OPEN: Submit/publish sdc & QA-reviewed code to journal

OPEN: Publish sdc & QA-reviewed code on a public website/GitHub
Aim of pilot work

• Facilitate planning - code sharing group set up and manager role recruited
• Locate suitable use cases and start investigations/solutions

Explore and develop workable processes and protocols:

➢ 'As is' and proposed workflows
➢ Governance and administration, resourcing
➢ SRS and user policies
➢ User guidance and templates
➢ Capability building activity: webinars, 1-1 drop in sessions, blogs and case studies

• Roll out early adopter call and training sessions
Use case 1: Wealth and Employment Dynamics in Britain (WED) project

- Project aims to transform understanding of wage and employment issues in Britain, from labour market entry, through job mobility and career progression to retirement decisions.
- Involves linkage of key data sources: Annual Survey of Hours and Earnings (ASHE), 2011 Census data for England and Wales, tax data, and income from occupation pensions and benefits.
- Project keen to be early SRS adopters for documenting and sharing Stata code:
  - Data manipulation operations to create new data
  - Testing our emerging standards on good code and review processes.

Data Creation Code Description

This file lists in detail the code files developed by the WED team to generate the ASHE datasets and supplementary files, and describes their functions, input and outputs.

The spreadsheet gives a simpler list. The presented file shows diagrammatically how the inputs and outputs of the programs link, and which programs call other programs.

<table>
<thead>
<tr>
<th>Global</th>
<th>ASHE</th>
<th>AUX auxiliary files</th>
<th>01_create_BSD_lookups</th>
<th>02_create_postcode_EN_lookup</th>
<th>03_create_nmw_lookup</th>
<th>04_create_survey_reference_dates</th>
<th>05_create_rural_urban_coa_lookup</th>
<th>06_create_CPI_index_lookup</th>
<th>DATA Code creating user files</th>
</tr>
</thead>
<tbody>
<tr>
<td>3</td>
<td>2</td>
<td>4</td>
<td>4</td>
<td>5</td>
<td>6</td>
<td>7</td>
<td>8</td>
<td>9</td>
<td>11</td>
</tr>
</tbody>
</table>

`$3_create_nmw_lookup`

Brief Description

Creating table of annual NMW rates with matching bands. Two files created:

1. `$nmw_group_file` for an age, year and quarter this gives you the nmw band (note apprentice pay eligibility needs to be calculated on separate information- only available in ASHE from 2013).
2. `$nmw_rate_file` for an nmw band and year the exact rate in pennies is given.

To use these files:

- Merge on the $nmw_group_file by age, year and quarter to get the $nmw_band.
- Adjust for apprentices if necessary.
- Merge on the $nmw_rate_file by year, quarter and band to get the exact rate in pennies for an individual.

Detailed Description

Stage 1: Import nmw data from Excel spreadsheet for ages between 16 and 120, from years 1999 to the latest year.

For each year, quarter and age, create a variable `$nmw_band` which says which nmw band a person should fall into (note that age bands vary over time and not all bands exist for whole period).

There are five `$nmw` bands numbered 1 to 5, and labelled as follows:

1. `$nmw_apprentice`
2. `$nmw_teen`
3. `$nmw_development`
4. `$nmw_adult`
5. `$nmw_nlw`

The labels in the global variable are the same without ‘nmw ’
Use case 2: Longitudinal Educational Outcomes dataset

- LEO is a de-identified, person level administrative dataset that brings together data on individual’s education, employment, earnings data and benefits claims.
- Asset links data provided by five separate government departments via the SRS.
- Dataset has the potential to provide transformative insight and evidence on the longer-term labour market outcomes/educational pathways of @38 million English learners.
- 9 projects/50 researchers using data, including government users.
- Early data manipulation work to create 'research-ready 'datasets/new variables.
- Some R code shared across ONS-led projects; target wider sharing in the SRS.
Use case 3: Large scale Covid survey analysis

- April 2020 new survey launched from ONS, Universities, Public Health England: the Coronavirus Infection Survey (CIS), available in the SRS
- Interviews with each individual in a household, including nose and throat swabs (infection rate) and blood samples (antibodies)
- Large project with 80 researchers with urgency and significant modelling asks. Directly informed government decision-making
- Varied software use: R users preferred ONS Google Cloud Platform with less granular data
  - GCP uses GitHub to share code - ten repositories set up for each key analytical pipeline
  - SRS - initial poor management of code; later built basic version-controlled code using a master folder
- Review the code repositories for lessons learned for large multi-sector projects
- Work with data owner to review publishing of analysts’ code in the SRS; distinguish added-value work from formal data documentation
Useful ONS resources

- Quality Assurance of Code for Analysis and Research
- Reproducible Analytical Pipelines (RAP) Champions
- Data Accelerator programme
- Reproducible Analysis — Coding for Economists ([aeturrell.github.io](https://aeturrell.github.io))
- Tips for Better Coding — Coding for Economists ([aeturrell.github.io](https://aeturrell.github.io))
Joint secondary anonymisation of categorial and numerical variables in sensitive time series microdata – novel approach for Statistical Disclosure Control of a sensitive microdata set published in BELab data laboratory¹

Eugenia Koblents and Alberto Lorenzo Megía,
Bank of Spain

¹ This presentation was prepared for the conference. The views expressed are those of the authors and do not necessarily reflect the views of the BIS, the IFC or the central banks and other institutions represented at the event.
Joint secondary anonymisation of categorical and numerical variables in sensitive time series microdata

A novel approach for Statistical Disclosure Control of a microdata set published in BELab data laboratory

Eugenia Koblents Lapteva and Alberto Lorenzo Megía

Abstract

In this paper, a Statistical Disclosure Control (SDC) approach for secondary anonymisation of sensitive time series microdata is proposed that allows the joint analysis of numerical and categorical key variables. This method has been developed at the Banco de España’s BELab data laboratory in order to protect confidentiality of the recently published CIR dataset. This dataset contains yearly microdata on loans to legal entities, including multiple variables describing the loans and debtors. The main challenge faced in this work is the fact that the set of key variables, i.e. those that may allow debtor re-identification, includes both categorical and numerical loan and debtor variables. Additionally, debtors may have multiple loans and loans may have multiple debtors, which makes the direct use of existing SDC software tools for microdata protection (mu-argus, sdcMicro) unfeasible. For these reasons, a novel SDC procedure has been designed and implemented in order to protect the debtors appearing in the CIR dataset against re-identification, while jointly analysing categorical and numerical variables and addressing time series data protection.

Keywords: Statistical Disclosure Control, numerical and categorical key variables, time series data protection.

JEL classification: C4 (Econometric and Statistical Methods: Special Topics)
# Contents

1. Introduction ............................................................................................................................. 3  
2. Secondary anonymisation of CIR dataset ................................................................................. 4   
   2.1. Key variable identification ............................................................................................... 6  
   2.2. Global recoding of categorical key variables ................................................................... 6  
   2.3. Creation of debtor profiles ............................................................................................... 7  
   2.4. Local suppressions performed on debtor profiles ............................................................ 8  
   2.5. Local suppressions performed on the original loans dataset ........................................ 9  
3. Conclusions and future lines of research ............................................................................... 11  
References .................................................................................................................................. 12
1. Introduction

This paper addresses Statistical Disclosure Control (SDC) of sensitive microdata in the context of the Banco de España’s BELab data laboratory [5]. The goal of SDC is to minimize the risk of re-identification of individual samples while minimising the information loss produced by the anonymisation in order to retain information utility [1]. The SDC process requires the identification of feasible disclosure scenarios and key variables that might allow the re-identification of individual samples in the microdata set under analysis.

This work entailed analysis of a dataset containing information on loans (in the following referred to as the CIR dataset), which has recently been published in a BELab safe data room [6]. The CIR (Central de Información de Riesgos) dataset contains yearly microdata on loans extended to legal entities, resident and non-resident in Spain, that were reported to the Central Credit Register (CCR) between 2016 and 2020. This dataset includes multiple variables describing loans and debtors, and does not contain information on the financial institutions involved. This dataset required both primary and secondary anonymisation to avoid debtor re-identification, and has some distinctive features that make the direct use of existing SDC software tools for microdata protection (mu-argus, sdcMicro) unfeasible. Additionally, existing SDC software tools have some other limitations that have also been addressed in this work.

As a result, a novel SDC procedure has been designed and implemented in order to protect debtors appearing in the CIR dataset against re-identification. The implemented SDC procedure makes use of the open source R package sdcMicro [2,3]. Both sdcMicro and mu-argus [1,4] are Eurostat-supported SDC software tools used by many public institutions, such as national statistical institutes and central banks. Both implement a broad variety of SDC methods for individual risk evaluation, microdata protection and information loss assessment. In this work, the sdcMicro package has been used because its creators claim that it is better optimised for large volumes of data [2,3] as compared with mu-argus. Pre- and post-processing of the CIR dataset has been implemented in Python. The proposed secondary anonymisation procedure has been designed in close collaboration with the Banco de España’s CIR Department and has been validated by a team of internal researchers who consistently work with the CIR dataset, to guarantee that the utility of the anonymised dataset is preserved.

Numerous previous works have addressed the independent protection of categorical and numerical key variables [1,8,10,11,12]. In particular, in [1] the authors provide an extensive overview of the main SDC methods for addressing microdata protection problems involving both categorical and numerical key variables. However, both types of variables are generally processed independently. In [8] the authors focus on the protection of outlying samples of numerical variables, and compare, for different masking methods, the information loss and disclosure risk related to outliers. A recent review of the state of the art of SDC and a discussion on future challenges (big data, machine learning, etc.) can be found in [11].

Previous attempts have also been made to jointly analyse categorical and numerical variables. In particular, in [7] the authors propose to exploit the hierarchy of categorical variables to compute a numerical mapping that quantifies their underlying semantics. This approach is similar to the one proposed in this work, in the sense that it aims to combine categorical and numerical variables by computing distances between categories, while we propose to transform numerical variables into...
categorical ones. More recently, in [9] the authors discuss the limitations of existing SDC software and the need to jointly assess disclosure risk and information loss when both categorical and numerical key variables are identified. However, to the best of our knowledge, a general solution to this problem has not yet been found.

The rest of the paper is organised as follows. In Section 2 we describe the secondary anonymisation method designed and implemented in this work, including a description of the CIR dataset and the steps of the anonymisation procedure. The numerical results obtained are also presented. Finally, Section 3 is devoted to the conclusions and future lines of research.

2. Secondary anonymisation of CIR dataset

The goal of SDC is to protect statistical data by producing safe datasets with low (individual) risks and high data utility, and that can be securely released without compromising data confidentiality. The SDC procedure for protecting sensitive microdata consists of the following steps [2]:

1. Deletion of direct identifiers, to guarantee primary confidentiality.
2. Identification of key variables, to address secondary confidentiality.
3. Measurement of individual risks based on sample frequency counts.
4. Application of SDC methods to protect high-risk observations.
5. Assessment of the resulting disclosure risk and the information loss produced by the anonymisation procedure.

Software tools such as sdcMicro and mu-argus provide implementations for a broad variety of SDC methods that can be used in steps 3-5. However, the CIR dataset has a number of peculiarities that hinder the direct application of standard microdata protection methods. Therefore, a specific procedure has been designed and implemented to address this problem.

The CIR dataset currently available at BELab contains data describing loans extended to legal entities between 2016 and 2020. Around 25 million records are available, containing 19 variables describing debtors and loans. A complete sample is available representing the whole population. The CIR dataset contains the following variables describing debtors:

1. Debtor ID (anonymised)
2. Residence
3. Institutional sector
4. Economic activity
5. Enterprise size
6. Legal form

On the other hand, the variables describing loans are the following:

7. Loan ID (anonymised)
8. Type of instrument
9. Residual maturity
10. Currency
11. Collateral type
12. Collateral coverage
13. Personal guarantee
14. Personal guarantee coverage
15. Investment region
16. Joint debtor
17. Number of joint debtors
18. Drawn amount
19. Undrawn amount

Variables 18 and 19 (drawn and undrawn amounts) are numerical variables while the rest of them are categorical. Debtor ID and loan ID have been anonymised by the
data provider prior to being shared with BElab to guarantee primary confidentiality. Loans are only active for a specific period of time. A detailed description of the CIR dataset is available in the user manual published by BElab [6].

Even though the original dataset contains information on loans, the sensitive entity to be protected is the debtor, which makes the direct use of existing SDC software tools unfeasible. The standard SDC procedure for microdata protection makes the assumption that each row in the dataset represents an individual respondent, which is not satisfied in this case. The CIR dataset includes debtors with multiple loans and loans with multiple debtors (joint loans). Additionally, several variables describing loans, such as investment region, currency and amounts, can also allow debtor re-identification, and must therefore also be considered as key variables.

Existing SDC software tools have other limitations which are relevant in this case. On the one hand, they do not allow a joint analysis of categorical and numerical variables and do not support time series data protection. On the other hand, the implemented anonymisation methods for numerical variables do not yield a good trade-off between re-identification risk and information loss in this case. Ideally, we would like to protect only those samples that turn out to be sensitive when numerical and categorical variables are jointly analysed and leave the rest of the samples unaffected. The computational cost should also be affordable even for large datasets.

The top/bottom coding method is very simple and fast and only affects a small number of samples (information loss is limited). However, this method does not allow the protection of samples that might turn out to be sensitive when numerical and categorical variables are jointly analysed. Thus, disclosure risk might not be sufficiently reduced under this approach. This method processes each numerical variable independently, ignoring correlations among variables, and requires the definition of individual thresholds for each variable that directly affect the resulting data utility and disclosure risk.

Alternatively, more complex methods based on micro-aggregation are widely used and recommended in the literature, since they allow a significant reduction in disclosure risk. However, their main limitation is the fact that all samples are affected by the anonymisation process, significantly reducing data utility in some scenarios. This family of methods also requires high computation times, which can hinder its application when working with large volumes of data.

Finally, perturbative methods, such as noise addition and rank swapping, have been discarded in this work due to the feedback received from the team of internal researchers, who claimed that data utility would be seriously affected by these transformations.

For all these reasons, a specific approach has been designed and implemented at BElab to address secondary anonymisation of the CIR dataset. The proposed method consists of encoding loan and debtor information into a so-called debtor profile, to ensure that the resulting dataset contains one single row per individual respondent and that standard SDC methods and tools can be used. Numerical key variables are discretised and incorporated into this profile, allowing the assessment of individual re-identification risks based on complete debtor and loan information. The proposed procedure consists of the following steps:

1. Identification of debtor and loan key variables.
2. Global recoding of selected key variables, reducing the number of classes.
3. Creation of a full profile for each debtor, including information on all of its loans (active at some point throughout the full time series).
4. Local suppressions performed on debtor profiles to ensure k-anonymity with the selected value of k.
5. Transfer of local suppression patterns identified for each debtor to the original loans dataset.

When new yearly data is incorporated into the dataset, the full process needs to be repeated (including all yearly data available to date) and a new anonymised time series dataset needs to be generated. Under this approach, each researcher has access to no more than one version of the time series dataset simultaneously. Otherwise, an intruder would be able to cancel the performed suppressions by comparing different datasets, since the suppression pattern would be different in both versions.

### 2.1. Key variable identification

Key variables are those that may lead to the disclosure of individual samples in feasible re-identification scenarios. Key variable selection is a challenging problem that requires close collaboration between the SDC expert and the data provider. In this case, even though the original dataset mainly contains information on loans, the entity to be protected is the debtor. For this reason, all variables describing debtors, except for the anonymised debtor ID, have been considered as key variables.

Additionally, selected variables describing loans have also been considered as key variables, since they can allow debtor re-identification in the defined disclosure scenarios. Anonymised debtor and loan IDs have not been taken into account for the anonymisation procedure, since they are internal identifiers that are not published externally and thus cannot be used for re-identification by potential intruders. Table 1 shows all the debtor and loan key variables identified in this case study.

<table>
<thead>
<tr>
<th>Debtor key variables</th>
<th>Loan key variables</th>
</tr>
</thead>
<tbody>
<tr>
<td>Residence</td>
<td>Currency</td>
</tr>
<tr>
<td>Institutional sector</td>
<td>Personal guarantee</td>
</tr>
<tr>
<td>Economic activity</td>
<td>Investment region</td>
</tr>
<tr>
<td>Enterprise size</td>
<td>Drawn amount</td>
</tr>
<tr>
<td>Legal form</td>
<td>Undrawn amount</td>
</tr>
</tbody>
</table>

### 2.2. Global recoding of categorical key variables

Once the set of key variables has been identified, global recoding is commonly performed on selected categorical variables by grouping existing classes [1]. This method significantly reduces disclosure risk while incurring an acceptable information loss, since the recoded classes and the grouped categories are selected based on expert knowledge in order to maximise data utility. Global recoding does not involve sample suppression but reduces the level of detail of all samples. In this work, this process has been agreed upon with the data provider and a number of internal researchers to guarantee high information utility for the resulting data. The process has been implemented in Python instead of using sdcMicro, because a large number of categories had to be grouped in this case. Table 2 shows the original and modified number of categories for the recoded variables. The reduction in disclosure risk...
achieved through this procedure cannot be numerically assessed, since at this point the data still does not contain one single row per individual respondent and is thus unsuitable for risk disclosure assessment using sdcMicro. However, global recoding significantly reduces disclosure risk when a large number of categories are grouped.

<table>
<thead>
<tr>
<th>Categorical variables</th>
<th>Original categories</th>
<th>Modified categories</th>
</tr>
</thead>
<tbody>
<tr>
<td>Institutional sector (debtor)</td>
<td>16</td>
<td>3</td>
</tr>
<tr>
<td>Economic activity (debtor)</td>
<td>167</td>
<td>21</td>
</tr>
<tr>
<td>Currency (loan)</td>
<td>56</td>
<td>4</td>
</tr>
<tr>
<td>Personal guarantee (loan)</td>
<td>5</td>
<td>4</td>
</tr>
<tr>
<td>Investment region (loan)</td>
<td>55</td>
<td>18</td>
</tr>
</tbody>
</table>

### 2.3. Creation of debtor profiles

The key step in the designed anonymisation procedure is the creation of a detailed profile for each debtor containing information on all of its loans throughout the time series. In the original CIR dataset each row contains loan and debtor information, while the entity to be protected in this case is the debtor. Loans can affect multiple debtors and debtors can be involved in multiple loans. For this reason, the original dataset does not contain a single row per individual respondent, which makes the direct use of standard SDC procedures unfeasible. To overcome these difficulties, the original dataset has been transformed in such a way that each row, called the debtor profile, contains all the information of a given debtor. All debtor and loan key variables, both categorical and numerical, need to be represented in the debtor profile in order to assess individual risks based on complete debtor information. To achieve this goal, categorical loan key variables have been encoded using one-hot encoding, while numerical loan key variables have previously been discretised.

All debtor key variables (Table 1, left-hand column) have been directly included in the profile. These variables correspond to fixed debtor attributes that are usually constant over time (company size, economic activity, etc.). However, in some cases these attributes might also change during the time series. In those cases, the most frequent value has been considered. Where there are multiple modes, the most recent value has been selected.

Additionally, categorical and discretised numerical variables describing loan operations (Table 1, right-hand column) have also been incorporated into the debtor profile using one-hot encoding. For categorical loan variables, an auxiliary binary variable has been created for each category of the original key variables. In particular, 18 auxiliary binary columns have been created for the “Investment region” variable, 4 columns for the “Currency” variable and 4 columns for the “Personal guarantee” variable. The presence and absence of loan operations with specific “Currency”, “Guarantee” and “Investment region” attributes has been encoded as 1 and 0 respectively in the debtor profile. For example, if a debtor has invested in Andalusia and Madrid, a 1 will appear on those two columns and the rest of the values for “Investment region” will be 0. Table 3 shows several examples of the resulting codification of categorical loan key variables.
On the other hand, the two numerical key variables describing loans (drawn and undrawn amount) have also been incorporated into the debtor profiles in the following way. First, the maximum of the two amounts for each operation has been computed. Then, this maximum value has been discretised according to the number of digits, thus taking only a small number of possible values (loans with less than 6, 7, 8, 9, 10 and 11 digits or more). This discretisation process assumes that an intruder might know the order of magnitude of a debtor’s operations but not the exact amounts. This assumption has been considered reasonable by the data provider.

A new binary variable has then been added to the debtor profile for each of these new categories. The existence of operations with a given number of digits for a given debtor has been encoded as 1 in the debtor profile matrix. For example, if a debtor has operations with 8 and 9 digits, there will be 1s in those columns in its profile and 0s in the rest of the columns representing amounts. Table 4 shows synthetic examples of the section of the debtor profile corresponding to the discretised amount variables.

As a result of this process, a full profile for each debtor has been created, including information on all its loans, which can then be processed using standard SDC software tools such as mu-argus or sdcMicro. In this particular case, a total of 1,430,503 debtors with active operations in any of the years across the time series has been obtained. A profile containing 37 variables has been created for each debtor, containing 5 debtor variables and 32 one-hot encoded loan key variables.

### 2.4. Local suppressions performed on debtor profiles

Once a detailed profile has been created for each debtor including the relevant debtor and loan information, the sdcMicro tool has been used to evaluate disclosure risk and anonymise the debtor information [2,3]. In particular, local suppressions to achieve k-anonymity with k=3 have been performed. A dataset satisfies k-anonymity when there are at least k samples with the same combination of key variables. sdcMicro allows the local suppression of specific values in the dataset in order to guarantee that k-anonymity is satisfied for all samples. Perturbative SDC methods for
protecting categorical key variables have been discarded in this work because they introduce randomness that can significantly reduce the value of the resulting data for researchers.

The disclosure risk assessment conducted in sdcMicro revealed that 68,450 debtor profiles (4.78%) do not satisfy k-anonymity with k=3. Individual risk for each debtor is computed as the inverse of its number of replicas in the dataset. Samples with an individual risk above a threshold of 1/k = 0.33 thus need to be protected. After performing 78,394 suppressions (corresponding to 0.15% of the debtor information), k-anonymity with k=3 is satisfied for all samples. The execution time for this process was 13 hours. Table 5 shows the results provided by the sdcMicro tool.

**Table 5. Number and percentage of samples violating k-anonymity in the original and anonymised debtor profiles, computed by sdcMicro.**

<table>
<thead>
<tr>
<th>k-anonymity</th>
<th>Original debtor profiles</th>
<th>Anonymised debtor profiles</th>
</tr>
</thead>
<tbody>
<tr>
<td>2-anonymity</td>
<td>46,936 (3.28%)</td>
<td>0 (0%)</td>
</tr>
<tr>
<td>3-anonymity</td>
<td>68,450 (4.78%)</td>
<td>0 (0%)</td>
</tr>
<tr>
<td>5-anonymity</td>
<td>95,733 (6.69%)</td>
<td>14,353 (1%)</td>
</tr>
</tbody>
</table>

2.5. Local suppressions performed on the original loans dataset

Finally, the local suppression pattern obtained for each debtor has been transferred to the original CIR dataset for each of the operations. For example, if a local suppression of the variable “Investment region=Madrid” has been performed for a given debtor, only those operations of that debtor with the attribute “Investment region=Madrid” will be affected by local suppressions, but not operations with any other investment region. On the other hand, if a discretised 10-digit amount has been suppressed for a certain debtor, only those operations with 10-digit amounts will be affected. This process produced a total of 4,300,076 local suppressions in the whole time series, which corresponds to 0.95% of the suppressed values, over a total of 1,430,503 rows and 37 columns. Table 6 shows a summary of the results.

**Table 6. Summary of the number of local suppressions per variable in the full time series.**

<table>
<thead>
<tr>
<th>Debtor and loan key variables</th>
<th>Number of suppressions</th>
<th>Percentage of suppressions</th>
</tr>
</thead>
<tbody>
<tr>
<td>Residence</td>
<td>70,720</td>
<td>0.27</td>
</tr>
<tr>
<td>Institutional sector</td>
<td>81,436</td>
<td>0.31</td>
</tr>
<tr>
<td>Legal form</td>
<td>528,404</td>
<td>1.98</td>
</tr>
<tr>
<td>Economic activity</td>
<td>3,006,248</td>
<td>11.29</td>
</tr>
<tr>
<td>Enterprise size</td>
<td>447,127</td>
<td>1.68</td>
</tr>
<tr>
<td>Currency</td>
<td>44,428</td>
<td>0.17</td>
</tr>
<tr>
<td>Guarantee</td>
<td>15,623</td>
<td>0.06</td>
</tr>
<tr>
<td>Drawn amount</td>
<td>2,993</td>
<td>0.01</td>
</tr>
<tr>
<td>Undrawn amount</td>
<td>2,993</td>
<td>0.01</td>
</tr>
<tr>
<td>Investment region</td>
<td>100,244</td>
<td>0.38</td>
</tr>
<tr>
<td>TOTAL</td>
<td><strong>4,300,076</strong></td>
<td><strong>0.95</strong></td>
</tr>
</tbody>
</table>

As can be seen in Table 6, the variables with the highest number of suppressions are the debtor key variables “Economic activity”, “Legal form” and “Enterprise size”, with a 11.29%, 1.98% and 1.68% of suppressions, respectively. Only 0.01% of numerical values of drawn and undrawn amounts have been suppressed.
Fig. 1 (left-hand chart) depicts the joint density estimation based on a 1% sample of the original data and the marginal densities for both numerical variables in logarithmic units. The plot shows that the joint density has two main modes. The marginal density of the variable “Undrawn amount” is highly concentrated at low values (the median is actually 0). Suppressed numerical values with added noise are also represented. It can be observed that multiple outlying samples have been suppressed. However, operations with lower amounts that turn out to be sensitive when analysed in combination with other key variables (e.g. the largest loans per sector or per region) have also been protected. By contrast, other outlying samples have been protected by suppressing some of the categorical key variables instead of the numerical ones. This suppression pattern has been obtained using sdcMicro, which identifies the optimal solution to achieve k-anonymity with the suppression of a minimum number of sample values. Obtaining this suppression pattern, which allows the protection of sensitive numerical samples within their whole range of values, is not possible without performing the described analysis.

For comparative purposes, Fig. 1 (right-hand chart) shows the suppression pattern obtained with top/bottom coding (with added noise), together with the threshold used for each variable. sdcMicro does not allow the joint protection of multiple numerical variables with top/bottom coding and each of them is protected independently. The suppression threshold for each variable has been adjusted so that the total number of suppressions is very similar for both methods (2,963 samples for top/bottom coding versus 2,993 samples for the proposed method). It can be observed that only samples with the highest values of one or both of the variables have been protected by top/bottom coding. Some of those samples are not really sensitive, as they satisfy k-anonymity and have thus been left unaffected by the proposed method. However, other samples that may indeed be sensitive when jointly analysing multiple key variables, have not been protected and disclosure risk may still be higher than desired.

Even though the number of suppressions is very similar in both cases, statistical distributions and summary statistics are affected in different ways, since top/bottom coding consistently suppresses the highest values in the population while the proposed method suppresses values in the whole range of values of each variable. To illustrate these differences, Fig. 2 shows the violin plot obtained from the original data, data anonymised using the proposed method and top/bottom coding, for the variables “Drawn amount” (left-hand chart) and “Undrawn amount” (right-hand chart) in logarithmic units. The plots show that only the tail of the distribution of both
variables is affected by anonymisation, while the rest of the distribution is preserved. It can also be seen that the proposed method affects the tails less than top/bottom coding.

![Fig. 2. Violin-plot of the “Drawn amount” (left) and “Undrawn amount” (right) variables. Original and anonymised data using the proposed method and top/bottom coding are shown.](image)

Table 7 shows a comparison of the main summary statistics obtained using top/bottom coding and the proposed method. The deviation with respect to non-anonymised data has been computed for each variable. It can be observed that the maximum value of the data anonymised with top/bottom coding is reduced by 99% on average with respect to the maximum value of the original dataset, while only a 68% reduction is obtained with the proposed method. Similarly, the mean and standard deviation is far more affected when using top/bottom coding than the proposed method. The median is very slightly modified by both anonymisation procedures. These values illustrate how disclosure risk and information loss are significantly reduced when applying a suppression pattern that jointly analyses categorical and numerical variables.

<table>
<thead>
<tr>
<th>Summary statistics</th>
<th>Drawn amount</th>
<th>Undrawn amount</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Top/bottom coding</td>
<td>Proposed method</td>
</tr>
<tr>
<td>Maximum</td>
<td>-99.6%</td>
<td>-83.7%</td>
</tr>
<tr>
<td>Mean</td>
<td>-24.0%</td>
<td>-12.9%</td>
</tr>
<tr>
<td>Standard deviation</td>
<td>-88.7%</td>
<td>-68.3%</td>
</tr>
<tr>
<td>Median</td>
<td>-0.03%</td>
<td>-0.04%</td>
</tr>
</tbody>
</table>

3. Conclusions and future lines of research

In this paper, a strategy for performing secondary anonymisation of microdata combining numerical and categorical variables is proposed. The described method has been evaluated using a real and especially sensitive time series dataset, recently published by the BELab data laboratory. This dataset contains information on loans, while the sensitive entity to be protected is the debtor. Additionally, the set of key variables include both categorical and numerical debtor and loan variables. These particularities make the use of standard SDC methods unsuitable for this problem.

The key step in the proposed method is the creation of a debtor profile that incorporates complete information on all loans active at some point during the full
time series. The profiles include debtor key variables as well as one-hot encoded loan key variables. Once complete debtor profiles are created, standard SDC software tools such as scdMicro or mu-argus can be used to assess disclosure risk, protect sensitive samples and evaluate information loss. In this work, local suppressions to ensure k-anonymity with k=3 are performed to protect debtor information. The resulting suppression patterns are finally transferred to the original loan dataset.

The proposed method has a number of advantages over alternative procedures. It allows the joint protection of categorical and numerical variables, and generates suppression patterns that protect sensitive samples only and leave the rest unaffected. Disclosure risk is thus significantly reduced with very limited information loss. Additionally, statistical distributions and summary statistics are less affected in comparison with alternative methods such as top/bottom coding. This procedure makes use of existing SDC software tools, is simple to implement and cost efficient, which makes it suitable for large datasets. It allows the protection of time series data and only requires the selection of the k-anonymity k parameter.

Future lines of research include modelling the uncertainty regarding the information available to intruders, and exploring the relationship between SDC and anomaly detection to optimise results in both fields.

References

2. Templ, M., Meindl, B., and Kowarik, A. (2013). Introduction to statistical disclosure control (SDC). Project: Relative to the testing of SDC algorithms and provision of practical SDC, data analysis OG.
JOINT SECONDARY ANONYMISATION OF CATEGORICAL AND NUMERICAL VARIABLES IN SENSITIVE TIME SERIES MICRODATA

A novel approach for Statistical Disclosure Control of a microdata set published in BElab data laboratory

Eugenia Koblents
Alberto Lorenzo

ELEVENTH IFC CONFERENCE ON “POST-PANDEMIC LANDSCAPE FOR CENTRAL BANK STATISTICS”
BIS, BASEL, 25 AND 26 AUGUST 2022
INEXDA SESSION: MICRODATA DISCLOSURE CONTROL: A PRACTICAL PERSPECTIVE

25/08/2022

STATISTICS DEPARTMENT
INTRODUCTION
Motivation and goal

- Banco de España launched BELab in July 2019 to provide access to the research community to high quality microdata via on-site and remote access. [https://www.bde.es/bde/en/areas/analisis-economia/otros/que-es-belab/](https://www.bde.es/bde/en/areas/analisis-economia/otros/que-es-belab/)

- In October 2021 CIR (Central de Información de Riesgos) Department of Banco de España provided a very sensitive dataset to BELab containing information on loans to legal entities extended between 2016 and 2020.

- Primary and secondary anonymisation was required to protect debtor confidentiality.

- Categorical and numerical debtor and loan key variables have been identified.

- CIR dataset contains multiple rows per debtor and loan (joint loans).

- A novel SDC approach for secondary anonymisation has been designed and implemented which allows to jointly analyse categorical and numerical key variables.

- The implemented approach makes use of the open-source R package sdcMicro for risk assessment and microdata protection and Python for data pre and post processing.
Challenges faced

- **Existing SDC software tools** (sdcMicro, mu-argus) have some limitations which are relevant in this case:
  - They require data to contain **one single row per individual respondent**, which often is not the case.
  - They do not support a **joint analysis of categorical and numerical variables**.
  - Implemented anonymisation methods for **numerical key variables** (top/bottom coding, micro aggregation, noise addition, etc.) do not yield a good trade-off between disclosure risk and information loss for this problem.
  - They do not support **time-series data protection**.

- **A novel secondary anonymisation approach** has been designed and implemented which overcomes these difficulties:
  1. Identification of continuous and numerical debtor and loan **key variables** that can allow debtor re-id.
  2. **Global recoding** of selected key variables reducing the number of classes and disclosure risk.
  3. Creation of **full debtors’ profiles** that incorporate information on all their loans throughout the full time series.
  4. **Debtor anonymisation**: local suppressions performed on debtor profiles with sdcMicro to guarantee k-anonymity.
  5. **Transfer of local suppression** patterns of debtors to the original loans dataset.

- When **new yearly data** is incorporated to the dataset the full process needs to be repeated, requiring that each researcher only has access to one version of the dataset to avoid the cancellation of local suppressions.
1. Identification of categorical and numerical debtor and loan key variables that can allow debtor re-id in feasible disclosure scenarios. This is a challenging problem that needs to be addressed in collaboration with the data provider:

<table>
<thead>
<tr>
<th>Debtor key variables</th>
<th>Loan key variables</th>
</tr>
</thead>
<tbody>
<tr>
<td>Residence</td>
<td>Currency</td>
</tr>
<tr>
<td>Institutional sector</td>
<td>Personal guarantee</td>
</tr>
<tr>
<td>Economic activity</td>
<td>Investment region</td>
</tr>
<tr>
<td>Enterprise size</td>
<td>Drawn amount</td>
</tr>
<tr>
<td>Legal form</td>
<td>Undrawn amount</td>
</tr>
</tbody>
</table>

2. Global recoding of selected categorical debtor and loan key variables by grouping existing classes to significantly reduce disclosure risk. This process is agreed with the data provider and data users to guarantee high data utility.

<table>
<thead>
<tr>
<th>Categorical variables</th>
<th>Original categories</th>
<th>Modified categories</th>
</tr>
</thead>
<tbody>
<tr>
<td>Institutional sector (debtor)</td>
<td>16</td>
<td>3</td>
</tr>
<tr>
<td>Economic activity (debtor)</td>
<td>167</td>
<td>21</td>
</tr>
<tr>
<td>Currency (loan)</td>
<td>56</td>
<td>4</td>
</tr>
<tr>
<td>Personal guarantee (loan)</td>
<td>5</td>
<td>4</td>
</tr>
<tr>
<td>Investment region (loan)</td>
<td>55</td>
<td>18</td>
</tr>
</tbody>
</table>
3. **Debtors’ profiles** are created, which contain information on all their **loans** extended in the whole time-series.

- **Categorical key variables describing loans** have been incorporated into the profile using **one-hot encoding** (an auxiliary binary variable has been created for each category of the original key variables).

<table>
<thead>
<tr>
<th>Debtor ID</th>
<th>EUR</th>
<th>USD</th>
<th>GBP</th>
<th>Other currencies</th>
<th>Madrid</th>
<th>Catalonia</th>
<th>Andalusia</th>
</tr>
</thead>
<tbody>
<tr>
<td>52364</td>
<td>1</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>1</td>
<td>1</td>
<td>0</td>
</tr>
<tr>
<td>76354</td>
<td>1</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>75345</td>
<td>1</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>1</td>
</tr>
</tbody>
</table>

- **Continuous key variables describing loans** are discretized according to the number of digits (loans with 1-6, 7, 8, 9, 10 and 11 digits) and are incorporated into the profile in the same way as categorical variables.

<table>
<thead>
<tr>
<th>Debtor ID</th>
<th>1-6 digits</th>
<th>7 digits</th>
<th>8 digits</th>
<th>9 digits</th>
<th>10 digits</th>
<th>11 digits or more</th>
</tr>
</thead>
<tbody>
<tr>
<td>52364</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>76354</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>0</td>
</tr>
<tr>
<td>75345</td>
<td>1</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
</tbody>
</table>

- As a result, **1,430,503 debtor profiles** (with one **single row per individual respondent**) have been created containing **37 variables**: 5 debtor key variables and 32 one-hot encoded loan key variables.
4. Debtor profiles contain one single row per individual and existing SDC software can be used to evaluate individual disclosure risks and to apply local suppressions to sensitive debtor profiles to guarantee k-anonymity (k=3).

Disclosure risk evaluation performed by sdcMicro before and after applying local suppressions to debtors’ profiles.

- 4.78% of debtors are at risk of re-id in the original dataset, when combining categorical and numerical key variables.
- 0.15% of debtors’ information has been suppressed (78,394 values of samples).
- As a result, all debtors in the anonymised dataset satisfy k-anonymity with k=3 (all individual risks are below 0.33).
5. Local suppressions obtained for debtors are **transferred** to the original loans dataset. **0.95%** of suppressions on average (mainly economical activity, legal form and company size). **0.01%** of numerical values suppressed.

<table>
<thead>
<tr>
<th>Debtor and loan key variables</th>
<th>Number of suppressions</th>
<th>Percentage of suppressions</th>
</tr>
</thead>
<tbody>
<tr>
<td>Legal form</td>
<td>528,404</td>
<td>1.98</td>
</tr>
<tr>
<td>Economic activity</td>
<td>3,006,248</td>
<td>11.29</td>
</tr>
<tr>
<td>Enterprise size</td>
<td>447,127</td>
<td>1.68</td>
</tr>
<tr>
<td>Drawn amount</td>
<td>2,993</td>
<td>0.01</td>
</tr>
<tr>
<td>Undrawn amount</td>
<td>2,993</td>
<td>0.01</td>
</tr>
<tr>
<td>TOTAL</td>
<td><strong>4,300,076</strong></td>
<td><strong>0.95</strong></td>
</tr>
</tbody>
</table>

- Suppression pattern (red dots) obtained with the **proposed method** (left) and **top/bottom coding** (right).
- The proposed method protects samples that turn out to be sensitive when categorical and numerical variables are **jointly analysed** (largest loan per sector or region, etc.), while top/bottom coding consistently suppresses **high-valued samples**.
The proposed method affects the data distribution and summary statistics less than top/bottom coding. The obtained suppression pattern yields a low information loss and disclosure risk, since only sensitive samples are modified.

- Only the tail of the distribution is affected by anonymisation. The proposed method affects the tails less than top/bottom coding, because less outliers are suppressed.

- Summary statistics are significantly less affected using the proposed method than top/bottom coding.

<table>
<thead>
<tr>
<th>Summary statistics</th>
<th>Drawn amount</th>
<th>Undrawn amount</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Top/bottom coding</td>
<td>Proposed method</td>
</tr>
<tr>
<td>Maximum</td>
<td>-99.6%</td>
<td>-83.7%</td>
</tr>
<tr>
<td>Mean</td>
<td>-24.0%</td>
<td>-12.9%</td>
</tr>
<tr>
<td>Standard deviation</td>
<td>-88.7%</td>
<td>-68.3%</td>
</tr>
<tr>
<td>Median</td>
<td>-0.03%</td>
<td>-0.04%</td>
</tr>
</tbody>
</table>
A novel secondary anonymisation approach has been designed and implemented to protect the CIR dataset as a result of a close collaboration between BELab and the data provider (CIR Department of Banco de España).

This procedure has a number of benefits over alternative procedures:

- It allows to jointly analyse and protect categorical and continuous variables. Information on loans is incorporated into the debtors data, yielding a very complete profile for each debtor and allowing to use existing SDC software.
- The joint anonymisation of categorical and numerical variables minimizes disclosure risk and information loss since only sensitive samples are affected. Only 0.95% of suppressions (0.01% of numerical values).
- The full time-series dataset is protected as a whole.
- Once the described procedure has been designed, its implementation and use is relatively simple and its computational cost is low, in comparison with alternative methods, such as micro aggregation.

Note that the full process needs to be repeated every time new yearly data becomes available. Researches cannot have access to more than one version of the dataset simultaneously, to avoid the possibility of cancelling local suppressions.

Future research lines:

- Address the possibility of modelling uncertainty on the information available to intruders.
- Analyse links between microdata protection and anomaly detection, since both topics are closely related.
Thank you for your attention!
Introduction to and application of SDC rules using self-developed tools\(^1\)

Jannick Blaschke, Matthias Gomolka, Christian Hirsch, Sebastian Seltmann and Harald Stahl, Deutsche Bundesbank

\(^1\) This presentation was prepared for the conference. The views expressed are those of the authors and do not necessarily reflect the views of the BIS, the IFC or the central banks and other institutions represented at the event.
Introduction to and application of SDC rules using self-developed tools

Jannick Blaschke, Matthias Gomolka, Christian Hirsch, Sebastian Seltmann and Harald Stahl

Abstract

To gain access to confidential microdata via a Research Data Centre (RDC), researchers and the output they produce must adhere to strict confidentiality rules. Typically, RDCs maintain explanatory documents and sophisticated tools to assist researchers in checking the compliance with those rules. However, – as a consequence of diverse and very complex requirements – the amount of available information on SDC rules is constantly increasing. We present concrete examples of three self-developed tools to help on-board researchers at the Deutsche Bundesbank without overloading them with information.

Keywords: Statistical Disclosure Control, output control, tools, INEXDA, data sharing, research data centre, data access, research, microdata

Contents

Introduction to and application of SDC rules using self-developed tools ........................ 1

1. A brief introduction to the work of Research Data Centres (RDCs) .................. 2

2. How tools support researchers in understanding and applying SDC rules .......... 3

3. Practical examples for self-developed tools supporting researchers at the RDSC of the Deutsche Bundesbank ................................................................. 6

   The Project start: Researchers would like to get an overview of all available information - The RDSC Landing Page................................................................. 6

   During the research: Researchers need context information that applies directly to their analysis - SDC tools................................................................. 7

   Finalising the project: Researchers need to ensure compliance with SDC rules - The Output Submitter ................................................................. 8

4. Discussion and future areas of improvement ................................................. 9

References ........................................................................................................... 10

Annex.................................................................................................................. 11

1 The authors would like to thank our colleagues at the RDSC for their valuable suggestions and feedback. All views expressed in this report are the authors’ personal views and do not necessarily reflect those of the Deutsche Bundesbank or the Eurosystem.
1. A brief introduction to the work of Research Data Centres (RDCs)

Much data collected for public purposes is considered to be a public good and should therefore be as broadly accessible as possible. While some of the data may be released to the general public without any restrictions, most data contain some sort of information that needs to be protected, e.g. personal or market-sensitive information.

One way to make this information accessible is to provide fully aggregated data that can be downloaded, e.g. from the website of the data-collecting institution. Fully aggregated data no longer allow direct or indirect identification of the reporting agents, e.g. households, banks or firms. Although fully aggregated data does ensure a very high level of data protection, much valuable information is lost in the process of aggregating the underlying micro data.

However, depending on the legal basis, anonymised micro data characterised by a lower degree of anonymisation than absolute anonymity\(^2\) may be used for independent scientific research. For this purpose, many data collecting institutions have established special access modes that ensure secure data usage by external researchers. Next to Scientific Use Files\(^3\), Research Data Centres (RDCs) are probably the most common of these access modes.

RDCs provide secure on-site access to confidential micro data for the purpose of scientific research (Ritchie, 2017, 2021). After their data access request has been approved by the RDC, external researchers can come to the premises of the RDC and access the micro data in a secure environment, where they have no access to the internet and are not allowed to use their personal laptops or phones. In addition, researchers must ensure that all results derived from confidential data that leave the secure environment of the RDC following a research visit no longer contain any confidential information. This so-called Statistical Disclosure Control (SDC)\(^4\) or output control is mandatory before results can be released.

The price that data providers and data-using researchers pay for this is a certain degree of complexity in accessing data. Starting with the application, to the access modalities, to arguably the most difficult part, the SDC, working in RDCs is often a new challenge for researchers. RDC staff are well aware that the access processes and rules can be complicated and often require RDC-specific knowledge.

This is particularly evident in the SDC. While this topic is essential for successful work in the RDC and good knowledge determines whether or not results can be published, researchers outside the RDCs usually do not encounter it at all. Therefore, new researchers first have to build up this knowledge, even though it is not part of

---

\(^2\) Such anonymised micro data is often referred to as formally or factually anonymised data. Formal anonymisation describes the deletion of direct identifiers such as names, addresses, and other identifiers (e.g. LEI) so that no direct identification is possible. Factual anonymisation includes additional perturbation measures so that an identification is only possible with an unreasonable investment of time, cost and manpower.

\(^3\) Scientific Use Files are usually factually anonymised datasets that are sent to approved researchers or can be downloaded from a password-protected area.

\(^4\) An example of an SDC document is the Bundesbank’s “Rules for visiting researchers at the RDSC” (Research Data and Service Centre, 2021).
their actual research. After building up the required knowledge, researchers also need to correctly apply this knowledge to their specific use cases. RDCs usually provide user documentation with information aiming to support researchers with both tasks irrespective of their level of knowledge. This strategy works when the number of accessible datasets is small and the complexity of these datasets is low.

In recent years, however, this strategy of a few documents that fit all users and all datasets has come under increased scrutiny. One reason for this is an increase in the number of available datasets as well as an increase in the linking possibilities of datasets in RDCs. In addition, datasets are also becoming increasingly more complex and originate from diverse backgrounds. This development makes it more cumbersome for researchers to comprehend and adhere to SDC requirements. It may also increase the need to amend existing requirements to reflect the new datasets.

These developments also have repercussions for how RDCs present information to their users. In this paper, we argue that ever-expanding user documentation is certainly not the most efficient and helpful way to address the growing complexity and diversity of datasets. Instead, more unstructured information by e.g. extending the SDC document increases the likelihood of researchers not finding the information they need when they need it. This in turn increases the likelihood that they will apply rules incorrectly, resulting in the submission of results that are non-compliant with SDC rules and thus rejected by RDC staff.

The objective of this paper is to present three tools that implement different approaches to address these developments. All three tools have been self-developed by the Research Data and Service Centre (RDSC) of the Deutsche Bundesbank. The first tool, the RDSC Landing Page, pre-sorts all available information according to characteristics deemed helpful to the user. The second tool, “nobsdes5/nobsreg5” and (sdcLog), helps users apply information by (semi) automating the SDC checking process. Finally, the Output Submitter applies selected SDC rules automatically to the researcher’s output, displays additional information if needed, and streamlines the submission process.

This paper is organised as follows. Section 2 discusses approaches to communicating complex information to a heterogeneous target group in the context of SDC. Section 3 presents the three self-developed tools from the RDSC that are meant to support researchers in learning and performing SDC. Finally, Section 4 concludes with a brief discussion and gives an outlook regarding future areas of improvement.

2. How tools support researchers in understanding and applying SDC rules

Researchers in RDCs generally need information about existing SDC rules and how to apply them to their research project. Let us take as an example a researcher who calculates the mean of a distribution and wants to check whether her result complies with SDC rules. Pointing this researcher to a potentially large pile of information would not be an effective way to support her in this situation. Instead, she would need supplementary information that helps her to isolate specific information on the calculation of the mean of a distribution. After that, she would have to filter this
information again depending on her personal skills, e.g. previous experience with SDC rules.

A natural first step to start is to pre-sort all available information into buckets that help researchers find the required documents faster. For example, one could choose categories according to researcher characteristics, the tasks that they perform, and the purpose of the document. Pre-sorting information according to researcher characteristics allows account to be taken of differences in knowledge owing to different levels of experience. At the one end of the spectrum are vastly experienced researchers who require only selective information on how to solve specific problems. First-time users, by contrast, need documents that teach them how to get started. We choose tasks as an additional sorting category because different tasks (potentially) require the application of a different set of rules.

While the first two dimensions are relatively straightforward, the third is often overlooked. Frameworks such as Diátaxis⁵ argue that that researchers’ success depends both on the theoretical understanding and their ability to apply this understanding to the task. Consequently, documents that cater to these different needs will also look vastly different. Documents aiming to enhance the theoretical understanding of researchers need to be descriptive, state the complete set of rules and explain the rationale behind them (e.g. “explanation” and “reference” documents). In contrast, documents aiming at facilitating the application of rules to a real-world task need to demonstrate how to solve a specific problem and guide researchers through a series of steps (e.g. “tutorials” and “how-to guides”).

---

**Pre-sorting heterogeneous information**

Structuring different documents* using two dimensions

---

* The different shapes symbolise the different document types, e.g. legal reference, tutorial, explanation.

Sources: Authors

Panel A and B of Figure 1 depict the transition from unsorted to pre-sorted information, i.e. from unstructured to structured information. For example, the x-axis

---

⁵ “The Diátaxis framework (…) adopts a systematic approach to understanding the needs of documentation users in their cycle of interaction with a product. Diátaxis identifies four modes of documentation - tutorials, how-to guides, technical reference and explanation. It derives its structure from the relationship between them.” For more information, see https://diataxis.fr/.
of Panel B could represent different groups of researchers with different experience levels while the y-axis could represent different tasks such as e.g. mean vs. regression. We represent documents with different purposes as differently shaped elements in the grid. Pre-sorting clearly speeds up information discovery compared to the base case (see Panel A), as researchers do not have to browse through all available documents. As such, pre-sorting creates metadata about available information. Pre-sorting is the idea behind the tool "The RDSC Landing Page" that we describe in more detail in the next section.

How useful this pre-sorting is hinges on (i) whether researchers correctly interpret the sorting categories and (ii) the scope and amount of different types of documents in each bucket. Let us begin with the former. This question relates to how researchers assess their own level of experience. For example, does a vastly experienced researcher refer to years working with data or years working with data in the RDC environment? Similarly, the task may not always be clear.

Furthermore, document characteristics may also affect the usefulness of pre-sorting for researchers. For an example, we shall turn to the bucket (A, x) in Panel B which features two distinctly different documents. It may take additional time for a researcher to go through both documents. Along the same lines, very long documents also require researchers to spend additional time finding information appropriate to their task.

To address this challenge, RDCs could either enhance their personal user support, e.g. by assigning staff to assist researchers in finding relevant information, or try to automate this process as far as possible. The advantage of automation is that it is resource-efficient and ensures an equal treatment for all researchers. For example, a tool could present or even apply SDC rules relevant to the current task, e.g. the SDC after the calculation of the mean of a distribution as in the example above.

Panel C of Figure 1 illustrates this point. The hatched area illustrates the tool which relieves the researcher of the decision between the two documents with different purposes that are available for the task and the level of the researcher’s experience. Without tools, on the other hand, the researcher would need to go through both documents and then apply the information to the task. Therefore, tools are an avenue to efficient information provision and application. The RDSC OutputSubmitter is an example of how to implement this idea. This tool automatically applies selected SDC rules to the researcher’s output and displays additional information if the output does not comply with applicable SDC rules.

In the next section, we present three examples of how self-developed tools support researchers in performing all tasks related to SDC checking. These examples are taken from the Research Data and Service Centre (RDSC), which is the RDC of the Deutsche Bundesbank. They loosely follow the process of a research project at the RDSC from start to finish. Readers should observe that tasks and therefore information needs of researchers might differ between the various RDCs. This could be due to differences in the underlying legal frameworks, the technical environment in which the data access is granted, or organisational decisions of the RDC.

---

6 Readers interested in a more detailed description of the RDSC are cordially referred to “Data Access to Micro Data of the Deutsche Bundesbank” (Schönberg (2019)).
In addition, the information need of a researcher might also vary between different phases of the project. For example, it might be most important for researchers to understand SDC rules at the beginning of a project while the actual application is only relevant when the project is nearing its end. At this final stage, researchers usually submit their results for SDC and need to ensure compliance with the RDC’s rules. The tools’ focus in the next section will reflect this.

3. Practical examples for self-developed tools supporting researchers at the RDSC of the Deutsche Bundesbank

The project start: Researchers would like to get an overview of all available information – The RDSC Landing Page

The first self-developed tool for automatic filtering of information that we would like to present is the RDSC Landing Page. As the number of data sets available in the RDSC has increased continuously in recent years, the number of rules to apply has also become more and more extensive. In addition, the complexity of the datasets themselves has also increased significantly. Examples of these new challenges include the adequate handling of missing values and duplicates as well as an SDC with multiple variables with entities worth protecting. Finally, the increasing availability of large datasets, and thus the need for advanced programming skills, pose new challenges for the SDC.

As a logical consequence, there are increasingly dataset-specific special rules as well as a variety of auxiliary materials that the RDSC provides to researchers’. However, depending on the requested datasets, those challenges might not apply to all research projects and thus not all researchers will need to familiarise themselves with the respective rules. To support researchers visiting the RDSC in finding the right materials at any time and without much effort, the RDSC has built a small application that is available offline in the secure environment. Here, researchers can find a brief summary of relevant information sorted by topic. This allows the RDSC to guide researchers through the information with comparatively little effort and to highlight particularly relevant parts. Detailed support materials, such as researcher guides, are usually only linked.

In our view, the greatest benefit of the RDSC Landing Page is its ability to highlight relations between similar information across different documents and thus point researchers to relevant information for which they were not actually searching. For example, if researchers would like to learn about working with large data, they will most likely go to the respective part in the section on data handling where they will find a helpful guide called “Working with large data at the RDSC” (Gomolka et al, 2021). However, the same report will also be displayed when searching for information on any of the datasets that the RDSC categorises as “large data”. Therefore, a researcher who is interested in e.g. a variable overview will likewise find this very helpful report, the existence of which she might otherwise never have known.

7 Examples of such documents are the three technical reports on “Working with large data at the RDSC” (Gomolka et al, 2021), “Statistical Disclosure Control (SDC) for results derived from aggregated confidential microdata” (Blaschke et al, 2022) and “Linking data for MFIs” (Stahl, 2020).
of. This is an example of a relation that is hard to represent otherwise, e.g. with pre-sorting of information.

Technically, the landing page is an NW.js application built from a set of RMarkdown files using the R package (distill) (Allaire, et al, 2018). This makes it easy to improve and extend the landing page, as the RDSC staff need only a working markdown knowledge. Editing or extending the app takes only a few minutes, which allows for quick iteration and development. Note that a mere website would not serve the same purpose, as modern web browsers block any links to local files, which is an essential feature of the landing page because it needs to work within the secure environment where all internet access is blocked.

Figures A.1 and A.2 in the Annex show two examples from the RDSC Landing Page.

During the research: Researchers need context information that applies directly to their analysis – SDC tools

The two software packages “nobsdes5/ nobsreg5” (for STATA users) and (sdcLog) (for R users), both developed by the RDSC, allow researchers to check whether the results they generated in their research project comply with the RDSC’s SDC rules. To do this, researchers run a special command directly after generating their results and get immediate feedback from the package.

Thus, researchers receive the appropriate information “just in time”. They do not need to go through a full cycle of output submission, checks by RDSC staff and subsequent feedback. Assuming they apply the tools correctly, they know right after the calculation of the result whether or not it complies with the most important SDC rules. Instant feedback allows for quick iteration in case of problems, which is especially helpful for researchers coming a long way to the RDSC premises.

Another strength of both packages is that they provide helpful information for those cases where the results do not comply with SDC rules. Therefore, researchers always know what kind of issue caused the non-compliance and can therefore better react to it. This significantly reduces the workload for both researchers as well as RDSC staff members when performing and checking SDC. However, it is important to understand, that the tools are only semi-automatic and need to be applied appropriately and correctly by researchers, e.g. by modifying the checking commands for their regressions or descriptive tables.

“nobsdes5/ nobsreg5” and “[sdcLog] are both available on the RDSC’s website, which means that researchers can also download them before they come to the Deutsche Bundesbank for their on-site work. In addition to the software packages, we also provide practical examples and guidelines to help researchers familiarise themselves with the software. Of course, this documentation is also available in the RDSC’s secure environment and can easily be found on the RDSC Landing Page (see Figure A.2).

---

Figure A.3 is an excerpt from the (sdcLog) vignette and shows how to use the tool to check whether a mean complies with SDC rules.

Finalising the project: Researchers need to ensure compliance with SDC rules – The Output Submitter

After researchers have concluded their analysis and performed SDC using one of the tools described above to the best of their ability, they need to submit their results to the RDSC for review and approval.

Especially in the case of first-time visiting researchers to the RDSC, many common mistakes and issues occur in this submission: the output is not stored in a permissible format, certain mandatory checks are skipped, or required supplementary information is omitted. The RDSC identifies these issues and discusses them with the researchers, enabling them to improve the quality of their submission to a publishable level. However, this frequently happens after the researchers have already left the RDSC’s premises, necessitating another visit, which costs the researchers additional time and effort.

It is therefore beneficial to enable even researchers not entirely familiar with all RDSC rules to identify as many potential issues with their output submission as possible. To achieve this, we have designed and built a self-service application capable of running a battery of automated tests against the intended output submission even before a human RDSC employee looks at the output. Only output which passes these automated checks will then be forwarded to the RDSC staff for final verification.

The properties of a good automated output submission verification process are:

**Simplicity:** The main beneficiaries of the tool are researchers not yet intimately familiar with the rules and requirements of the RDSC. The usage should therefore be simple enough so that anyone can use it. Time spent learning to use the verification tool could have been spent learning the rules or improving the output and should be minimised. Even experienced researchers can benefit from a simple tool, as it reduces the mental load and allows them to focus on their analysis.

**Speed:** An automated tool that runs near-instantaneously enables a fast, iterative workflow. Researchers can produce a tentative output, check it for issues themselves, and immediately fix them. Less time and mental energy is spent on non-issues.

**Correctness:** Errors in output verification (automated or otherwise) fall into two categories: false positives and false negatives. While perfect correctness is desirable and should be strived for, it is not a realistic goal. Therefore, the tool needs to be able to handle such errors gracefully. False positives occur when the checks produce a finding that is not actually a problem. Researchers must be able to mark the false positive as such and explain themselves. The RDSC staff will then choose whether to accept the explanation. False negatives occur when the tool failed to recognise a genuine problem with the output. To mitigate this, it is important to keep both researchers and RDSC staff aware of which kinds of issues can or cannot be automatically detected. This allows human efforts to be focused on the areas too difficult for the automated process to recognise.

Researchers can launch the tool we have created simply by clicking a shortcut in their main project directory. It auto-detects and displays information about the
project and the researcher starting the submission. It collects confirmations from the researcher required for regulatory reasons and then proceeds to run the automated checks. These currently include the following:

- Does the project directory structure conform to our expectations?
- Does the analysis code at some point run SDC checks?
- Did the analysis code produce log files?
- Do these log files indicate any issues during the SDC checks?
- Does the code run any commands typically associated with SDC violations because they aggregate away data required for SDC checks? (This would cause only a warning, not outright rejection of the submission.)

The researchers have the opportunity to stop (and later rerun) the tool at any time to go back to their own analysis code and make improvements. They can also make comments regarding any finding of the tool. Once they are content with their output submission, the tool compiles all the information it gathered into a small report for the RDSC staff and allows the researchers to submit it by email.

We developed the RDSC Output Submitter ourselves in-house as a native desktop application using web-technologies (javascript, html, css, node). This helps us to modify our tools easily and quickly based on researcher feedback and experience regarding errors in the checks.

Figures A.4 and A.5 in the Annex show two examples from the RDSC’s Output Submitter.

Aligning the three tools by information purpose

Information purpose spans from understanding-oriented (e.g. rules in a legal text) to task-oriented (e.g. applying rules to specific results).

**Figure 2**

![Diagram showing the alignment of tools: Landing Page (Understanding-oriented), Output Submitter, Nobsdes5 / Nobsreg5 and sdcLog (Task-oriented)]

Sources: Authors

4. Discussion and future areas of improvement

Due to its importance and comparatively high resource and time costs, SDC is one of the most debated topics in the field of RDCs. This paper presents three self-developed tools intended to support researchers in performing SDC in an RDC secure environment. All tools share the same goals of supporting researchers in understanding and applying SDC rules. However, the tools follow different
approaches in pursuit of these objectives. At one end of the spectrum is the RDSC Landing Page that helps researchers faster understand the rules. The “nobsdes5 / nobsreg5” and (sdclog) tools support researchers in applying SDC rules and accordingly are to be found at the other end of the spectrum, with the OutputSubmitter located somewhere in the middle. Ultimately, the different approaches reflect the changing focus of researchers during the project’s lifecycle.

All tools in this paper present an improvement over a situation in which only unstructured information is available to researchers. One obvious area of improvement would be to move from semi-automated to fully automated tools. While this would clearly benefit researchers as they would receive more help in performing SDC it also complicates the development of these tools, putting some constraints on RDCs to do this.

Ultimately, the different approaches reflect the changing researchers’ needs during the project’s lifecycle. At the beginning of the project researchers focus most on understanding the rules they need to adhere to. While working on their analysis a researcher’s focus shifts from understanding towards applying SDC rules to check compliance of the generated results. After the analysis is complete, the focus shifts back towards understanding as the focus now is on submitting results for output checking.

References


Annex

RDSC Landing Page – Getting started

This is what researchers see when opening the tool. After reading a brief summary of the most important recommendations, they can easily navigate to other topics.

Figure A.1

Getting started

This app is intended to be be an anchor for researchers who need any kind of information about their visit at the Research Data and Service Centre (RDSC).

You find a range of relevant topics in the top right. Additionally, there is a search field in the top left area.

Top recommendations

1. Even at the very start of your project you should keep in mind that eventually you will have to perform Statistical Disclosure Control (SDC) after obtaining results. This can be very challenging if you modify your data too much. For more information see Output Submission.

2. Always ensure compliance with the prescribed folder and code structure to avoid complications when submitting the output.

3. Familiarize yourself with our SDC packages (robust5 and robreg5 in Stata and irdocog in R) as otherwise your output might be rejected and you will have to come again to our premises to fix your code.

4. If you don’t find the information you need in the landing page, please send an email to info@data and describe what you are looking for.

Sources: Authors
The tool also provides detailed information on SDC rules and available tools for different programming languages as well as recommendations on how to use them.

**RDSC Landing Page – Output Submission**

**Stata**

In Stata, the compliance of output with our data confidentiality rules can be checked using the `nobides5` and `nobisreg5` commands provided by the RDSC.

For help, type the following in the Stata console:

```
help nobides5
help nobisreg5
help maundice
```

Examples and further documentation are available for all researchers in the following folder:

```
examples_rds
```

**R**

In R, you can use the `sdclog` package, which offers similar functionalities as the Stata commands mentioned above.

To read the documentation, see

```
?sdclog::sdc_descriptors
?sdclog::sdc_model
```

Sources: Authors
(sdcLog) – Excerpt from the documentation

The following example for (sdcLog) is taken from the vignette and shows a simple example for the application of (sdcLog) to check whether a mean complies with SDC rules.

**Simple cases**

Consider the case that the mean for val_1 has been calculated and is now to be output as a result:

```r
sdc_descriptives_DT[, .(mean = mean(val_1, na.rm = TRUE))]
#> mean
#> 1: 20.16835
```

Before this result can be released, it must be checked whether all RDC rules for calculating this value have been followed. Thus, the underlying data is checked for compliance with the RDC rules.

This is the simplest case, the descriptive statistic (mean) was calculated for the variable val_1 without further specifications. Required arguments of `sdc_descriptives()` are the data set (data), the ID variable (id_var) and the variable for which the statistics were calculated (val_var):

```r
sdc_descriptives(data = sdc_descriptives_DT, id_var = "id", val_var = "val_1")
#> # SDC results (descriptives)  
#> OPTIONS: sdc.n_ids: 5 | sdc.n_ids_dominance: 2 | sdc.share_dominance: 0.85  
#> SETTINGS: id_var: id | val_var: val_1 | zero_as_NA: FALSE  
#> ✓ Output complies to RDC rules.
```

Since there are no problems at this point, the function runs without warnings and returns (invisibly) a list of information containing options, settings and the checked criteria `sdc_n_ids` and `sdc_dominance`.

Options and settings are always printed to show that all specifications are set according to RDC rules. From the output above follows that there are at least 5 distinct entities required (sdc.n_ids: 5) and that dominance is defined as 2 entities (sdc.n_ids_dominance: 2) with a value share of more than 85 percent (sdc.share_dominance: 0.85). This reflects the standard values for the options. For details on setting options see the separate vignette on options.

The settings show again which arguments were specified in the function call and vary depending on the `sdc_function`. This is important if the result from `sdc_descriptives()` is not printed right away.

Sources: https://cran.r-project.org/web/packages/sdcLog/vignettes/intro.html
RDSC Output Submitter – Adding check properties

Screenshot right after launching. Here researchers can read a brief introduction on how to use the tool and indicate which results they wish to submit.

Sources: Authors
RDSC Output Submitter – Automated checks

Screenshot after running automated checks, before submission to the RDSC staff, giving the opportunity for correction or comment.

A. Used lines

If you submit all files in the specified transfer sub-folder, you will have used 399 (16%) of the 2500 available lines of output (see principle O.4.2 of the RDSC Rules):

399 / 2500 (16%)

B. Cases of non-compliance with the RDSC Rules

Checks for obvious cases of non-compliance with the RDSC Rules.

The following list shows general violations:
1. No disclosure-control functions (nobsdes5, sdc_model, ...) have been called (see principle O.2.7).

The following table shows violations in your code:

<table>
<thead>
<tr>
<th>#</th>
<th>File</th>
<th>Line</th>
<th>content</th>
<th>Explanation</th>
</tr>
</thead>
<tbody>
<tr>
<td>2</td>
<td>notreal.dta</td>
<td>n/a</td>
<td>n/a</td>
<td>.dta files should not be submitted, rather use .csv files.</td>
</tr>
<tr>
<td>3</td>
<td>something.bak</td>
<td>n/a</td>
<td>n/a</td>
<td>This submitted output-file is of an unusual type (see principle O.5.3).</td>
</tr>
</tbody>
</table>

Your comments

If you are of the opinion that the cases listed in (B) do not constitute a violation of the RDSC Rules, you can comment here. Please use the IDs to refer to a specific check result.

Example for a comment on problem with ID = 2:
#2 - I believe this is unproblematic because of...

Sources: Authors
Introduction to and application of SDC rules using self-developed tools

J. Blaschke, M. Gomolka, C. Hirsch, S. Seltmann and H. Stahl (Deutsche Bundesbank)

11th Biennial IFC Conference on “Post-pandemic landscape for central bank statistics”
Session 2 – Microdata disclosure control: a practical perspective
25 August 2022

The views expressed here do not necessarily reflect the opinion of the Deutsche Bundesbank, the INEXDA network or the Eurosystem.
Introduction to and application of SDC rules using self-developed tools

1. A brief introduction to the work of Research Data Centres (RDCs) (1|2)

RDCs provide secure on-site access to confidential micro data for scientific research

Data selection
Application
Application review
Contract
Data provision

Secure environment at an RDC

Research
Statistical Disclosure Control (SDC)

Publication checking

RDC team
✓ Personal user support
✓ Provision of support documents
✓ Provision of SDC software

Jannick Blaschke, Research Data and Service Centre (Deutsche Bundesbank)
25 August 2022
Page 2
Introduction to and application of SDC rules using self-developed tools

1. A brief introduction to the work of Research Data Centres (RDCs) (2|2)

Provision of support materials (e.g. documents, software) can potentially lead to an information overflow

Scenario 1 - **Little** user support needed
- Small number of datasets
- Easy data structure
- Small dataset size
- Homogeneous legal framework → Similar rules for data access and SDC

Scenario 2 - **Much** user support needed
- Large number of datasets
- Complex data structure (e.g. multiple IDs, missing IDs)
- Large dataset size
- Heterogeneous legal framework → Dataset-specific rules for data access and SDC
Introduction to and application of SDC rules using self-developed tools

2. How tools support researchers in understanding and applying SDC rules

Possible solution for scenario 2: Pre-sorting information and automating as far as possible

Case A
- No structure of material
- No Customized filtering

Case B
- Structured information (e.g. by researcher’s characteristics and purpose of the document)
- No Customized filtering

Case C
- Structured information (e.g. by researcher’s characteristics and purpose of the document)
- Customized filtering

Jannick Blaschke, Research Data and Service Centre (Deutsche Bundesbank)
25 August 2022
Introduction to and application of SDC rules using self-developed tools

3. Practical examples for self-developed BBk tools - RDSC Landing Page

The Project start: Researchers would like to get an overview of all available information

Features

- Structured overview of all available resources (e.g. documentation, software)
- Show relations between similar information across documents → Point researchers to relevant information they were not actually searching

Structured information

Technical set-up

- NW.js application, which is built using Rmarkdown and the R package {distill}
- Advantages: Easy to modify, possibility to open links to local files
Introduction to and application of SDC rules using self-developed tools

3. Practical examples for self-developed BBk tools - **SDC tools**

*During the research:* Researchers need information that applies directly to their analysis

---

**Features**

Researchers can use commands after generating a result (e.g. descriptive or regression table) and get immediate feedback, if the table will pass SDC or not

- No need to program checks themselves
- Prerequisite: Researchers need to correctly apply the commands

**Example from nobsdes5:**

```
.nobsdes5 id x, by(year) notab
```

**Disclosure problem:**
Share of largest two IDs > 85%
Smallest number of distinct IDs (id) of variable x for year: too small
Introduction to and application of SDC rules using self-developed tools

3. Practical examples for self-developed BBk tools - RDSC Output Submitter (1|2)

Finalizing the project: Researchers need to ensure compliance with SDC rules

Features

- Automated checks for selected rules (e.g. file format, folder structure)
- Warnings for potential rule breaches
- Automated count of remaining lines of output

→ Customized filtering of information

Technical set-up

- Web-technologies (javascript, html, css, node)
- Runs locally, allowing access to research project code, logs and output files
Introduction to and application of SDC rules using self-developed tools

3. Practical examples for self-developed BBk tools - RDSC Output Submitter (2/2)

Example

Explicit reference to rule that was breached.
Future areas of improvement

Move from semi-automated to fully automated tools. While this would clearly benefit researchers as they get more help in performing SDC it also complicates the development of these tools putting some constraints on RDCs to do this.
Thank you

Jannick.blaschke@bundesbank.de

Website: www.bundesbank.de/rdsc
Contact: fdsz@bundesbank.de