Irving Fisher Committee on Central Bank Statistics

IFC Bulletin

No 37

Integrated management of micro-databases

Deepening business intelligence within central banks' statistical systems

Proceedings of the Porto Workshop, 20–22 June 2013

January 2014

BANK FOR INTERNATIONAL SETTLEMENTS

Contributions in this volume were prepared for the IFC Porto Workshop held in Porto on 20-22 June 2013. The views expressed are those of the authors and do not necessarily reflect the views of the IFC, its members, the BIS and the institutions represented at the meeting.

This publication is available on the BIS website (www.bis.org).

© *Bank for International Settlements 2014. All rights reserved. Brief excerpts may be reproduced or translated provided the source is stated.*

ISSN 1991-7511 (online)

ISBN 92-9197-975-9 (online)

Contents

Governor's message to the workshop
Carlos da Silva Costa 1

Key points of the Porto Workshop on Integrated Management of Micro-databases
João Cadete de Matos 3

Opening of the workshop

Motivation
João Cadete de Matos 5

Session A: Central Credit Registers I

Challenges and prospects for setting-up a European Union shared system on credit
Ana M de Almeida and Violetta Damia 7

Malaysia's experience in managing the credit registers
Rosnizam bin Saari 11

The Central Bank of Armenia's project of a credit registry
Liana Kirakosyan 15

The evolution of loan level data in Ireland
Rory McElligott and Martin O'Brien 18

Session B: Security-by-security Databases

Which value added can a securities database platform provide to a central bank?
Frank Mayerlen 24

The Czech National Bank's approach to security-by-security data collection and compilation
Milan Nejman 28

Coping with challenges in establishing security-by-security database in Thailand
Wongwatcharapaiboon Weeraya 31

Micro-databases on securities statistics as a key issue for satisfying user needs
Luís Teles Dias 36

Session C: Central Balance Sheet Databases and Business Registers

Building business registers to monitor entrepreneurial dynamics
Homero Gonçalves and Mário Lourenço 42

Using registers as the basis to statistical production – linking issues and potential bottlenecks
Jaanus Kroon and Ain Paas 46

Legal Entity Identifier or the utility approach to shared, high-quality basic micro-data
Francis Gross 51

Efficient ways of dealing with accounting data from enterprises
Margarida Brites 57

ECCBSO: a tool for the analysis of the non-financial sector of the economy (from individual to consolidated accounts)
Manuel Ortega 61

Session D: Central Credit Registers II

Credit Information System – the use of micro-data in credit statistics
Renato Baldini Jr 65

A flexible approach to credit statistics: the use of the Portuguese Central Credit Register for statistical compilation
Paula Casimiro 68

Improving the breakdown of the Central Credit Register data by category of enterprises
Jean-Pierre Villetelle 74

Using credit registry data to assess the risks in the financial system – the *Banco de Cabo Verde* experience
Ana Semedo 80

Application of micro-data for systemic risk assessment and policy formulation
Karen Lee Fong Ling 83

Session E: Balance of Payments / International Investment Position

Micro-data based BoP data quality management
Zsuzsanna Fekete and Péter Bánhegyi 88

The use of micro-data to enhance the compilation and analysis of BoP statistics: the case of South Africa
Stefaans Walters 92

How granular should data be to compile BoP statistics?
Carla Marques 96

Challenges in moving to micro-data based external sector statistics – Japan's case
Hidetoshi Takeda 101

The Bulgarian National Bank's experience with loan-by-loan data collection for the compilation of external statistics
Emil Dimitrov 103

Session F: Prospective Data Repositories for Statistical Purposes

Understanding money markets via surveys
Peter Neudorfer 108

The use of payments data to improve monetary and financial analysis
Filipa Lima 111

Is micro-data reporting the future for interest rate statistics?
Patrícia Antunes, Inês Correia and Rita Lameira 115

The use of central bank databases for macroprudential analysis
Laura Vajanne 120

Panel Discussion

Summary of the parallel sessions (A-F) and the panel discussion
Enrico D'Onofrio, Jesmond Pulé, Timur Hülagü, Jean-Marc Israël,
Jacques Fournier, Gregor Bajtay, Katherine Hennings 123

Session G: Micro-databases Integration and Business Intelligence

Summary and conclusions of session on "Micro-databases Integration and Business Intelligence"
Miguel de Castro Neto 132

A Research Data and Service Centre (RDSC) at the Deutsche Bundesbank– a draft concept
Ulf von Kalckreuth 135

Business intelligence tools and micro-data related to the South African equities market: application and experience
Zeph Nhleko 141

Data management – running systems and ongoing projects in the National Bank of Romania
Andreea Varjoghe 146

Business intelligence in securities statistics: an evolutionary approach
Maria do Carmo Aguiar 149

The use of micro-data in the financial system information model of Banco de México
Alejandro Gaytán González 153

Session H: Links between Micro- and Macro-Data I

Summary of session on "Links between Micro and Macro Data I"
Aurel Schubert 158

Integrating micro-databases for statistical purposes
Paula Menezes and Luís D'Aguiar 160

Use of tax micro-data for the compilation of the Chilean National Accounts
Ivette Fernández 165

Micro-data – the Austrian concept with more granularity as a lesson learned from the financial crises
Johannes Turner, Erich Hille and Günther Sedlacek 168

From data to information and from information to policy making – the story of the "Integrated Company and Industry Analysis Platform"
Çakir Şaban Murat 171

Session I: Links between Micro- and Macro-Data II

Integration of different micro-databases – a significant value added for statisticians
Francesca Monacelli 179

Do macroprudential tools require micro-data?
Perttu Korhonen 183

The use of micro-data to develop remaining maturity data for debt securities and loans
Kuniko Moriya 186

Using trade repository data for systemic risk monitoring
Erik Heitfield 190

New micro-data on household wealth in Denmark: perspectives for national accounts
Martin Oksbjerg 192

Message to the Workshop on "Integrated Management of Micro-databases"[1]

Carlos da Silva Costa[2]

Ladies and Gentlemen,

Welcome to this Workshop and to the city of Porto.

It is indeed a very great pleasure to welcome you all – for a number of reasons.

First of all, as I studied and lectured at the School of Economics and Business of the University of Porto, I have a special attachment to this city. And it always pleases me to know that people keep coming to Porto, enjoying what this city has to offer and benefiting from its hospitality.

Secondly – and this is an important issue –, I think that the availability of information at the micro level is a key tool to understand the behaviour of both firms and private individuals. For this purpose, having a good data collection system, one that also makes use of the existing micro-databases, is a necessary condition, although not a sufficient one – in addition, we need to know how to get the most out of the available data.

That leads me to my third reason. As central bankers, our responsibilities in this area have now increased significantly, given the fact that Information Technology (IT) systems are providing us with new capabilities to draw on the existing micro-databases and to put their data to good use. However, in order to do just that, we need to develop econometric methods and other analytical tools, capable of allowing us to exploit those databases thoroughly. And that is why I think that your task, as statisticians, is so crucial.

Compiling statistics is the basis of your work; but we need to know how to proceed once this task is completed. In other words, we need to understand reality and, for that purpose, the work that you are doing in terms of exploring the existing micro-databases is indeed a very significant one. I am sure that the papers that you will be presenting here and the speeches that you will be delivering on this occasion will be instrumental in taking a step forward in this area. As regards *Banco de Portugal*, I can assure you that we take very seriously the statistical potential of micro-databases and we think highly of the work that our staff is doing, both in terms of data collection and in data mining. It certainly is very important for our own policy decisions.

I come now to my final remark. In order to carry out macro-supervision comprehensively, to understand the behaviour of companies and individuals, and to realize what is happening as regards financial reality, we need to know how diverse reality is and look beyond the aggregate level. Then we will be able to see what is going on at the level of the different groups of individuals or firms and to look for

[1] Transcript of the video message addressed to the Workshop participants.

[2] Governor, Banco de Portugal.

the reasons that are behind this diversity. Your work is a very important catalyst for reaching such an understanding.

To conclude, I just want to say that you are more than welcome to visit us at *Banco de Portugal*, should you would like to have a closer look and a better understanding of the work that we have been doing in the field of statistics.

Thank you very much for being in Porto. Please enjoy your stay. I hope that the outcome of this Workshop will provide a significant benchmark for your future progress in this area.

Again, thank you very much and goodbye.

Key points of the Porto Workshop on Integrated Management of Micro-databases

João Cadete de Matos, Aurel Schubert and Paul Van den Bergh

On 20 and 21 June 2013, the Irving Fischer Committee on Central Bank Statistics (IFC) co-sponsored a Workshop on Integrated Management of Micro-databases with Banco de Portugal and the European Central Bank. The meeting was addressed by Mr Carlos da Silva Costa, Governor of the Banco de Portugal who noted that "*..., as central bankers, our responsibilities in this area have now increased significantly, given the fact that IT systems are providing us with new capabilities to draw on the existing micro-databases and to put their data to good use. ... If we want to do macro-supervision, if we want to understand the behaviour of companies and individuals, if we want to realize what is happening as regards financial reality, we need to know how diverse reality is and we need to look further than the aggregate level and try to see what is going on at the level of the different groups of individuals or firms and to look for the reasons that are behind this diversity.*" Mr João Cadete de Matos, Director of the Statistics Department at the Banco de Portugal, presided over the Workshop in cooperation with Aurel Schubert, Director General Statistics at the ECB, and Paul Van den Bergh, Head of the IFC Secretariat at the Bank for International Settlements. The workshop was attended by 142 officials from 44 central banks from all regions. A total of 42 papers were presented.

The workshop yielded the following insights:

1. Central banks generate and have access to huge amounts of micro-data. This includes raw data underlying statistical data collections (based on censuses and surveys); administrative data sets such as security-by-security databases, credit registers, central balance sheet registers, payment and settlement systems, trade repositories; and third-party data sources (institutional and commercial). Finally there are Big Data from the Internet.
2. Micro-databases have been around for quite a long time but central bank statisticians have been constrained to fully exploit their potential. IT solutions are now becoming available that facilitate the processing of huge quantities of data not only from the same source but also from separate sources.
3. There are clear benefits in using micro-data, such as:
    - reducing the respondents' reporting burden,
    - enhancing data quality and quality control procedures,
    - increasing the efficiency of the statistical production process,
    - improving responsiveness to users' requirements, by expanding the range of available statistics and providing higher flexibility and agility, in particular in the context macro-economic analyses and risk-based financial stability and micro-supervisory analyses,
    - providing an excellent tool for economic analysis and research.
4. There are also various constraints to using micro-data, in particular those dealing with the problems of record matching, data coherence, under-coverage, missing variables, reference data consistency over the various

domains, differences in concepts and statistical units and data updating equally reflected across domains.

5. Business Intelligence tools can be used to integrate micro-databases. Such tools include data warehouses, registries, central repositories or dictionaries; building smart cubes; ETL tools; using aliases; anonymisation of data; digital signatures for accessing data. These are not new but are becoming increasingly important for statistical IT applications.

6. Business Intelligence tools are not by themselves the solution to managing and integrating micro-databases. As important are:

   - appropriate governance of statistical functions in the central bank,
   - improving the skill mix of data managers (eg registration of data, data validation),
   - good management of key common identifiers (eg business registers) and reference data (eg bank identification numbers),
   - good cooperation between economists, statisticians and IT experts,
   - appropriate solutions to legal, confidentiality and privacy constraints to facilitate access by users,
   - use of data standards such as SDMX (statistics), XBRL (accounting), DDI (surveys),
   - education of users, including through online tools.

7. Maximising the benefits while reducing the costs of using micro-data will take time and should be carried out stepwise. New opportunities will be discovered when integrating micro-data bases, including meeting demands from policy makers and data analysts that were previously difficult to accommodate.

The papers of the workshop are available at http://www.bis.org/ifc and will be compiled into an IFC Bulletin. The findings of the workshop will be reported to BIS Governors as part of the yearly IFC Annual Report.

Motivation

João Cadete de Matos[1]

National central banks generate and have access to huge amounts of data. In spite of that, the issue of "data gaps" has been attracting a lot of attention within the central banks' community for some time now. Particularly in the wake of the financial crisis of 2007–09, the statistical systems have been experiencing additional pressure to expand their reach with a view to covering areas where the shortage of appropriate data had been more damaging – *e.g.*, information for the purpose of financial stability analysis and systemic risk assessment.

However, trying to keep up with the rapid changes of the economy and continuously adapting the statistics to new phenomena has some serious limitations. Conventional data collecting systems cannot simply keep on expanding indefinitely to cope with the need to fill the information gaps perceived by the users or in anticipation to their future data requirements.

In this context, it makes sense to exploit the largely unused statistical potential of the available micro-databases covering different areas of the economy and the financial markets – as in the case of, *inter alia*, records from existing central credit registers and business registers, and data related to the prudential supervision function. Once statistically edited, these micro-data might play an important role in enhancing the efficiency of central banks' statistical systems. The granular nature of this information, together with a good coverage of the relevant population, offers increased flexibility as regards the compilation of new statistics (*e.g.*, those related to financial and other structural innovations) and a more rapid response to *ad hoc* data requirements from the users. In general, this approach is technically easy to implement and with relatively low costs associated.

The evolution in network and communication protocols, database systems and multidimensional analytical systems has somewhat removed the potential disadvantages of having to deal with the huge amounts of data normally associated with the handling of micro-databases. In addition, these developments created the objective conditions for the statistical systems based on the so-called stove-pipe model, in which statistics in individual domains have developed independently from each other, to evolve to the next level, which is that of coherent and fully integrated data systems, enabling rapid data exploration, multidimensional analysis and cross-referencing of multiple sources with different granularities. With this objective in mind, many organizations have been developing, for the last two decades, their own data warehousing projects, ranging from combining multiple legacy systems to developing user interface tools for analysis and reporting.

This ongoing paradigm shift seems all together indispensable and irreversible, but is one that involves innumerable difficulties and challenges. One way to deal with these problems is to discuss them within the central banks' community and try to learn from each other experiences. The Porto Workshop on Integrated Management of Micro-databases is just one step in that direction.

[1] Statistics Department, Banco de Portugal, jcmatos@bportugal.pt.

Structure of the workshop

In keeping with the Workshop's preliminary programme (see annex), the one and a half days of the event will be organised as follows:

1. On the first day, the whole morning and part of the afternoon will be dedicated to discuss the potential and actual uses of micro-databases for central bank statistics, in particular for macro-analysis. To accommodate as much as possible the high number of proposals for presentations on this topic that were received so far from central banks and international organizations around the world, this part of the Workshop will be organised in parallel sessions, devoted to describing the characteristics and uses of existing micro-databases, by statistical domain (*e.g.*, central credit registers, central balance sheet databases, security-by-security databases, balance of payments/international investment position data), as well as the possible uses for statistical purposes of other, less traditional, data depositories, such as data from payment and settlement systems and the *Bankscope* commercial database. In the ensuing panel discussion, the participants will attempt at drawing conclusions from the earlier discussions.

2. The second half of the afternoon and most of the morning of the following day will be centred on data integration issues, essentially from a front-end (*i.e.*, the users' side) standpoint. Participants will be encouraged to describe the way in which their organizations retrieve information from different sources (*i.e.*, from multiple micro-databases) and assemble it in a unified view, as well as the techniques they use and the problems they face in so doing. Discussions are also expected to tackle the methodological, statistical and technical issues associated with combining aggregate data with micro-data, the use of micro-data to generate macroeconomic and statistical information and the possibility to drill down through the data hierarchy to examine increasingly granular levels of detail associated with combining aggregate and survey data. The use of business intelligence technologies to help decision-makers use data, documents, knowledge and analytical models to identify and solve problems in a more effective way, will be addressed here as well.

Challenges and prospects for setting-up a European Union shared system on credit

Ana M de Almeida[1] and Violetta Damia[2]

1. Background

Central credit registers (CCRs) are granular databases operated by national central banks (NCBs) containing loan-level or borrower-level information locally tailored to provide for exchange of credit information within the financial system, especially among banks, and additionally they serve to support micro-supervision analysis. Currently, there are three main uses of CCRs: (1) to enable bank supervisors to accurately assess credit risk in supervised financial institutions; (2) to support financial transactions by assisting credit institutions in the evaluation of risk; and (3) for economic analysis.

Since 2007, the European System of Central Banks (ESCB) has been exploring the potential statistical use of CCRs, in particular to understand to which extent their content may be enhanced and adapted to euro area and European Union (EU) statistical needs, to alleviate the statistical reporting burden and increase transparency. In this context, several related ESCB initiatives have not only proven the analytical usefulness of CCRs but have also shown that, where applicable (since in some countries there is no CCR), the differences in terms of coverage, attributes and data content are often substantial pointing to the need for (1) harmonisation in concepts and definitions, (2) over time, convergence in data coverage and content.

2. ESCB Task Force on Credit Registers

Against this background, an ESCB Task Force on Credit Registers (hereinto mentioned as the TF), comprising experts from the statistical as well as from credit registers' areas, was mandated in 2012 to investigate three main issues: (1) identify a core set of information to meet main users' needs and the necessary data attributes and level of harmonisation of definitions/methodologies; (2) consider the governance, legal and confidentiality issues; and (3) explore the identification of entities[3] and of loans and the CCRs' links to other data sources such as micro databases and business registers.

In addition, the TF was assigned to work on a 'pilot area' with the aim to study the identification of a core set of information to meet main users' needs, in terms of analysis envisaged, required coverage of the data in focus, available definitions and

[1] Statistics Department, Banco de Portugal, Ana.Margarida.Almeida@bportugal.pt.

[2] Directorate General Statistics, European Central Bank, violetta.damia@ecb.europa.eu.

[3] Including prospects for the use of the, to be established, Legal Entity Identifier (LEI).

concepts and the related necessary data attributes. The pilot was based on a proposal for sharing a subset of data, prepared in cooperation with ECB internal users from the Financial Stability, Economics and Research departments. The users were then asked to assess whether the dataset (or subsets of it) were indeed fit-for-use for their data analyses, and provide recommendations, also in comparison with other available data sources.

Overall the pilot exercise showed that loan-by-loan, or at least borrower-by-borrower, data availability may further support, including at country level, a number of monetary policy relevant issues relating to the provision of credit and the functioning of the transmission mechanism. In addition, from a financial stability point of view, such granular data allow for a more accurate analysis of credit risk, while the granularity of information on lending could complement (or even replace) existing data sources for monitoring the degree of financial integration within the euro area. Nonetheless, all user areas stressed that the analysis performed had largely been hampered by the lack of further cross-country harmonisation, limited country coverage and incompleteness of the data (lacking further information on collateral, ratings, etc.), thus paving the way for (longer-term) developments towards more harmonisation and data comparability in the future.

In addition, a light fact-finding exercise on the legal frameworks prevailing at national level among the participating members has shown that relevant confidentiality and data protection obligations apply, arising under both the EU and national laws, and that accessibility to CCR raw data is often restricted for supervisory purposes.[4] The overall TF work resulted in a number of recommendations, which refer to the gaps on credit and credit risk data, and include proposals to coordinate efforts to improve granular data among NCBs on lenders, borrowers and credit attributes, as well as methodological issues, covering the following areas:

(1) reporting framework and methodological issues, with a strong encouragement to collect loan-by-loan information;

(2) coverage and determination of lenders, borrowers and types of credit;

(3) core set of data, i.e. required attributes on lenders, borrowers, credit, currency, collateral and maturity etc.;

(4) other information requirements, i.e. additional attributes that the CCRs are encouraged to incorporate, such as information related to the identification of entities and loans, performance of a borrower/loan, probability of default, loss given default, financial guarantees, inception and maturity dates of the loan, effective annualised interest rate and accrued interest; and

(5) other considerations regarding the database management, such as historical data and revisions, data frequency and timeliness, and threshold.

[4] In addition, data sharing with other CCRs follows the rules of the initiated in 2005 Memorandum of Understanding (MoU), on the exchange of information among CCRs, signed by nine NCBs (AT, BE, CZ, DE, ES, FR, IT, PT, RO), and based on reciprocity. The MoU defines the framework to exchange information among the CCRs to provide this information to the reporting institutions to obtain a better overview of the indebtedness of the borrowers in an environment of financial integration, in which cross-border lending is becoming increasingly important. Data aggregation or anonymisation are in general minimum requirements for providing access for non-supervisory users of CCR data, with similar access available to both national and cross-border users on this basis.

3. Way forward and challenges ahead

The CCRs' content could become a primary source of information on credit exposures and credit risks to support the macro and micro-supervisory tasks in the context of the forthcoming (not yet established) Single Supervisory Mechanism (SSM). With a view to defining the best strategic approach towards meeting the increasing user demands for more granular data and flexible data sources in a pan-European set of credit and credit risk data, a number of work-streams have been identified within the recently established (in 2013) ESCB Task Force on Analytical Credit Datasets (also involving countries that do not maintain a CCR), which has been mandated to extend the work of the first TF on the following key issues:

(a) In the *short run*, to (1) monitor CCRs' convergence to the benchmarks identified by the previous TF via a repetition of the pilot exercise, possibly also embracing the emerging SSM requirements, and (2) to investigate the appropriate legal steps for a regular exchange of data serving analytical/policy purposes.

(b) From a *longer term* perspective, and in view of the upcoming implementation of the SSM, to analyse any further ECB needs on granular credit data and overall consider how datasets could be shared across the Eurosystem/ESCB, so as to ensure an in-depth analysis of the data for lenders and borrowers with significant cross-border loans, while limiting duplication with current or enhanced CCRs, or any similar datasets. In this attempt, a new template to collect granular data from NCBs, including attributes on lenders, borrowers, credits and methodological aspects may be developed. In addition, based on the legal framework (including that for the SSM), to investigate the most appropriate legal solution to have the new data collection mechanism in place.

(c) Given that the work focused thus far on non-financial corporations, the need for data on other borrowing sectors can be explored. In particular, as the household sector is usually not adequately covered by CCRs, to investigate the extension of coverage possibly to housing (including mortgage) loans, but also to inter-financial corporations' lending and lending to the government sector (taking into consideration other similar loan-level datasets, such as the European Data Warehouse and "private credit bureaus" that operate in many EU countries).

In particular regarding the second work-stream, the work may be further supported by the establishment of an IT solution (Analytical System on Credit – AnaCredit) for receiving, storing and disseminating credit and credit risk information on a euro area (or a EU-wide scale), which would be sourced from national CCRs or other similar datasets and would include the most important attributes on loans, lenders (possibly abridged from the ESCB Register on Financial Institutions and Affiliates database, RIAD) and borrowers. The system may process some very granular information, *e.g.* on significant loans and borrowers, and more aggregated data (as combination of other, individually less significant). The remainder of the granular information would be handled at national level, under the responsibility of NCBs which would coordinate and cooperate with any relevant national data sources.

Such a tool may substantially support credit and credit-risk analysis required for the conduct of monetary policy, micro-prudential supervision, economic analysis, financial stability, research and statistics, in particular under the on-going turmoil of financial markets. It will enable the necessary flexibility to address in a timely

manner needs on granular data, *e.g.* credit data broken down by economic activity, size of firms, new credit granted (whether, or not, yet drawn), arrears, etc., while the reporting burden is kept to a minimum. Actually, the efforts to set up or enhance such granular credit datasets are expected to be offset, overall by (i) the additional information feedback to reporting agents (as a normal feature of CCRs or similar loan-level datasets) for them to assess the creditworthiness of their (actual or potential) customers and (ii) the substitution effect of the new granular datasets in comparison with additional breakdowns, *e.g.* in Monetary and Financial Statistics (MFI) balance sheet statistics or in (additional breakdowns in) supervisory reports.

To meet the ESCB data needs, methods would have to be worked out in order to overcome the difficulties associated with the data scope, coverage, definitions, reporting framework, as well as the interoperability of credit registers and their links with other sources. In collaboration and with the active participation of all the entities involved, one more coherent and integrated system would have to be developed for the creation of interconnected statistical databases to be used according to the increasing demand for data on credit.

References

Asset-backed securities (ABS) Loan-level Initiative: http://www.ecb.europa.eu/mopo/assets/loanlevel/html/index.en.html

ESCB Register on Financial Institutions and Affiliates Database, RIAD

"General Principles for Credit Reporting.pdf": http://siteresources.worldbank.org/FINANCIALSECTOR/..._text.pdf

Legal entity identifier (LEI): http://www.financialstabilityboard.org/list/fsb_publications/tid_156/index.htm

Mandate of the Joint STC/FSC Task Force on Analytical Credit Datasets

Memorandum of Understanding on the exchange of information among national central credit registers for the purpose of passing it on to reporting institutions: http://www.ecb.int/press/pr/date/2010/html/pr100428_2.en.html

Report of the ESCB Working Group on Monetary and Financial Statistics / Working Group on Credit Registers Joint Task Force on Credit Registers

"Report of the Expert Group on Credit Histories.pdf": http://ec.europa.eu/internal_market/consultations/..._en.pdf.

Malaysia's experience in managing the credit registers

Rosnizam bin Saari[1]

Introduction

The credit reporting institutions play an important role in the area of credit risk management and the promotion of a sound credit culture in the financial system of the country. It is recognised globally that a matured and developed financial sector must have effective and efficient systems to manage credit risks in order to sustain the surveillance and financial needs of the government, regulatory authorities, industries and consumers. This explains why more and more developing economies have taken the cue from the developed economies in establishing credit reporting institutions.

Based on World Bank survey over 185 countries via Doing Business 2013, 145 countries are having credit reporting institutions. It can be a Public Credit Registry (PCR) which is owned and operated by a government agency, mostly by the central bank, or a Private Credit Bureau (PCB) which is owned and operated by a commercial enterprise or a combination of both. In Malaysia's case, the credit reporting landscape has evolved over the last 30 years from the PCR operated by Bank Negara Malaysia (the Bank) as the sole credit reporting institution in the country to more PCBs operated by a few other agencies.

Public Credit Registry (PCR) operation in Malaysia

PCR in Malaysia is established under the Central Bank of Malaysia Act 2009 and it has been in operation since 1982 as the first credit reporting institution in Malaysia. The Act empowers the Bank to collect customers' credit information from lending institutions and furnishes the information back to the institutions for credit assessment, to the customers for verification of data accuracy reported by the lending institutions and to the registered credit reporting agencies for credit assessment services. The operation of PCR by the Bank is well supported by an on-line system known as the Central Credit Reference Information System (CCRIS) which is the key enabler for efficient and reliable data collection and dissemination of credit information. Since the implementation of CCRIS in 2001, it has enabled real time access to both positive and negative credit information of the borrowers. The system contains about 7.9 million customer profiles, which cover individual and non-individual borrowers, with more than 22 million active credit accounts. Figure 1 illustrates the development of PCR support systems in Malaysia since 1982.

[1] Statistical Services Department, Bank Negara Malaysia, rosnizam@bnm.gov.my.

Development of PCR support systems in Malaysia

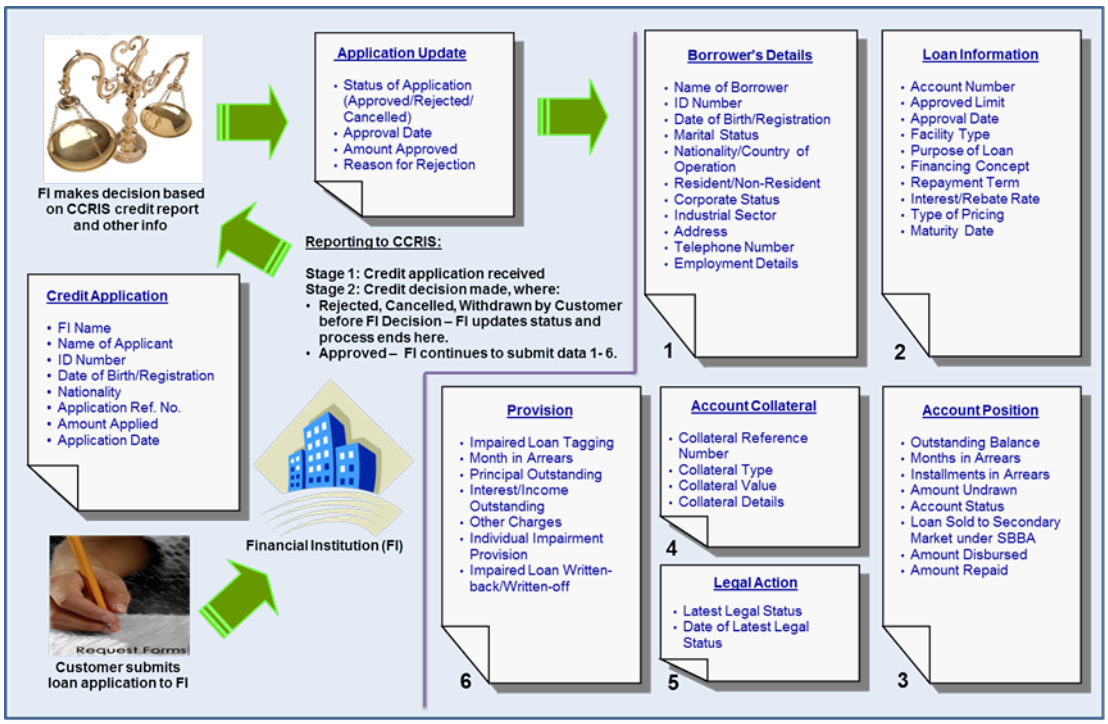
Figure 1

Period	Mode of Data Submission	Mode of Data Dissemination	Reporting Threshold
1982	Hardcopy	Hardcopy	Minimum threshold 1 million ringgit for banks and ½ million ringgit for finance companies of loan outstanding
1988	Diskette	Hardcopy	
1994	Computerised batch submission	Hardcopy	
1999	Computerised on-line and batch submission	Batch on-line	
Since Oct 2001	Real-time on-line and batch submission with implementation of CCRIS	Financial institutions access to real-time on-line enquiry for credit reports	No minimum threshold

The comprehensive nature of the information facilitates the development and validation of credit scoring models by the lending institutions, leading to more robust credit decisions and responsible lending culture. The detailed information provided through CCRIS allows lending institutions to look beyond collateral and cash flow statements to understand the borrower's track record in making payments and predict future payment pattern. This undoubtedly contributes to higher loan growth and better asset quality. The positive contributions can be seen as outstanding loans grew at an annual average rate of 8.9%, while the overall impairment ratio showed a significant decline from its peak of 16.2% in 2001 to 2.2% in 2012. Figure 2 illustrates the reporting stages in CCRIS and the data involved in each stage.

Reporting stages and data involved

Figure 2



In the early stages of its operation, the participation in PCR was limited to banking institutions supervised by the Bank, mainly for regulatory purposes to support the Bank's role as regulator. However, for better analysis and credit decision making, the coverage has been extended to non-bank financial institutions such as hire purchase, leasing and cards companies as well as non-financial institution such as government agencies which provide credit for specific purpose, with participation in the CCRIS on a voluntary basis. Since the latter two types of institutions are not governed by the Bank, selected preset conditions are imposed on them to ensure that data quality and accuracy is well preserved in CCRIS. To date, there are 78 reporting institutions participating in the Malaysia's PCR. These include 27 Commercial Banks, 16 Islamic Banks, 13 Investment Banks, 11 Development Financial Institutions, 4 Insurance Companies, 3 Payment Systems Operators and 4 Other Financial Institutions.

Information collected from all lending institutions is summarised in the form of credit reports which are accessible via CCRIS with minimal charges, depending on the types of report. As for the public, the credit report can be obtained directly from the Bank with no charge via its centralised contact centres at the Bank's headquarter and 5 nationwide offices. In 2011, the Bank launched its mobile customer service coach which complements the existing channels of contact centre to provide greater access and opportunity to the public in semi-urban and rural areas to retrieve credit report.

Development of Private Credit Bureau (PCB)

Compared to PCR, PCBs are a relatively recent development. Although PCBs have existed in Germany, Sweden and the United States for nearly a century, they emerged in many other high-income countries, including France, Italy, and Spain, as recently as the 1990s (World Bank, 2013). In Malaysia, the credit reporting landscape entered a new chapter when the first PCB was established in 1990 to provide another source of information for lending institutions in accessing customers' credit worthiness. There was no specific provision to govern the operation of the PCBs until 2010 when the Credit Reporting Agencies Act was enacted by Parliament of Malaysia to regulate those institutions that conduct credit reporting business. To date, there are 4 PCBs in Malaysia that are supervised by the Registrar Office of Credit Reporting Agencies (RCRA), a division under the Ministry of Finance.

The Act specifies the requirement to preserve customers' privacy. PCBs are required to notify the customers that their information are being collected and PCBs are required to obtain the consent of the customers on disclosure of information to third parties. As data accuracy and security are among the key elements in the Act, PCBs are required to take reasonable steps to ensure that information held are accurate, complete, not misleading and up-to-date. PCBs are also obligated to put in place the necessary safeguards, in term of systems, procedures and processes, to ensure the security and safety of information. To protect consumer rights, the Act has given the privilege to the customers to access their own credit information collected by the PCBs. No charge is imposed on the customer for correcting any information that is inaccurate, incomplete or misleading.

While PCR is focused on borrowings mainly from banking institutions and other selected non-bank financial institutions for macro-prudential and other policy purposes, PCBs offer additional credit information collected from a wide variety of

financial and non-financial entities as well as other sources such as published information to facilitate accurate credit decision making by the lending Institutions. Apart from credit information collected from various sources, including the data from CCRIS, credit scoring and rating are the value added information offered by the PCBs in their reports which differentiate the services provided by PCBs from PCR. The access to the report are chargeable and rates are varies from one PCB to another, depending on the services provided by the PCBs.

Conclusion

Over the past three decades, credit reporting institutions play a vital role in supporting financial stability and promoting credit market efficiency in Malaysia. It is proven that through effective credit reporting systems guided by overall legal and regulatory framework and support received from the market players, the allocation of new credit is increasing year after year while the default risk is reducing. Contribution and collaboration between PCR and PCBs has improved the depth of credit market transactions and broaden access to finances. This was reflected by Malaysia's number one ranking for five consecutive years in term of Ease of Getting Credit in the Doing Business Report, published jointly by the International Finance Corporation and the World Bank. Information sharing among the market players contributes to the responsible credit culture by discouraging excessive debt and rewarding responsible borrowing and repayment.

References

World Bank (2012). "Global Financial Development Report 2013: Rethinking the Role of the State in Finance". Washington, DC. Retrieved 14 April 2013 from: http://econ.worldbank.org/WBSITE/...:8816097,00.html

Ministry of Finance Malaysia (2012), "Summary of Rights Credit Reporting Agencies Act 2010 (Act 710)". Putrajaya. Retrieved 19 April 2013 from: http://www.treasury.gov.my/index.php?option=...=en

World Bank (2013). "Doing Business 2013: Smarter Regulations for Small and Medium-Size Enterprises". Washington, DCD. Retrieved 14 April 2013 from: http://www.doingbusiness.org/data/exploretopics/getting-credit

The Central Bank of Armenia's project of a credit registry

Liana Kirakosyan[1]

Introduction

The Credit Registry of Central Bank of Armenia was set up in 2003. It is an information system that collects, maintains and, upon request, provides information to the creditor concerning different borrowers and loans granted to them. The main objectives of Credit Registry are:

- To help banks to assess the creditworthiness of their clients, avoiding non-diligent and over-indebted borrowers and impose discipline on them.
- To provide supervisors with credit information that could be useful in monitoring credit risk in the entire system.

Information provided to the credit registry

The participation to the registry is mandatory. The reporting organizations are the banks, credit organizations and foreign banks' branches operating in the Republic of Armenia. Data exchange among participants is takes place via local network called CBAnet. Information provided to the CR covers all the legal entities (except for creditors) and natural persons (including individual entrepreneurs) that have liabilities to the banks and credit organizations. Liabilities include:

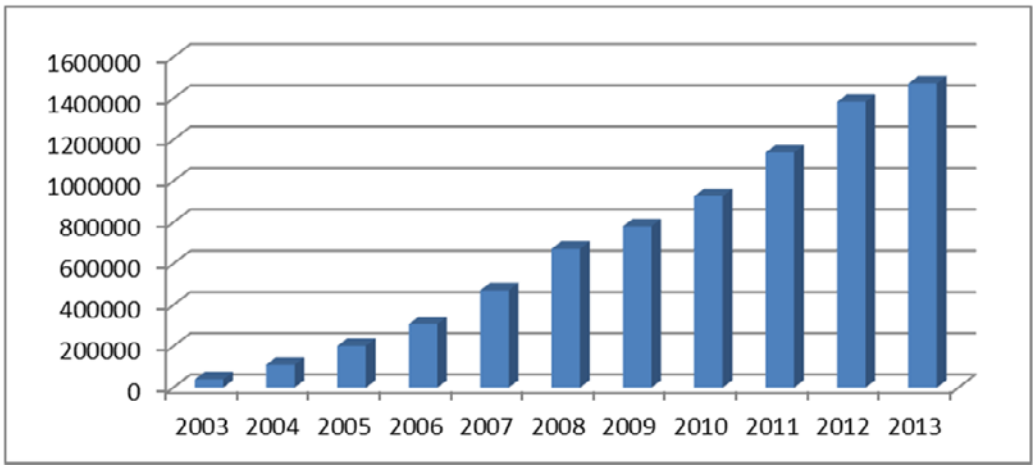
- Loans drawn, including credit lines and overdrafts
- Credit cards
- Financial leasing
- Factoring
- Guarantees issued
- L/C contracts
- Obtaining of securities through repurchase agreement (REPO agreement)
- Any other transaction that generates a monetary obligation to the creditor.

The credit registry collects information about all the loans, the amount of which exceed AMD 1.5 millions and the ones that carry an outstanding amount or interest and are classified as monitored, non-standard, doubtful or bad loans according to the "Procedure on Classification of Loans and Receivables and Creation of Possible Loss Reserves for Banks operating in the Territory of the Republic of Armenia"

[1] Statistics Department, Central Bank of Armenia, liana.kirakosyan@cba.am.

(approved by the CBA Board Resolution 63, 23/04/1999 and the MoFE Order 214-A, 04/06/1999). If a borrower receives several small loans from the same creditor with total amount exceeding AMD 1.5 millions, these loans are reported to the registry separately. In case of any change in a loan data already inputted into the Registry, including changes in classified loans, the creditor must submit such changes to the Credit Registry within three business days following the making of the change. In case of a credit line, credit card and overdraft, changes in data is submitted on the first business day of each week, based on the Sunday data of the previous week.

Number of loans from 2003 to 2013



Content of CR database:

- Clinet identification data (tax code/passport number, legal form, company name/name-lastname for natural persons, address, residency, executive director).
- Data on loan (code of reporting creditor, contract number, total amount, current balance, maturity, number of days past due, past due principal and interest, types of collateral, value of collateral, classification by economic activity, according to NACE, by region, where most part of the credit was used, etc.).
- Data on guarantors, owners, related persons.

Information provided by the credit registry

The users of information collected in the registry are the creditors, borrowers and Central Bank of Armenia. Information provided by the CR is categorizes as:

1. Current information

   1.1. Information in a standardized format

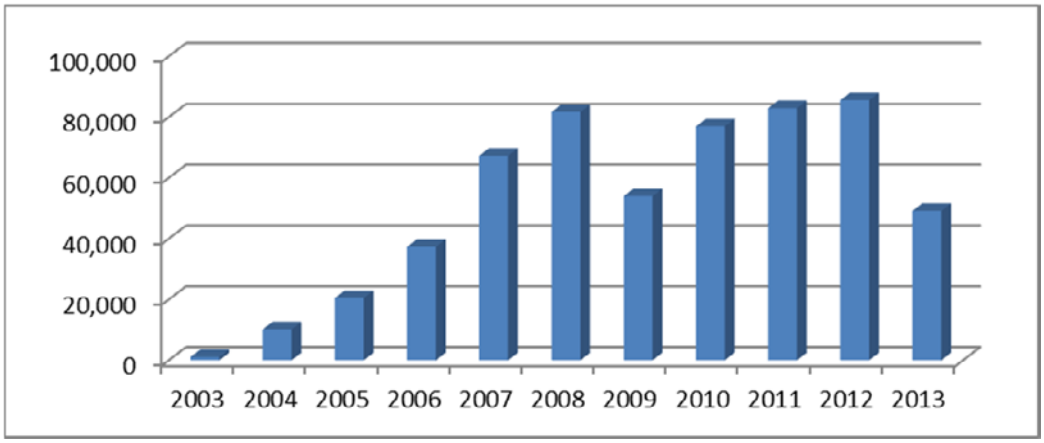
   1.2. Non-standardized information on the borrower available in the Credit Registry (for example data on guarantees provided by the borrower, owners of a company, a credit by reference, etc.)

2. Regular information covering all the borrowers of that creditor, based on this information banks classify their loans

To obtain current information, the creditors make an electronic inquiry to the Credit Registry and get the information within three business days.

The creditors bear responsibility for the content in writing from the related persons, guarantors and owners. They can use the information from the Credit Registry solely for provision of loan to the borrower or appraisal of the borrower's creditworthiness. The creditor can request information only about the parties who have applied for a loan, or who have obtained the loan but have not repaid it. Where the information from the Credit Registry contains inconsistency with the data obtained by the creditors, they inform the Unit to rectify such inconsistency and provide the verified information. The borrower may also obtain information about itself, available with the Credit Registry, by paying a fee determined for such service.

Number of inquiries from 2003 to 2013



The data from CR is also used by different divisions of CBA. For example it is used for carrying out rating, for off-site supervision, for different analysis, etc.

References

"Procedure on Classification of Loans and Receivables and Creation of Possible Loss Reserves for Banks operating in the Territory of the Republic of Armenia", Central Bank of Armenia, Yerevan, 23.04.1999

"Procedure for Creation of Information System of Creditworthiness of Clients of Banks, Credit Organizations, Branches of Foreign Banks operating in the Republic of Armenia, that is of Credit Registry, and Procedure for Participation in Credit Registry", Central Bank of Armenia, Yerevan, 29.03.2005

The evolution of loan level data in Ireland

Rory McElligott[1] and Martin O'Brien[2]

Introduction

In recent years developing a granular credit database has been a key part of the Central Bank of Ireland's (CBI) strategy to tackle the issues faced by the domestic financial sector. The initial motivation for this development was the need for detailed and extensive data on the loan-books of Irish-headquartered banks for prudential purposes. This has since broadened to the establishment of an official granular credit database for wider use by the CBI, industry and consumers. This paper discusses the progress to date in this area, the main challenges faced, and the lessons learned for the future.

In early 2011 the CBI first received detailed loan level data from the Irish-headquartered banks. The initial dataset was developed in a relatively short timeframe by external consultants; primarily for the purpose of loan-loss forecasting as part of a capital requirements assessment. The Central Bank has since taken complete control of the process, and is currently focusing on developing the necessary infrastructure and improving the data quality. A secondary outcome of the dataset has been a rich stream of policy relevant research at the CBI. The flexibility of the granular data allowed analysts at the CBI to gain a previously inaccessible picture of the state of the Irish loan books. The many publicly available papers provide only a sample of the extensive research being undertaken within the CBI.

However, this success is not the end of the road, with the existing dataset having only partial coverage of the banking population. Work is ongoing to establish a Central Credit Register (CCR), aimed at encouraging a more informed credit market for all participants. Legislation to facilitate this is currently before the national parliament and the CBI will have a key responsibility in operationalizing the system. The register will expand the coverage from a limited number of institutions to the full banking population and into the non-bank financial sector. It is hoped the CCR will be fully operational by 2016.

First Steps

On 28 November 2010, the Irish Government agreed to a €67.5bn EU-IMF financial support programme, with accompanying conditional targets. One such target was related to bank recapitalisation, including a capital assessment review to be completed by the CBI by end-March 2011. This review, the Financial Measures

[1] Statistics Division, Central Bank of Ireland, rory.mcelligott@centralbank.ie.

[2] Statistics Division, Central Bank of Ireland, martin.obrien@centralbank.ie.

Program Report (FMP), included a bottom up analysis of the banks' loan portfolios and required detailed loan level data. External consultants were retained to complete the capital assessment review, including the data sourcing. Progress was exceptionally quick, given the conditionality of the process as part of the EU-IMF programme and the potential negative impact that poor quality data may have had on the results for individual banks. The review, including the collection, quality remediation work, and analysis, was completed just over four months after the signing of the EU-IMF programme.

As outlined in the FMP, the Central Bank conducted a data integrity and verification exercise to ensure robust outputs. The validation activities included:

1. Data tape to balance sheet reconciliation: The review found that data tapes for all banks reconciled to unaudited financial accounts within 0.02% of notional value.
2. Loan file sampling and testing: the data provided by the banks were compared to their source systems and the results provided a more detailed understanding of the strengths and weaknesses of the data.
3. Review of bank IT systems: this highlighted that several banks had multiple and fragmented systems, but had procedures in place to address data quality issues or passed comprehensive testing.
4. Assessments of loan portfolios: in-depth assessments of loan portfolios were conducted by reviewing and re-underwriting individual loan files: For example, the banks' existing internal ratings were risk assessed and benchmarked to a common external scale.

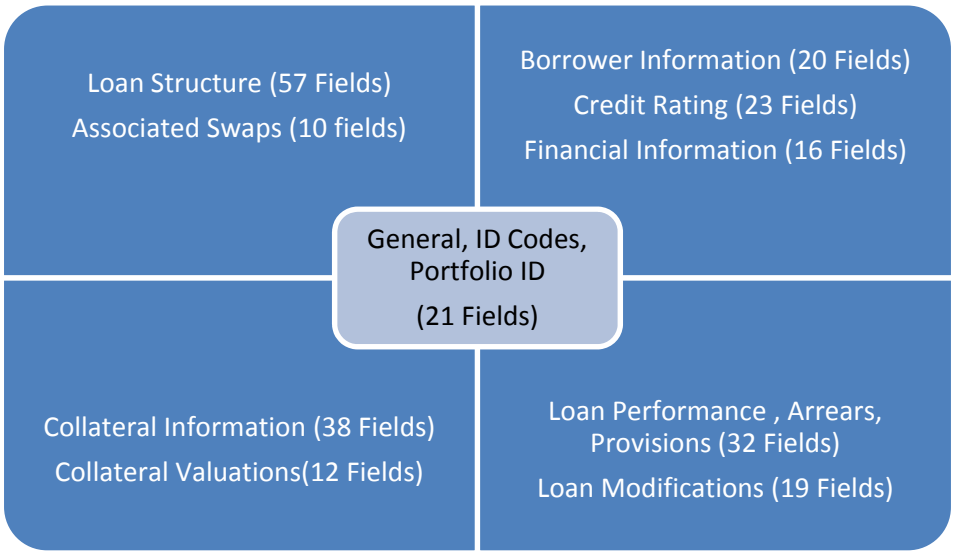
The results of the validation activities highlighted some areas of concern including significant error rates for certain portfolios. In addition, significant portions of data were missing from the individual loan files (particularly for Commercial Real Estate and SME loans).

Moving the process in-house

While the approach outlined above was successful and extremely quick, it also came at a substantial financial cost. A more sustainable solution was required for the medium to long term and the CBI initiated a project to develop the internal capabilities to collect and process the loan level data. A granular data request was sent to three pillar banks (Bank of Ireland, Allied Irish Banks (including EBS) and Irish Life and Permanent) in Q1 2012 for loan level data over the following asset types: Retail Mortgages, Retail Non-Mortgages, Micro-SME, Corporate/SME and Commercial Real estate. In total, 248 fields were requested for each loan, with some relevant only to certain portfolios or loans. The fields covered include information on the original loan contract, borrower details, loan performance (see Figure 1). Coverage levels of these fields by individual Banks range from 56%-67%. To date, three data drops have occurred with reference to December 2011, June 2012 and December 2012.

Information Collected on a Loan Level Basis

Figure 1



The importance of policy application

Analysts within the CBI quickly embraced the new dataset because of the richness of micro level data. As a result, highly policy relevant research emerged that supplied a previously inaccessible picture of the state of the Irish loan books (e.g. Kennedy et al, 2011, Lawless et al, 2012, McGuinness, 2011). More recent work has concentrated on improving the understanding and predication of potential losses within the portfolio (Kelly, 2012) and testing the interaction between certain macroeconomic policies and the loan portfolio (Kelly et al, 2013). These papers represent only the tip of the iceberg on the analysis being conducted within the CBI addressing the policy issues such the large overhang of mortgage arrears in Ireland. The adoption of these loan level datasets by researchers is vitally important for the future of the loan level data initiative. Not only do the policy outputs justify the investment to date, but the flexibility of such datasets emphasises to senior policy makers the importance of having such data available. It is important that this buy-in by senior management is maintained. These early analytical papers help maintain the initial momentum and support the business case to further develop the project.

Example of output from loan level data

Figure 2

			Negative Equity	
			No	Yes
Arrears 90+ days	No	Number of properties	242360	134007
		Negative equity (€ Mn)	-	9081
		Total arrears (€ Mn)	-	-
		Average balance (€)	120501	258761
		Total balance (€ Mn)	29200	34676
	Yes	Number of properties	9598	9515
		Negative equity (€ Mn)	-	798
		Total arrears (€ Mn)	113	160
		Average balance (€)	130783	272991
		Total balance (€ Mn)	1255	2598

Source: McGuinness (2011), based on 31 Dec 2010 data.

Progress towards a full Central Credit Register

Ireland has committed under the EU-IMF programme to develop a Central Credit Register. The legislation underpinning the creation of a national credit register is currently before the national parliament. At the time of writing (late-April 2013), it is expected that enactment may occur either before the summer recess or in the autumn. Once the Bill is passed into law, the CBI will be charged with establishing and maintaining the register. It is also envisaged that the CBI will have access to the CCR database for the performance of all its functions, thereby maximising the wider use of the CCR data in an appropriate manner. There is uncertainty around the go-live date but a conservative estimate is no later than 2016. Under the latest Memorandum of Understanding the Irish authorities must ensure that the Register is at an advanced stage of development by end-2013 subject to passing of the necessary legislation. A number of key issues are outstanding at this stage:

- Unique identifier: The CBI is pushing for the inclusion of personal public service number (PPSN) to assist in creating a single borrower view across various banks and finance companies. The Data Protection Commissioner has concerns about collecting such a unique identifier. Such concerns are valid but from the, admittedly narrow, view of an end user, the usefulness and accuracy of the end database will be much reduced by the absence of a widely used unique identifier. Ireland does not currently use postcodes, so in the absence of PPSN, matching people becomes even more difficult;
- Access: The extent of access for several key groups is not yet finalised. Open issues are the frequency of access that lenders should have, level of access for

other groups such as new market entrants, European public registers, overseas lenders etc.;

- Scope of Register: what groups of borrowers and lenders will be captured;
- Management of day-to-day operations: the most appropriate operational structure for the CCR within the CBI (or potentially outsourced) given the CBI rights and responsibilities as owners of the database.

There are still considerable operational obstacles to be overcome and practical decisions and scoping out of the required IT systems, information security arrangements, charging structures etc. Critical data definition issues are also on-going, with input from CBI Economist/Statisticians. Only when this is further developed can industry begin to actively position itself to meet its obligations.

Some lessons to date

Statisticians have been actively involved in assisting design of the credit register but this will not be a credit register designed for statisticians. As with other CCRs, the primary objectives will be to have a dataset for prudential and credit reporting purposes. Many of the principles outlined in various reports such as harmonisation of definitions across borders, international sharing of data etc. (World Bank, 2011), are unlikely to be fully incorporated at inception. To maximise the usability of the CCR data, the ongoing involvement of statisticians is important in the data definitions and mapping, where applicable, CCR data back to other standard classifications (ESA, etc.). This also promises to speed up the delivery of new statistics to meet user requirements in a way which minimises the cost to industry. In establishing/manipulating granular credit data, the CBI has faced a new challenge in developing staff skill sets to manage and integrate micro-databases. These challenges will continue to be significant as the possibility for linking the CCR data with other micro-databases for policy/research uses are explored. The need for commonly used identifiers is key to the success of such initiatives.

References

Central Bank of Ireland (2011). "The Financial Measures Programme Report", March 2011.

European Commission (2013). "Ireland Memorandum of Understanding on the Specific Economic Policy Conditionality (Eight update)", 12 March 2013.

International Monetary Fund (2010). "Memorandum of Economic and Financial Policies", http://www.imf.org/external/np/loi/2010/irl/120310.pdf , December 2011.

Kelly, R (2011), "The Good, The Bad and The Impaired: A Credit Risk Model of the Irish Mortgage Market"; *Central Bank of Ireland Research technical paper*, 13/RT/11, November 2011.

Kelly, R & McQuinn, K, (2013). "On the hook for impaired bank lending: Do sovereign-bank inter-linkages affect the fiscal multiplier?", *Central Bank of Ireland Research technical paper*, 1/RT/13, March 2013.

Kennedy, G & McIndoe Calder, T (2011), "The Irish Mortgage Market: Stylised Facts, Negative Equity and Arrears", *Central Bank of Ireland Research technical paper*, 12/RT/11, November 2011.

Lawless, M & McCann, F (2012). "The Irish SME lending market – a snapshot, December 2010"; *Central Bank of Ireland Economic Letter Series*, Vol. 2012, No. 3.

McGuinness, A. (2011). "The Distribution of Property Level Mortgage Arrears", *Central Bank of Ireland Economic Letter Series*, Vol. 2011, No. 6.

World Bank (2011) "General Principles for Credit Reporting", September 2011.

Which value added can a securities database platform provide to a central bank?[1]

Frank Mayerlen[2]

Introduction

Since spring 2009 the European System of Central Banks (ESCB) runs the current "Phase 2" of its Centralised Securities Database (CSDB) platform. The CSDB covers debt securities, equity securities as well as investment fund shares where each security is identified by its ISIN code and defined by more than 50 individual attributes. Regarding the geographical coverage, the system includes all of the above securities issued by European Union (EU) residents, all securities potentially held by EU resident investors and all securities issued worldwide which are denominated in euro. In spring 2013 the CSDB covers about 7 million active securities in total.

Drawing from the CSDB experience this note elaborates in Section 1 on the conceptual advantages of such a system while Section 2 provides some further background on how to operate and further develop a securities database platform.

Section 1: why should a central bank have a securities database platform?

Also driven by the financial crisis the availability of reliable and complete information on securities became crucial for central banks. Securities data plays not only a role for the production of aggregate statistics, e.g. when security-by-security (s-b-s) reporting system are used, but also for the detailed statistical analysis or for the conduct of tasks in economic analysis or research. Prudential supervision is another field of interest.

Consistent integration of several data sources for different users in one platform

Securities data referring to a country's reporting agents may be directly collected but respective data referring to non-residents must usually be purchased from commercial data vendors. While many commercial providers claim to have "full coverage", reality often differs and, dependent on the purposes of the data, more than one commercial provider is required. Hence there is no single source for

[1] This Paper should not be reported as representing the views of the European Central Bank (ECB). The views expressed are those of the author and do not necessarily reflect those of the ECB.

[2] Directorate General Statistics, European Central Bank, frank.mayerlen@ecb.europa.eu.

worldwide securities data. As a consequence individual users and user areas of a central bank often maintain individual "handmade" securities databases. This is usually inferior in quality and inefficient, it may lead to inconsistent results and there is the risk that the data usage may not always be compliant with commercial license restrictions or rules of statistical confidentiality.

Against the above it is advisable to set-up a single securities database platform with centrally managed access which can serve all uses of a central bank. With the set-up of the CSDB the ESCB has developed a single and consistent securities database platform which even serves all 27 national central banks (NCBs) of the EU.

Iterative approach – increasing required completeness and accuracy over time

It is desirable to set-up and enhance a system like CSDB in iterations. Indeed, the current CSDB platform has been launched in 2009 to support the production of aggregate statistics where the reporting agents would provide data with a very lean s-b-s dataset which is then enriched with CSDB reference information e.g. on issuer country, issuer sector, security classification and price. In a second step the CSDB has been used to look deeper into the aggregate statistical data and also to prepare very timely proxies for data which is still reported as aggregate, e.g. in the government finance area. Only in a third step the data usage is gradually being expanded to "non-statistical" use, e.g. to support economic analysis or prudential supervision related work.

Wide data usage – with appropriate health warnings

For the statistical usage but in particular when expanding beyond that, expectations regarding data quality have to be managed carefully not only on the user side but also on the side of the data producers. The trade-off to be managed is that data at the level of the individual instrument cannot always be of book-keeping accuracy but still better than any individual alternative. Hence, while being transparent with users and issuing the appropriate "health warnings", one should not shy away from making the data available, given that there is usually no superior solution.

Section 2: operating and further enhancing a securities database platform

Drawing on the CSDB experience, this section describes some practicalities to be considered when operating and enhancing a securities database platform. After a brief overview some thoughts on further development and required staff profile are provided.

Operation

The CSDB is fed daily by several commercial data providers and monthly by more than 20 EU NCBs. The average daily processing covers about 400,000 security records and about 2.5 million prices. Multi source input data may partly overlap and even be inconsistent between different data providers. One of the strengths of the

CSDB system is its fully automated cleansing of the information into one single compound record per security, hereby automatically resolving inconsistencies and gaps based on statistical rules.

On-line access and joint data quality management

The CSDB system is hosted and managed by the European Central Bank (ECB) and provides real-time on-line access to all 27 EU NCBs, who can not only see but also verify and correct the data, if needed. Indeed, the monthly CSDB data quality management (DQM) process is organised in a way where each EU NCB is responsible for the securities issued by issuers resident in their country while the ECB takes care of the securities issued by issuers resident outside the EU. Moreover, the ECB is responsible for "data source management activities", i.e. for the direct contact with commercial data providers to address repetitive or structural problems in their input data directly at the source.

The DQM framework is fully harmonised and supported by dedicated CSDB system functionality, which provides six DQM metrics to prioritise and steer the DQM work as well as the necessary system functionality to correct data, if needed. DQM metrics are fully automated and compare massive amounts of data. As an example, the update of the metric identifying relevant updates in the CSDB data compares 23 attributes for two points in time for over 6.3 million instruments, i.e. about 300 million data points.[3] The DQM framework has been formalised in an ECB Guideline, which will apply as of 1 July 2013.[4]

Data distribution and usage

NCBs are provided monthly with a full extract from the CSDB covering about 9.5 million securities.[5] This end-month data extract, together with revision for previous months, is provided with a high timeliness of only about six working days. In parallel CSDB data can be accessed on-line via a business intelligence tool (Business Objects).

The CSDB DQM process so far supports the use of the data in five different statistical areas: external statistics, statistics on financial vehicle corporations, investment fund statistics, securities holding statistics and statistics on government securities funding. Work is on-going to support also the timely production of securities issues statistics in the future.

Furthermore, as of 2013 CSDB data is made available for dedicated non-statistical purposes, following an agreed procedure. Given that the non-statistical use is not covered by the DQM Guideline, it has to rely on data "as is" and will not be subject to additional DQM work beyond what is anyway done to support the statistical use.

[3] This metric covers all EU issuers and issuers in selected countries outside EU such as US, JP, RU, CH, BR.

[4] See Guideline ECB/2012/21 http://www.ecb.int/ecb/legal/pdf/l_30720121107en00890103.pdf.

[5] The monthly data provision covers also matured securities, which may be needed to revise statistical data.

System maintenance and further development

The current "CSDB Phase 2", went live in April 2009 with Release 2.0. Since then 2–3 maintenance releases have been rolled out per year, with Release 2.9 being currently in operation. Release 2.10 is expected to go live in summer 2013. The relatively high frequency of maintenance releases has proven crucial in improving system functionality in quick but manageable iterations in particular to support new needs as well as efficient DQM work by NCBs and the ECB. Given the huge amount of data which is processed by the CSDB it is a must to continuously increase system automation and to improve efficiency and performance of the system operation.

Further system enhancements could aim at a higher frequency for the data production process, e.g. weekly or even daily, and could also cover the inclusion of further financial instruments which are identified item-by-item with a unique instrument identifier, e.g. securitised derivatives. From a more operational perspective the next development steps may aim at an automated data provider feedback regarding the supplied data quality.

Required staff profile

Running a data intensive operation like CSDB requires a stable and qualified core team seamlessly covering the business perspective and IT. The team must be able to prioritise and manage new needs and development iterations, to handle and analyse substantial datasets and to quickly assess and resolve any related problems. It must also keep close contact with all stakeholders including system development, testing, operation and final data usage.

Compared to the earlier collection and production of aggregate statistics this adds an entirely new dimension to the required staff profile.

The Czech National Bank's approach to security-by-security data collection and compilation

Milan Nejman[1]

Introduction

The increasing complexity of the current economic system poses a major problem for compilers of statistics all over the world. As the central bank of the Czech Republic, the Czech National Bank (CNB) is responsible for collecting data and compiling different kinds of statistics, including securities (holdings) statistics. To keep up with coming changes, it was vital for the CNB to set up a securities holdings data collection system that would be robust in its ability to deal flexibly with new data user requests. This résumé provides some insight into the process of implementation of such a system and provides basic information on its main features.

Historical circumstances and the need for disaggregated data

In April 2006 the Czech National Bank (CNB) underwent a substantial organisational change. Three independent supervisory bodies were integrated into the CNB to create a single financial supervisor in the Czech Republic. This was an opportunity to start negotiations between statisticians and supervisors towards the implementation of a joint securities holdings data collection system. Such cooperation would lead to considerable synergies, including the avoidance of double reporting. Due to the need for flexible and high-quality data, a disaggregation trend prevailed. This was reflected in progress in the replacement of aggregated data sources by disaggregated ones. To be more specific, as of December 2006, 27.3% of the overall volume of cross-border assets was collected on a disaggregated (security-by-security) basis, but by December 2012 the percentage had increased to 99.8%.

Description of the securities holdings data collection system at the CNB

Starting in 2006, all securities data sources were implemented and/or gradually adjusted with respect to accommodating the needs of the different parties concerned. The most important stakeholders for these data sources at the CNB are the Balance of Payments Division, the Monetary and Financial Statistics Division, the

[1] Monetary and Statistics Department, Czech National Bank, milan.nejman@cnb.cz.

Financial Accounts and Economic Statistics Division and the divisions of the Financial Market Supervision Department. The individual securities holdings data sources, together with their basic characteristics, are presented in Table 1.

Securities holdings data sources at the CNB as of December 2012 Table 1

Respondent	Respondent ESA2010 sector	Data collection channel	Level of aggregation	Frequency	Timeliness
Non-financial corporations (securities held via nonresident custody or in self-custody)	S_11	direct reporting	s-b-s	Q	Q + 30 days
Monetary and financial institutions	S_121, S_122	direct reporting	s-b-s	M	M + 15 days
Investments funds	S_123, S_124	direct reporting	s-b-s	M	M + 15 days (MMF's), M + 25 days (IF's)
Non-bank custodians	S_126	direct reporting	aggregate	M	M + 20 days
Insurance companies	S_128	direct reporting	aggregate	M	M + 20 days
Insurance companies	S_128	direct reporting	s-b-s	Q	Q + 45 days
Pension funds	S_129	direct reporting	s-b-s	M	M + 15 days
Short-term bond system	S_121	indirect reporting	s-b-s	M	M + 3 days
Custodians (banks + non-banks)	S_122, S_126	indirect reporting	s-b-s	M	M + 21 days

It is apparent that most financial sector institutions are covered by monthly security-by-security (s-b-s) direct reporting. The first exception to this rule concerns the sector of insurance companies, where two partly overlapping data sources exist. A monthly source is used for portfolio investment purposes (compilation of cross-border assets), whereas a quarterly s-b-s source serves for supervisory purposes. Unification of the two data sources (shift to monthly frequency) could be solved with the start of Solvency II reporting. The second exception applies to the aggregated data source of non-bank custodians. Here, however, a change to s-b-s reporting is not planned due to marginal relevance (low volume of holdings). As to timeliness, for the data sources with monthly frequency the reporting deadline is 25 days after the end of reference month, while for quarterly frequency the deadline increases to 45 days.

The challenges ahead

Seven security-by-security data sources have been implemented within the Czech National Bank over a period of five years by different CNB divisions. One of the fundamental problems is a lack of consistency between these sources. The reports differ in:

- Data formats
- Reporting frequencies

- Reporting deadlines
- Methodology (*e.g.* valuation base: market value versus book value)

To solve the problem of different data formats and laborious manual processing of data, a project has been set up at the CNB to automatically convert all the separate security-by-security reports into one database. This database will serve as a basis for compiling a who-to-whom matrix at the national level. Furthermore, the database will be used for data reporting to the Securities Holdings Statistics Database (SHSDB), an ESCB-wide project with the objective of collecting s-b-s holdings by institutional sectors of euro area/EU reporting countries for both direct holdings and indirect holdings (third party holdings).

Conclusions

The statistical and supervisory bodies at the Czech National Bank have worked together to set up a common securities holdings data collection system that avoids double reporting and adapts flexibly to unforeseen data gaps. A disaggregation trend has prevailed and most securities data sources have been implemented on security-by-security basis and with monthly reporting frequency. The envisaged way forward is to set up a database where all the separate reports are automatically converted. This will become a basis for compiling a who-to-whom table and will be used for international reporting to the Securities Holdings Statistics Database (SHSDB).

References

- European Central Bank (2002). "Task Force on Portfolio Investment Collection Systems, Final Report", Frankfurt am Main.
- Nejman, M., Cejnar, O. & Slovik, P. (2011). "Improving the quality and flexibility of data collection from financial institutions", *Proceedings of the IFC Conference on "Initiatives to address data gaps revealed by the financial crisis, Basel, 25–26 August 2010*", No. 34, pp. 52–59.

Coping with challenges in establishing security-by-security database in Thailand

Wongwatcharapaiboon Weeraya[1]

Introduction

The series of abrupt global financial crisis during 1990s-2000s have spelled out demand for the more comprehensive data as to help early detect the imbalances which can potentially trigger instability in the financial system. Such data needs require the greater extension of data collection breadth and depth, and even calls for the data granularity down to the micro level.

Another aspect of data adequacy is calling upon the re-orientation of data design where greater flexibility is a key component. In this regard, data set collected should be sufficient for today needs and be flexible enough to respond to potential use in the future.

Developing Security-by-Security Database in Thailand

The Bank of Thailand decided to develop Financial Market Instrument system (FMI) by using security-by-security approach. The data requirement was focused only the benefit of the BOT internal users. The project had started off in 2006 with the limited scope to only debt securities, equity securities and derivatives trading in the equity market. In parallel, portfolio investment in Thailand data (liabilities side) which had formerly been collected via quarterly survey from custodians in aggregate format was redesigned to integrate with the new system. The new portfolio investment in Thailand data is acquired at the security-by-security and investor-by-investor level to accommodate various needs from BOT internal users.

This FMI project was scheduled to finish in December 2008 but due to the technical issues concerning system performance discovered during the user acceptance test, the delay of the project is unavoidable. Eventually, the BOT's security-by-security database was officially launched to production on September 1, 2009.

The FMI system comprises of two main components, the financial instrument reference data and the transaction and end-position data. The financial instrument reference data are pool of data acquired from various sources, namely BOT, SEC, TSD and ThaiBMA, as identified in Table 1.

[1] Statistics and Information Systems Department, Bank of Thailand, weerayaw@bot.or.th.

Sources of FMI Data

Table 1

Source of Data	Role	Financial Instrument Reference Data	Transaction and end-position data
Bank of Thailand (BOT)	Registrar	Government and state enterprises debt securities reference data (only securities BOT acts as the registrar)	New issuance and outstanding of government and state enterprises debt securities classified by holder groups -- Scrip only
Securities and Exchange Commission (SEC)	Regulator	Corporate debt securities reference data (only securities under the authority of SEC)	New Issuance of corporate debt securities (only securities under the authority of SEC)
Thailand Securities Depository (TSD)	Registrar	Equity securities reference data (only securities traded in the Stock Exchange of Thailand)	Outstanding of securities classified by Holder groups –Scripless
Thai Bond Market Association (ThaiBMA)	Information Center	All debt securities reference data (only securities registered at ThaiBMA)	Market price and debt securities trading summary
Stock Exchange of Thailand (SET)	Exchange Market	-	Market price and equity securities trading summary
Custodians	Custodian	-	Outstanding of securities classified by Holder groups Outstanding of Thai securities held by non-resident (Portfolio Investment in Thailand)

The transaction and end-position data are collected from multiple sources as indicated in Table 1. In order to ensure the coverage and accuracy of the outstanding classified by holders, consolidation among data sources are mandatory.

The position of portfolio investment in Thailand is derived from the data submitted by custodians at the investor level. Unfortunately, most of the data provided by the custodians still in the global custodian level not in the end investor level as desired.

Challenges

Challenges of S-B-S project development are the multi-fold complication. Since data are scattered among authoritative bodies holding different missions and goals which stem into the differences of data characteristics. The difficulties we faced due to these differences are as follows:

1. Data Consistency

Duplication of data and how to incorporate the data from various sources

The foundation of the security-by-security database is the financial instrument reference data. Therefore, every effort must be made to ensure the accuracy and the coverage of the integrated reference data. The problem is how we can identify the duplication of data between two sources in order to merge them correctly. The answer must rely on the unique identifier of security, ISIN code, but not in the case of Thailand. The ISIN code is not mandatory and still not widely used in Thailand.

The trading symbol is more popular and more preferable to reporting entities but is available only the securities traded in the exchange market or registered with the authorities or associated agencies.

Another reference data that always raises the issue is the Involved Party information. Lack of unique identification code of non-resident is the impediment for monitoring capital inflow, distinguishing the direct investment from portfolio investment data, and compiling comparable statistics.

2. Complex and difficult-to-maintain architecture of data acquisition and preparation system

Data quality management process is fundamental to all data management systems; however; in the security-by-security database, the issue is more complicated than others. The data quality management process must be done at the security level. The lack of security unique identification code makes this matter worse. The transaction and position data must be validated against the financial instrument reference data by using the unique identifier of the securities. However, since not all of the data has the unique identifier, other methods of validation must be introduced in order to cope with the data with no unique identifier such as short term debt securities.

The integration of outstanding is not a minor issue as well. To ensure the data consistency as previously mentioned above, the data preparation module must be equipped with statistical methods or pre-defined rules in order to automate the processes as much as possible which make the data preparation more complex.

Moreover, since most of the reclassification and calculation processes are done in-house, the lack of some parts of reference data might cause the problem. The data adjustment at the security-by-security level is essential to complete the data before compiling statistics. The problem is amplified with the goal of minimizing data timeliness which is quite conflict with the amount of works burdened from security-by-security data quality management.

Even though the processes may seem complicated and troublesome, for the data quality managers, the mistakes can be effortlessly identified and investigated because every record is validated at the security-by-security level. The quality of data can be guaranteed if the data are properly passed through all of the processes.

In summary, the design of the security-by-security data acquisition and data preparation architecture should consider a trade-off between costs and benefits, i.e. the complexity of the system and the degree of accuracy of data and/or the load of data preparation processes burdened on the data quality managers.

3. Selection of Business Intelligence tool and multi-dimensional model design

The extremely large amount of transaction data combined with a great number of dimensions to form the FMI data cubes might be the reason of the poor performance of our chosen OLAP tool which was discovered during the UAT and this leads to the inevitable change of the OLAP tool and the six months delay of the project.

From BOT's S-b-S Database to TFIIC

In the year 2010, the BOT working under the umbrella of Capital Market Development Committee had initiated the separate track, called Thailand Financial Instrument Information Centre (TFIIC), which is a joint collaboration of five institutions, namely the Bank of Thailand, the Public Debt Management Office, Ministry of Finance, the Office of the Securities and Exchange Commission, the Stock Exchange of Thailand and the Thai Bond Market Association. The concerted effort among the agencies who are involving in security life cycle from issuing, trading and holding of securities is the key success factor to launch TFIIC to production in 2013.

Identical to BOT's FMI, TFIIC consists of two main components. The financial instrument reference data is managed by SEC which is the authority responsible for the approval of new issuance for both equity and debt securities of resident corporations, either in the domestic or offshore markets. The transaction and end-position data are submitted by BOT which already has the FMI system in place.

Challenges of the Thailand Financial Instrument Information Centre (TFIIC) Development

The challenges of TFIIC are somewhat similar to the BOT's FMI but there are a few issues which far more complicated.

1. Balancing between organization benefits and public benefits

TFIIC is primarily intended for the benefit of public to build the information framework for security data exchange and dissemination. Fortunately, those five agencies who aware of the significance of the TFIIC show strong commitment and allocate sufficient budget and staffs. Despite of that, some benefits should be reciprocated to those devoted organizations. In this case, TFIIC provide the information exchange in the security-by-security level for the related agencies to fulfill their missions and goals.

2. Flexibility of the TFIIC data model

Another technical challenge the working team concerns the most is the flexibility of data model. Since the expectation of stakeholders is that the data model must be designed in a way that could support all types of securities or at least less painful to add new features due to the financial innovation.

Next Step of the BOT's Security-by-Security Database

In the year 2013, to make BOT security statistics more aligned with international standard, the issues such as scrip holding and security repurchase transaction will be taken into account to make an enhancement on BOT's Financial Market Instrument system (FMI). The security statistics on "from whom to whom" basis from

FMI will then benefit the compilation of other statistics, i.e. financial account, and international investment position, to be more comprehensive.

Conclusion

Security-by-security database is an outset attempt of micro data to answer to wider range of needs from monitoring capital inflows to monetary policy formulation. Other than the flexibility to cope with the rapidly changing data requirement, the security-by-security information system can reduce the burden of reporting entities, improve the data quality despite the cumbersome data quality assurance processes shifted from reporting entities to the compilers, and move forward the compilation modality.

Despite the challenges confronted throughout the project, BOT's security-by-security database and TFIIC are still considered to be worth investment for Thailand. The success of Thailand case is built upon the inter-agency co-operation which is the most crucial factor, the devotion of relating parties, and the endorsement of the Capital Market Development Committee.

References

Chaikhor, S. & Powattanasatien, P. (2010). "Financial Market Data Gaps Through Inter-Agency Collaboration", *BOT Stat Symposium*, August 18, 2010.

Garnjarerndee, P. "Debt securities statistics: the Bank of Thailand's experience", *IFC Bulletin No. 29*.

Micro-databases on securities statistics as a key issue for satisfying user needs

Luís Teles Dias[1]

1. Introduction: the dawn of item-by-item reporting system for statistics

Banco de Portugal started collecting data from financial institutions specifically for statistical purposes during the 70's. The data acquisition model relied on traditional aggregated reporting schemes, consisting typically of forms designed to answer pre-defined requirements. This model of reporting has been the standard for many years in the collection of data for statistical and supervisory purposes by *Banco de Portugal* and is still applicable for some very relevant statistics (*e.g.* Balance Sheet and Interest Rates of Monetary Financial Institutions, although in the latter case some partial developments in new directions have already taken place).

The first entire item-by-item reporting framework for statistical purposes developed by *Banco de Portugal* was the Securities Statistics Integrated System (SSIS) which was launched in 1999 (the system for collecting data to compile Balance of Payments statistics developed in 1993 had already a major component based on the report of individual operations with non-residents). The SSIS is managed by the Statistics Department.

2. The development of the Securities Statistics Integrated System

The SSIS was designed to store, manage and explore data on securities issues and portfolios on a security-by-security and investor-by-investor basis except for investors in the households sector, whose data are aggregated by the investor's country. The scope of financial instruments covered by the SSIS is securities other than shares (short- and long-term) and shares and other equity. The collected data include, on a monthly basis, stocks and transactions. The ISIN code is used for the identification of the securities.

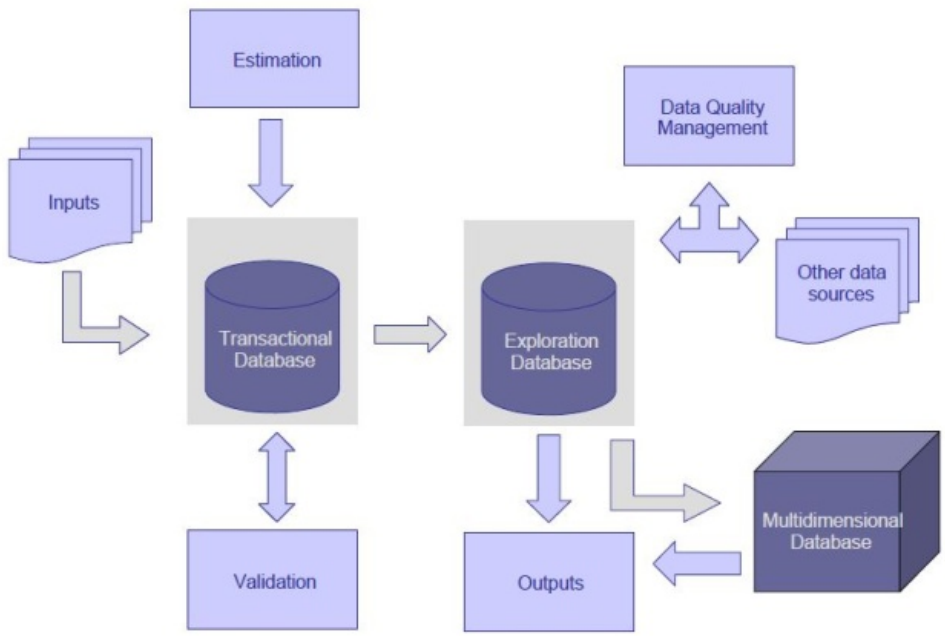
In the case of issues, the SSIS assembles data on securities issued by resident entities in Portugal, irrespectively of the fact that those issuances take place in the Portuguese market or in external markets. A variety of sources are used such as the Lisbon Stock Exchange, the Portuguese Securities Market Commission, the Portuguese Treasury and Debt Management Agency, commercial databases, etc. Concerning securities portfolios, detailed information is collected on holdings by residents of domestic and foreign securities, as well as on the holdings of domestic

[1] Statistics Department, Banco de Portugal, luis.teles.dias@bportugal.pt.

securities by non-resident investors. These data are reported mainly by custodians, which include banks, dealers and brokers. Direct reporting by resident investors with relevant portfolios deposited outside the Portuguese financial system is also applicable.

The architecture of the SSIS is based on three databases (two relational and one multidimensional). Collected data are stored and validated in the "transactional database". Data estimation of missing information is also done in this database. Validated and "enriched" data are copied daily to the "exploration database". A second level of quality control is performed on aggregate data, by carrying out consistency tests and comparisons with other information sources. Statistical outputs are produced from the "exploration database" as well as from the "multidimensional database". A graphical representation of this architecture is presented below.

The SSIS Architecture — Figure 1



The original purposes that inspired the development of the SSIS covering, in an integrated model, issues and holdings of securities were basically the following:

- the production of statistics on issues of securities by the resident sectors
- the production of statistics on securities portfolios of institutional sectors
- the development of "from-whom-to-whom" tables crossing issuers and holders
- as input data for other statistics, namely:
  - i. Monetary Financial Institutions (MFIs) and Investment Funds (IFs) statistics, allowing, for instance, the calculation of price revaluations of the portfolios of MFIs and IFs
  - ii. Balance of Payments and International Investment Position (BOP/IIP) statistics, contributing to the calculation of the Portuguese portfolio invested abroad and the foreign portfolio invested in Portugal

iii. Financial Accounts, regarding instruments F.3 (securities other than shares) and F.5 (shares and other equity)

In recent years the SSIS was "linked" with the Centralised Securities Database (CSDB) managed by the European System of Central Banks (ESCB), acting as a contributor but also as a receiver of the data stored in the CSDB. Presently, the ESCB is developing a Securities Holdings Statistics Database (SHSDB) for which the SSIS will be also involved in a two-way communication of data.

3. Flexibility: the main advantage of item-by-item reporting models

The main benefit of statistical compilation systems based on item-by-item reporting is their flexibility. This flexibility is manifold, either for respondents or compilers:

- it increases the ability of the system to deal with changes in the statistical requirements especially in cases where further details in existing breakdowns are needed (in most cases these situations don't imply any intervention in the reporting system);
- changes in the reporting scheme are easier to implement as they typically consist of additional granular items (new dimensions) that will not need to be transformed of aggregated by respondents;
- it prevents data redundancy, promoting in practice the principle that "data should be collected only once";
- it enables a more efficient data quality management; and, above all,
- it improves dramatically the responsiveness to *ad hoc* requests.

One important finding in economics over the 90's has been that the analysis of aggregate statistics does not give policy makers an accurate view of the functioning of the economy. Analysis of micro-data has become crucial allowing, for instance, analysts to calculate marginal, rather than average effects.

4. Some examples of the SSIS response to new demands

One of the most remarkable examples of the responsiveness of the SSIS happened in September 2008, one day after the announcement of the bankruptcy of Lehman Brothers. The ECB requested the Eurosystem to provide within the next 24 hours all available information regarding holdings of shares and debt of Lehman Brothers and subsidiaries with a set of details that included the country and sector breakdown (with further breakdowns for the financial sector to distinguish OFIs, IFs, FVCs, ICPFs and others) of holders. A list of relevant ISINs was distributed to facilitate the query. The request was obviously very urgent, sensitive and exceptional. In a couple of hours, Banco de Portugal was able to reply to the ECB providing the full answer to the request and even a number of relevant additional ISIN were identified and also provided. Fortunately, Portuguese institutions were minor exposed to Lehman Brothers. The security-by-security model of the SSIS and its high level of coverage (both for issues and holdings) were of paramount importance to address this important *ad hoc* request in such an effective way.

The financial crisis that affected a vast numbers of economies in the course of last 5 years was particularly severe in Portugal, leading the country to request international financial assistance from the EU/IMF in May 2011. Since then, the demand for higher detailed statistics, either on regular and also on *ad hoc* basis, has increased dramatically, namely regarding the financial sector, the public sector or the indebtedness situation of firms and households. It was a huge challenge for the Statistics Department of *Banco de Portugal* but very powerful tools were available, namely the three micro-data databases managed by the Department:

- The SSIS (already described)
- the Central Credit Register (CCR), containing information on an individual basis regarding the loans granted by the resident credit institutions to all resident legal and natural persons, based on a model very close to the loan-by-loan approach
- the Central Balance Sheet Database (CBSD) with the balance sheet and income statement of all non-financial corporations on an individual basis

The existence of these micro-data databases was a key success factor for an efficient and effective response to the myriad of new data requests for monitoring the development of the Financial Assistance Programme to Portugal and the preparation of the quarterly review missions.

The two following cases are particularly interesting examples of what has been said in the previous paragraph about the advantages of having statistical compilation systems based on item-by-item models and to build micro-data databases from that:

- The production of "amortization plans" for banks, general government and state-owned enterprises (SOEs), with the amortisation/redemption amounts, on an annual basis until 2020 and from 2021 onwards, of debt securities, domestic loans and external loans. The contributions for this output were from the SSIS, the CCR and BOP/IIP, respectively;
- The production of a new statistical product called "non-financial sector indebtedness" aggregating the amounts of non-consolidated debt resulting from issues of debt securities (held by residents and non-residents), domestic loans, external loans and trade credits. Data are presented combining and crossing different dimensions of analysis, namely: debtor and creditor institutional sectors, type of instrument, original maturity, sector of economic activity and size of companies. This output was the result of an integrated approach of data (largely micro-data) from the SSIS, Monetary and Financial Statistics, BOP/IIP, Financial Accounts, CCR and CBSD. It was the first time that *Banco de Portugal* has used an integrated approach with such a high number of different statistical domains. The result was an innovative achievement at international level and led to a new chapter in the monthly Statistical Bulletin.

An excerpt from the "non-financial sector indebtedness", crossing the debtor and creditor sectors

Figure 2

K.1.3 DÍVIDA[1], POR SETOR DEVEDOR E FINANCIADOR
Posições em fim de período
Valores não consolidados

DEBT[1], BY DEBTOR AND CREDITOR SECTORS
End-of-period positions
Non-consolidated values

Fonte/Source: Banco de Portugal

10^6euros

		Dez 09	Dez 10	Dez 11	Dez 12	Fev 13	Mar 13	
		Dec 09	Dec 10	Dec 11	Dec 12	Feb 13	Mar 13	
		1	2	3	4	5	6	
1	Administrações públicas	158 591	185 844	214 378	241 085	244 713	244 764	General government
	Financiado por:							Financed by:
2	Administrações públicas	14 384	19 477	25 598	32 927	32 248	32 659	General government
3	Setor financeiro[2]	23 943	46 518	50 243	58 314	58 035	57 953	Financial sector[2]
4	Empresas	4 313	6 261	6 643	6 776	6 589	6 756	Corporations
5	Particulares[3]	10 714	10 447	8 964	7 898	7 960	7 972	Private individuals[3]
6	Exterior	105 238	103 140	122 930	135 170	139 881	139 424	External
7=[13]+[19]	Empresas públicas	37 919	45 064	48 846	45 888	46 863	47 317	Public corporations
	Financiado por:							Financed by:
8	Administrações públicas	780	675	5 409	9 062	9 302	9 765	General government
9	Setor financeiro[2]	10 228	17 312	15 359	13 709	14 465	14 570	Financial sector[2]
10	Empresas	5 463	5 720	6 871	5 299	5 316	5 322	Corporations
11	Particulares[3]	136	134	145	143	145	152	Private individuals[3]
12	Exterior	21 312	21 223	21 062	17 676	17 634	17 509	External
13	Empresas públicas incluídas nas administrações públicas[4]	14 939	20 763	23 290	25 225	25 474	25 887	Public corporations included in the general government[2]
	Financiado por:							Financed by:
14	Administrações públicas	135	134	5 267	8 918	9 167	9 629	General government
15	Setor financeiro[2]	3 697	9 462	7 034	6 190	6 226	6 239	Financial sector[2]
16	Empresas	302	232	117	122	136	141	Corporations
17	Particulares[3]	0	1	2	17	27	32	Private individuals[3]
18	Exterior	10 804	10 934	10 870	9 979	9 919	9 845	External

5. Conclusion

The three examples presented in this paper showed that the SSIS, combining data on issues and holdings of securities, having both components based on an item-by-item model (security-by-security and investor-by-investor), was a key issue for satisfying users' needs overtime.

It is important to recall that the SSIS was launched in 1999, after a period of 2 years of development. The users' needs for statistics regarding securities have changed a lot during these years, in particular since May 2011 when the Financial Assistance Programme to Portugal began, but the SSIS remained unchanged since its inception (only a few technological improvements were made). In particular, the option of adopting a security-by-security and investor-by-investor model for holdings of securities was a very innovative solution in 1999 and proved to be the right option to introduce the maximum degree of flexibility in the system and to extend its longevity maintaining its relevance. The SSIS is presently under reformulation, triggered by technological constrains and by the need to increase the automation of a number of procedures more related to the collection of data on issuances. The model and the data requirements will remain almost unchanged.

The development of compilation systems for statistical purposes based on micro-data is a conscious strategy that *Banco de Portugal* has been following over the past years. Apart from securities, other micro-data databases where developed for loans, interest rates and individual accounts of non-financial corporations or are under study (*e.g.* for deposits). From the respondents point of view the strategy of requesting granular data has been clearly perceived as a way to reduce their reporting burden and to simplify their reporting procedures. For the compiler, it's a

strong investment in maximizing flexibility although incorporating higher costs in terms of data storage, computer processing and data quality management.

The use of micro-data databases and item-by-item reporting by the Statistics Department of *Banco de Portugal* has been absolutely vital to ensure a level of excellence of its statistical production. The ability to deal with the multiple new statistical demands that arose from the Financial Assistance Programme was a perfect example of the appropriateness of that strategy and a strong stimulus to pursuit it even deeper in the future.

Building business registers to monitor entrepreneurial dynamics

Homero Gonçalves[1] and Mário Lourenço[2]

Overall context

The in depth knowledge of the economic agents' characteristics within a certain economy has always been one of the goals of economic and financial statistics. Given the relevance of non-financial corporations (NFC) contribution to Portuguese economy (roughly 56% of Portugal's gross value added, in 2012) it is vital to have comprehensive, reliable and detailed information on this institutional sector.

This paper discusses how the integration of business registers and other administrative sources of information enabled the Statistics Department of the Banco de Portugal to create a single enterprise database with annual information on virtually all non-financial corporations, while providing some examples of the analysis this data, by itself, allows, namely in what regards the monitoring of entrepreneurial dynamics.

Framework

Banco de Portugal has developed and manages several micro-data databases (e.g. Central Credit Register, Securities Statistics Integrated System, Central Balance-Sheet Database) covering many features of NFC's activity. However integrated information on each firm's individual characteristics was not easily available (e.g., size, NACE, age, legal nature), with most of this information being scattered across numerous data sources and organizations.

Without a single statistical units file joining all the sources deemed relevant, it was possible for different users to consider different classifications with negative consequences on the comparability and reliability of official statistics, namely regarding its harmonization.

In order to overcome these limitations, the Banco de Portugal developed a new business register that combines information from several databases making use of a considerable number of administrative and statistical sources which provided information on firms' individual features.

Identifying the available information in order to determine how to characterise Portugal's NFC was the first challenge. Prior experiences in cooperating with other entities (such as IES – Simplified Corporate Information, a breakthrough which enabled Portuguese enterprises to single-handedly report their annual accounts to

[1] Statistics Department, Banco de Portugal, hgoncalves@bportugal.pt.

[2] Statistics Department, Banco de Portugal, mfllourenco@bportugal.pt.

several institutions; or SICAE, a single channel through which the companies are categorized via NACE codes according to their economic activity) gave way to a better knowledge of the strengths and weaknesses of each kind of report and database.

The following step was to determine which should be the sources deemed more relevant in each case, determining a hierarchy of sources. For instance, IES was considered a priority source in all accounting related characteristics, such as the level of turnover and assets, as well as the number of employees (and, consequently, the enterprise size which is derived from these items). The Ministry of Justice, responsible for SICAE's management and for the legal registries in Portugal, was deemed the source with the highest priority in what relates to the geographic location of the enterprise, its legal nature, and its economic activity. Other data sources, like Statistics Portugal, would then add to this information in cases where it was missing.

The implementation of this hierarchy depended, nevertheless, on the availability of a single key which enabled the correspondence between the several sources of information. Given that in Portugal all firms are identified by a unique fiscal number, the choice of the primary key was simple. The final outcome was the implementation of an integrated database, automatically and constantly updated according to the most relevant and recent information, for a time-frame that spans from 1991 onward (reconstructing more than 20 different yearly populations estimates of NFC operating in Portugal).

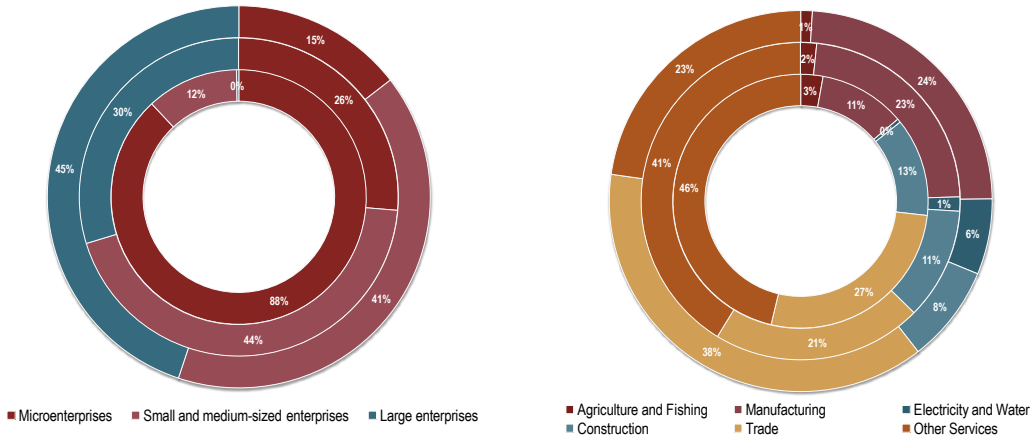
Some applications on monitoring entrepreneurial dynamics

The integration of the above mentioned data sources and the fact that they are automatically updated enables the users to analyse the structure and track the dynamics of Portugal's NFC sector. Figure 1 shows this sector's structure by size and economic activity in 2011, based on the number of enterprises, number of employees and turnover. It is possible to notice the significant weight of microenterprises in number of firms (88%) and the fact that large companies are the most relevant when turnover is considered (45%, although they only represent 0.3% of the firms). On the other hand, by economic activity, the weight of *Other Services* when considering the number of enterprises (46%) is partially compensated by the greater relevance of *Manufacturing* when considered the turnover of Portugal's NFC (24%).

Structure of Portugal's non-financial corporations (2011)

By enterprise size (left) and economic activity (right), based on the number of enterprises (inside), number of employees (middle) and turnover (outside)

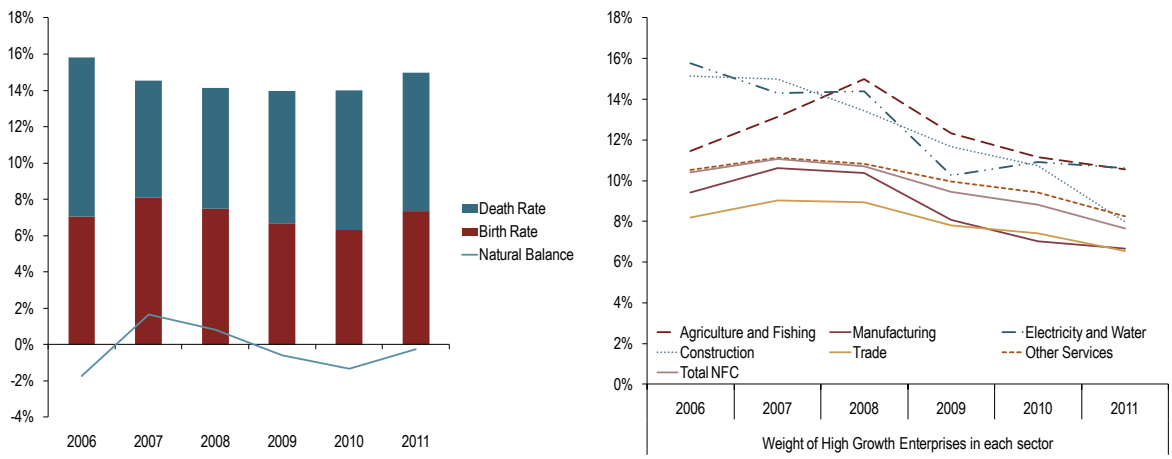
Figure 1



The fact that this information is organize as time-series enables the analysis of the dynamics of Portugal's NFC, namely through the computation of birth and death rates or by using the average variation in turnover to identify High Growth Enterprises (HGE) (Figure 2). The available data reveals negative natural balances in the demographics of Portugal's NFC over the last three years (-0.3% in 2011), as well as a reduction in the weight of HGE (from 11% of the total enterprises in 2008 to 8% in 2011). *Agriculture and Fishing* and *Electricity and Water* gather the largest proportion of HGE in 2011(11% of its enterprises), while *Construction* has registered the largest decrease in this number over the last five years (from 15% in 2006 to 8% in 2011).

Portugal's non-financial corporations' dynamics

Figure 2



Conclusion

This new micro-data database highlights the benefits of inter-institutional cooperation in statistics. Indeed, gathering data from different official sources enabled the collection of a new dataset that not only allows a better use of other micro-data, assuring their harmonization and an easy cross-reference between datasets (single identifier and single attributes per enterprise allow more straightforward comparisons and stable results across several fields of statistical data and outputs), but also, by itself, the monitoring of business and entrepreneurial dynamics.

References

Banco de Portugal (2010), "Structure and Dynamics of Non-Financial Corporations in Portugal", *Central Balance-Sheet Studies no. 2*, December 2010.

Banco de Portugal (2012), "Sectoral Analysis of Non-Financial Corporations in Portugal 2011/2012", *Central Balance-Sheet Studies no. 8*, November 2012.

EUROSTAT/OECD (2008), *Manual on Business Demography Statistics – 2007*.

Gonçalves, Homero (2011), "How can administrative databases help us understand the funding behaviour of non-financial corporations?", *Irving Fisher Committee Bulletin no. 34*, pp. 295–320, November 2011.

Using registers as the basis to statistical production – linking issues and potential bottlenecks

Jaanus Kroon[1] and Ain Paas[2]

Introduction and background

Central bank statisticians must deliver relevant and reliable data to meet growing user 'needs with better quality and quicker at the same time avoiding increases in reporting burden and budget. The scope of the central bank statistics has been reached more far of its "traditional". Keeping statistical production internally coherent and consistent is getting as a challenge when measuring an increasing number of phenomena at increasingly more detailed level.

The key to deal with the situation is the widespread use of administrative registers and records, interlinking these sources to everyday work process. It needs public approval, promoting legal base, unified identification code system across registers, reliable register records and cooperation among administrative authorities and statisticians.

According to the Official Statistics Act, the Eesti Pank (Bank of Estonia; NCB), the other main producer of official statistics besides Statistics Estonia (NSI), shall primarily use data collected in administrative records and databases, data generated in the course of the activities government authorities and legal persons or collected by them, if such data allow the production of official statistics complying with the quality criteria of official statistics.

Currently there are the following major administrative registers involved into statistical process of the Eesti Pank:

Central bank statisticians must deliver relevant and reliable data to meet growing user's needs with better quality and quicker at the same time avoiding increases in reporting burden and budget. The scope of the central bank statistics has been reached more far of its "traditional". Keeping statistical production internally coherent and consistent is getting as a challenge when measuring an increasing number of phenomena at increasingly more detailed level.

The key to deal with the situation is the widespread use of administrative registers and records, interlinking these sources to everyday work process. It needs public approval, promoting legal base, unified identification code system across registers, reliable register records and cooperation among administrative authorities and statisticians.

According to the Official Statistics Act, the Eesti Pank (Bank of Estonia; NCB), the other main producer of official statistics besides Statistics Estonia (NSI), shall primarily use data collected in administrative records and databases, data generated

[1] Statistics Department, Eesti Pank, jaanus.kroon@eestipank.ee.

[2] Statistics Department, Eesti Pank, ain.paas@eestipank.ee.

in the course of the activities government authorities and legal persons or collected by them, if such data allow the production of official statistics complying with the quality criteria of official statistics. Currently there are the following major administrative registers involved into statistical process of the Eesti Pank:

Register	*Source*	*Domain*	*Field*
Business and Commercial Register	Centre of Registers / Ministry of Justice	- Statistical base register - External Sector Statistics - Financial Account Statistics - CSDB/SEC statistics - HFCS (Household Financial and Consumption Survey)	- Survey frames of non-financial entities - Legal ownership (FDI and FATS) - Equity and reserves - Profits / reinvested earnings - Balance sheets for financial accounts - NACE/institutional sector
Register of licensed supervised entities	FSA of Estonia	- Statistical base register - RIAD	- Reporting frames of financial entities - NACE/institutional sector
Estonian Central Register of Securities	Tallinn Stock Exchange /NASDAQ OMX group	- Monetary and Financial Statistics - External Sector Statistics - Financial Account Statistics - CSDB/SEC statistics - HFCS	- Transactions and stocks of securities by investors and issuers, registered in Estonia
Information System of State Accountancy	Ministry of Finance/ State Shared Service Centre	- External Sector Statistics - Financial Account Statistics	- Central and local government related financial/non-financial flows and stocks
Tax Register	Tax and Customs Board	- External Sector Statistics - HFCS	- Items related to wages and income - Survey frames of non-financial entities (VAT data)
Land Register	Estonian Land Board	- External Sector Statistics	- Prices and transactions of real estate (FDI)
Real Estate Register	Centre of Registers /Ministry of Justice	- External Sector Statistics - HFCS	- Real estate stock (FDI, household property)
Traffic register	Estonian Road Administration	- HFCS	- Vehicles owned by households

The aim of this note is to briefly describe the current practice and role of *Statistical base register* in statistical production process of Eesti Pank, highlighting some linking issues and bottle-necks related to other mentioned data sets.

Current practice

The *Statistical base register* of the Eesti Pank is the cornerstone of statistical production of all domains of statistics the bank is responsible, established 1995. In the phase of creation and development of the register it was initially mostly a copy of administrative register of legal entities, provided by Ministry of Justice, supplemented by survey management system for direct reporting. Currently, close links to Statistical Register of Entities of Statistics Estonia having an essential role to supplement the register, promoted by co-operation and data exchange between NCB and NSI.

As mentioned, the most important administrative data source is *Business and Commercial Register* which is an online public service offered by the Centre of Registers and Information Systems (RIK) and the register is based on the central database of registration departments of the courts. This central database includes digital data of the commercial register, the register of non-profit associations and foundations and the commercial pledge register. The register contains historical data since 1995. Register related Company Registration Portal is an internet environment that enables electronic submission of documents to the Department of Courts without the mediation of a notary. Through the portal it is possible:

- to register a private limited company, sole traders, general and limited partnerships within 2 hours,
- to change the data in the commercial register,
- to file annual reports, according to the taxonomy (XBRL format),
- to submit notices about the changes in the main field of activity,
- to submit changes in the telecommunication numbers,
- to submit changes in the list of the members of the supervisory board and/or in the management board and in the list of auditors.

The basic units of the register are the legal units with the Business Identity Code (BIC) as the anchor identifier for all other statistical and administrative registers. The register consists of visualised part showing current and former relations between legal persons (BIC) and private individuals (identifier Private ID code). All previously mentioned is important from statistical point of view - the environment of legal units is extremely flexible and dynamic. The Eesti Pank updates statistical register on the bases of legal register once a month.

Another important data-source for the register of Eesti Pank is *Statistical Register of Entities* (SRE) of Statistics Estonia. The co-operation between two providers of official statistics in this field started in 2007, due to widened statistical field (FATS) and has increased in recent years. The SRE provides annually the Eesti Pank with detailed data on units in the new frame (frozen in November each year) – units' adjusted contact data, principal activity, status of activity and institutional sector code. During the year data on changes in the frame are communicated to the bank. The bank also receives data on members of enterprise groups (incl. UCIs). Eesti Pank as a producer of outward FATS in turn provides to SRE with data on ascending chain subsidiaries of Estonian enterprises located abroad. Regular statistical frame survey of SRE takes into account Eesti Pank needs.

Previously mentioned data sources are supplemented by data driven from the collection system (contact information and other relevant information) and from the register of licensed supervised entities provided by Financial Supervisory Authority (the list of investment funds).

The major duties of *Statistical base register* of the Eesti Pank are:

- to be a survey frame and tool for data collection,
- to be a data source to identify and classify macroeconomic activities and sectors of legal entities,
- to be a direct data source for FDI and FATS statistics regarding ownership and annual economic indicators (pre filled-in questionnaires with combination of securities register data),

- to be indirect data source for cross-checking, comparing and validating reported micro-data,
- it is foreseen that the register will be linked to statistical reporting portal (*eAruanne*) to identify board members of entities, having direct access to determine persons responsible for statistical reporting (identifier: Private ID Code).

Lessons learned

A fundamental concept in the use of registers is the *common code system*, such as the business and private identity codes is the key to extensive cross-use of all above mentioned administrative registers, enabling to link all relevant fields for statistical production in an efficient way. This also enabled many improvements in the coherence and consistency of macroeconomic and business statistics, using the same industrial and sector classification for individual entities in different domains, avoiding an increase in reporting burden.

Registers can help to improve the quality of statistics but need *remarkable investments* into data-base systems, maintenance and statisticians the more so because higher requirements need to be followed in the processing of personal data. Possible cost-savings can be expected hopefully in the longer term.

As the Statistical base register has different input data feeds, there is always need to find a *good compromise regarding the timing of synchronisation*. Business and Commercial Register is live (even regarding annual reporting) while the Statistics Estonia updates SRE frames annually or Institutional Sectors biannually. In Estonia, small and open economy with flexible registration of new businesses, even a single uncovered and/or incorrectly classified entity may become as a big player easily, distorting statistical output.

Although the major responsibility in classification of units by Institutional Sectors according to the ESA is borne by NSI, *the co-operation in determining sectors plays a key role in achieving statistical coherence*. As the NACE code is not always treated as legally binding field of Business and Commercial Register (the initial data source), the special decision trees were made-up in co-operation of NCB, NSI and RIK to help operators to fix or update of activity codes. This is particularly important and difficult for financial activities. The classification rules, based on NACE, are regularly discussed and amended in co-operation with NSI and NCB. Even though using the rules the institutional sector of over 99% of units can be determined automatically, the exceptions demand individual approach. The final lists of financial sector and government sector entities are validated in co-operation with NSI and NCB as well. Although there is good synergy between two organisations, the main bottleneck has been the limited resource.

The main challenge is *how to ensure the correct, stable and consistent classification of institutional units/counterparties in all registers and reports*. At the moment there are only lists of financial sector and government sector entities available for reporters. It is essential to disseminate regularly BIC codes with validated NACE and Institutional Sector Codes to key reporters (credit institutions). The preparatory work for that is already started.

References

"Using Administrative and Secondary Sources for official Statistics". A Handbook of Principles and Practices. United Nations, New York and Geneva, 2011.

"Register-based statistics in the Nordic countries". Review of best practices with focus on population and social statistics. United Nations, Geneva, 2007.

"Use of Registers and Administrative Data Sources for Statistical Purposes". Best Practices of Statistics Finland. Helsinki, 2004.

Legal Entity Identifier or the utility approach to shared, high-quality basic micro-data

Francis Gross[1]

Containers, bar code, global trade: does technology or policy drive change?

A recent study[2] demonstrates that the standardisation of the shipping container contributed significantly more to globalisation than all trade agreements taken together. The radical reduction of transportation cost brought by the shipping container created a powerful drive for whole industries to seek the benefits of low production cost in faraway, then poor countries and enabled those countries to at last draw benefit from their poverty in the form of competitive advantage and grow, some of them to now being wealthy. Trade agreements merely removed barriers to that drive, but they usually came behind technology.

The advent of the bar code pushed that development further, enabling international trade to develop sophisticated global, just-in-time supply chains for the production of sophisticated goods, adding to international trade in raw materials and finished goods a complex web of trade in semi-finished goods and services in which a single component can travel through many countries before the finished product is goes into is delivered to a customer. Bar code, combined with information technology and cheap transport combined to profoundly change the structures of whole industries in very few years. Apple doesn't manufacture a single Ipad; retail is now dominated by global giants whose global data-integration creates dominant buying and marketing power.

How do globalisation and the digital age affect statistics?

Statisticians can be seen as "measurement engineers", middlemen who collect information about the "real" world they measure and produce information about that world for users who act in and on that world.

Statisticians put values on categories (e.g. exports, GDP, IIP) defined by users on the basis, for instance, of economic theory, that builds the models, mental or computer-based, with which the users structure their analysis and action. Those categories tend to be relatively stable over time, giving the long time series so prized by users. Limited by technology and driven by the need to keep cost low, statisticians have become used to collecting data specifically tailored to the

1 Directorate General Statistics, European Central Bank, francis.gross@ecb.europa.eu.

2 Quoted in the article: http://www.economist.com/news/...-trade-humble.

categories they are asked to measure. The system that results is highly stable, and more so since it is largely anchored in legislation, which is slow to change.

In a context of fast, deep and global change, categories used by statisticians and their users, and/or data collection schemes serving them can become obsolete or irrelevant, see settlement data for Balance of Payments. In such times, new categories emerge and trigger new data requirements. Such changes can be resisted by force of habit or be addressed through proxies based on strong assumptions, due to lack of resources and skills to obtain adequate data.

Globalisation and the advent of the digital age have de facto united our world by creating a dense web of long range, real-time economic relationships, which are changing the dynamics of the economy and finance in every single country and community. That change continues at high speed and is likely to continue for many years. The increasing complexity and velocity of developments is accompanied by disaggregation and de-integration of previously large, easy to understand economic units. For instance, large companies can now be clouds of thousands of legal entities without a real "home base" or networks of changing, competing suppliers. The world has become more "granular" with many more elements interacting globally, and ever less visibly, at high speed.

In physical systems, what ensues form such developments is turbulence. As an example, whereas a laminar flow of air around a wing can be described and predicted with a few simple equations and a few data elements (air temperature and pressure, speed, attitude), a turbulent flow around the same wing requires high performance computing and thousands of measurement points at high frequency. That also means that measurement and analysis made for laminar flows fail at precisely the moment where the behaviour of the wing becomes critical. The analogy comes back to statistics if, with the wing moving very gradually towards a turbulent regime, measurement is adapted a just little and complemented with a few additional observations, the pilot adding ever more instinct to the instruments. Until the plane stalls, that is, and then flying becomes a very different exercise.

The financial crisis has been the stalling of a good part of financial statistics. In a very short period of time, the financial system has become very turbulent, affecting the real economy and society, and that our measurement instrument was not up to the challenge at the worst moment. Luckily, turbulence has been contained, but we now know that we have a very different system from what our measurement instrument was built for. The system has gone past a threshold, into a new phase where it can behave in a very different and dangerous mode. Some parts of statistics need to change fundamentally to deliver what is needed: fast measurement that works in turbulence.

How can we know where and when an object thrown up in the air will hit the ground?

- ***Rock***: simple physics and measurement allow predicting precisely place and time of impact.
- ***Gravel***: assume that gravel is a rock and apply (1). Roughly right – some spread in space.
- ***Sand***: assume that sand is gravel, just a little finer, and apply (2). Still roughly right – more spread in space and now also time.
- ***Powder***: assume that powder is sand, just a little finer, and apply (3). It becomes messy.
- ***Nanopowder***: assume that nanopowder is powder, just a little finer, apply (4)…

… and find out that nanopowder floats away in the air, poisons your water, short-circuits your PC and penetrates your lungs where it causes cancer. The system has crossed a threshold, the old assumptions can no longer be applied, even as approximations. ***We need new thinking***.

Measuring a complex, fast, granular, turbulent, global, digital financial system

Users of statistics must be able to respond to unforeseen, hard to predict, rapidly changing and potentially dangerous situations. Faced with turbulence or crisis, they will need new statistical products, tailored to the situation, very close to real time, most of which cannot be delivered by the traditional statistical processes. We need new thinking.

Some communities have crossed the threshold to new thinking in measurement, such as meteorologists with their satellites and large scale computer models. Some aspects:

- Data is collected in very large amounts, permanently.
- Data covers the whole relevant system, for weather the Earth, whereby granularity can vary locally.
- Data is highly standardised.
- Data granularity is increasing with each new generation of satellites, increasing the usefulness of the output, e.g. longer, more precise weather forecasts.
- Data collection has become very cheap, through technology and economies of scale.
- Different data sources (weather, geology, etc.) use the same geographical coordinates, enabling their combination to produce new services for the public and the private sector alike, e.g. agriculture, conservation, etc.
- Data is analysed and transformed into useful information through very large scale computing facilities and, still.
- Output such as weather forecasts come very near real time.
- These processes are facilitated by technically- and scientifically-minded communities that cooperate globally.

However, economy and finance are very different from weather. Weather is a natural system whereas economy and finance are entirely man-made. At the micro-level, the molecules that make up weather follow fairly well-known rules of physics

and chemistry, whereas economy and finance are at the mercy of human psychology, at individual and mass level.

Also, it has taken decades for global weather scientists and governments to build these systems, both technical and organisational, and to acquire the funding and the political support they now enjoy.

Weather therefore doesn't offer an easy blueprint for a quick solution for the measurement challenges faced by measurement in economy and finance, but it inspires a few thought that can help to shape our own new thinking:

- Solutions will emerge and evolve progressively, probably over decades.
- Solutions, or at least components, will need to be at the scale of the system, *i.e.* global.
- Solutions will require many communities to cooperate; that demands learning from all.
- A common vision can unite these communities around a shared goal and guide our work.
- We must begin with an ambitious, concrete project that benefits all within a finite time.
- That first project should also have transformational power to ease further progress.

Good, cheap basic data in a global, public Utility: a possible first project

Making basic micro-data for economy and finance good and cheap by building a global Utility as a public good could be such a first project. Basic micro-data is understood as identifiers and reference attributes allowing to identify and describe objects of relevance, for instance a legal entity, a financial instrument or a contract (share, bond, loan, etc.), links among them and changes to them. Basic micro-data can be seen as equivalent to the basic data collected by weather satellites or to bar-code data in supermarket supply chain: it is the layer of data that connects the real world and the data world.

Today, good basic data is so expensive, or even impossible, that many products or processes that need it become inefficient and mediocre, or even gradually ineffective. So, in that context, what would make good basic data? Has data become worse with time? No. The same data that was good fifteen years ago can be bad now, because:

- To be good in a global system, data must be globally standardised, fit for IT;
- To be good in a fast system, data must be real time;
- To be good in a complex system, data must be granular.

Basic financial data must now be all three at the same time: the definition of good has changed, crossing a threshold where the old methods don't work anymore. Just as nanopowder is not a rock, just ground finer. We need new thinking.

The Legal Entity Identifier (LEI)

Luckily, in the field of basic micro-data for finance, all sectors, private and public, facing similar challenges and the same obstacles, have shown that they are ready for new thinking. The experience of developing the initiative for a Global Legal Entity Identifier System (GLEIS), initially for the financial sector, has delivered global agreement that there can only be one Utility. In the words of Jean-Claude Trichet, former ECB President: "Each one needs it. There can only be one. No one can do it for themselves. So we must do it together."[3] Industry and authorities, across all their components, have been cooperating towards a global solution, which is now being implemented under a G20 Charter. The GLEIS will be worth billions to industry through cost savings and better risk control, and invaluable for authorities who couldn't deliver many new functions without. And it has the transformational power to facilitate further steps.

Currently, the Regulatory Oversight Committee (ROC), a global body of now 53 public sector institutions established under a G20-endorsed Charter, is preparing the creation of the central building blocks of the GLEIS, namely:

- The Global LEI Foundation, based in Switzerland (GLEIF)
- The private sector, 15 to 20 strong Board of Director of the GLEIF (BoD)
- The Central Operating Unit (COU), controlled by the BoD / the GLEIF.

The COU will be a start-up and the hub of a network of Local Operating Units (LOUs), that will conduct the registration, validation and other operations of the GLEIS. Beginning from the current "pre-LEI" implementation serving OTC reporting in the USA and the EU, the GLEIS has vocation to grow universal and global. One can hope that it would be successful and ultimately evolve beyond covering entities to become a "Financial Object Identifier".

Statistics and the GLEIS: what strategy for statisticians?

Statisticians will be well-served by the GLEIS as an infrastructure in many respects. However, the GLEIS will reach its potential only if it becomes the operational data infrastructure of the financial industry, worldwide, hence helping industry to solve the data quality problem that costs it so dearly.

Statisticians are by definition well placed to help design standards for basic reference data. They also have an interest to do all they can to help make reporting cheap, large scale, high quality, high frequency, high timeliness as well as highly flexible and easily extensible.

Hence, statisticians should adopt a positioning extending their strategic remit upstream along their supply chain. There they should actively influence the processes in and around the private sector industries, including legislation where they have expertise too, to lead to effective global adoption of data standards in the shape of a global Reference Data Utility as a shared strategic infrastructure, of which the GLEIS will be the first building block.

[3] Said in a meeting with EU leaders at the ECB on 10 Sep 2010.

Is it legitimate for statisticians to influence the world they are supposed to measure? Clearly yes. Indeed, not doing so would see statisticians have a much worse influence on that same world, simply by delivering sub-standard products in fields so critical to the safety of our society.

We need new thinking.

Efficient ways of dealing with accounting data from enterprises

Margarida Brites[1]

Introduction

Accounting data of non-financial corporations can play an important role in the fulfilment of the mission and tasks entrusted to national central banks. The "language of business", as is known, reflects all aspects of a business' financial performance; accounting documents, thus, are essential to reaching an accurate picture of a business's true financial well being. Armed with such knowledge, central banks can make accurate economic and financial analysis about non-financial corporations, assess financial stability properly, take the right decisions in the field of prudential supervision and, last but not least, can compile and disseminate better statistics and appropriate indicators about non-financial corporations.

Based on the experience of *Banco de Portugal*, this paper aims to illustrate the advantages and potential uses of corporate accounting data by national central banks, namely in statistical domains.

The central balance sheet database

Banco de Portugal runs a Central Balance Sheet Database (CBSD) since 1983, based on accounting data of individual firms. As from 2006, annual CBSD data has improved significantly and has been based on mandatory financial statements reported in fulfilment of firms' statutory obligations, namely with the Tax Authority, the Ministry of Justice, *Banco de Portugal* and Statistics Portugal. This is the result of close cooperation among the four above mentioned public entities, which, seeking a reduction in corporations' reporting burden, have decided to integrate in a single format their demands on annual business accounting data, to be submitted by companies only once a year. The statement, IES – *Informação Empresarial Simplicada*, is submitted by each company with a delay of about seven months after the reference period.

The CBSD has been managed by the Statistics Department since 1999 and, presently, manages two databases:

- the annual database has annual accounting data of nearly all non-financial companies in Portugal (about 370 thousand companies) and contains all IES statements that have been submitted; this database has very much detailed data (more than 3 thousand items);

[1] Statistics Department, Banco de Portugal, mpbrites@bportugal.pt.

- the quarterly database has quarterly accounting data from around 3,500 corporations/year, representing almost 45% of the total turnover of the non-financial corporate sector. It comprises the answers to a quarterly survey conducted jointly by Statistics Portugal and *Banco de Portugal*; this database covers about 70 items.

The quality issue

Given the strategic importance of this type of data, the *Banco de Portugal* has developed a very thorough set of procedures to ensure a high level of quality of the accounting data, as well as its consistency with other micro data collected by the Statistics Department.

Following a first level of quality checks that take place when companies submit their accounting data, *Banco de Portugal* carries out further quality controls to find out basically missing data and misclassifications in the reported data. In this domain, it has been crucial cross-checking CBSD data with other statistical data, namely from the Central Credit Register, the Securities Statistics Integrated System and the Balance of Payments / International Investment Position databases.

Besides, other corrections are undertaken with the aim to adjust reported data to statistical standards and common compilation criteria. This is done because the alternative solution (data adjusted from the reporters' side) could be worst and give rise to greater and uncontrolled errors.

In sum, control checks are done at different levels of data granularity, including the individual one so that data concerning the same firm are, in the end, coherent and comparable among the different statistical domains and regardless the reporting agent. Furthermore, micro data cross-checking noticeably boost data value and enables further exploitation of interlinks among different databases.

Accounting data and Statistics

The main objective of the CBSD is to contribute to a better understanding of the activity and situation of non-financial corporations. So far, this has been largely accomplished through the production of statistics. Overall, CBSD can be particularly useful in the following statistical domains:

- to produce a comprehensive set of statistics about non-financial corporations. E.g., Chapters A and G of the Statistical Bulletin of *Banco de Portugal*;
- to derive non-financial corporation sector for National Accounts, both financial and non-financial. In Portugal, CBSD data has been used for compiling trade credits, own funds, inter-company loans, pension funds, and loans granted by private shareholders;
- to estimate several items for Balance of Payments / International Investment Position. Portuguese CBSD main contributions cover external trade in services, trade credits, direct investment, loans granted by foreign credit institutions, inter-company loans;

- to produce sectoral benchmarks. A great number of Sector Tables is produced by *Banco de Portugal*, with economic and financial indicators for non-financial corporations aggregated by sector of activity and/or size class. These tables are made available on the BP*stat* |Statistics online in the *Banco de Portugal* website. Additionally, some harmonized sectoral aggregates are made available in the BACH database (Bank for the Accounts of Companies Harmonised), allowing for international comparison among companies in Europe;
- to return to companies valuable CBSD output. CBSD in Portugal deliver to each company the respective Enterprise and Sector Table, containing annual indicators for both the company and its sector of economic activity and size class. This provides companies with information about their position compared to other companies of the same economic activity and size and also from other countries, being, therefore, an important tool for management and decision taking; and,
- CBSD is a good source for updating business registers, which, in turn, plays an important role in statistics.

Finally, detailed accounting data of every firm undoubtedly embodies a powerful tool to respond to any users' requirements, reducing or eliminating potential data gaps in the field of corporation statistics.

Other uses

Alone or combined with other information, CBSD data has also proved to be of great value for pursuing other central bank statutory obligations. For example, data on the financial situation of companies, especially their level of debt and rate of return, is crucial to explain companies' investment decisions and further forecast the productive capacity of the economy. Additionally, data on the financial situation of companies is also essential to evaluating monetary policy impact on their financing and economic performance. Finally, in the field of financial stability, assessing companies' capability to fulfill their debt obligations is essential to detect potential risk sources of losses for financial institutions.

In short, analysis based on a firm-level accounting database is unlimited, irrespective of the purpose, namely because it allows a careful assessment of discrepancies among companies. This is what is done, inter alia, in the CBSD Studies, where some specificity beyond aggregates is explored. Moreover in the context of macroeconomic analysis and for financial stability purposes, indebted firms are better assessed by using individual accounting data in particular to detect sources of weakness.

References

Banco de Portugal (2011). Central Balance-Sheet Studies, 6, December 2011 – New Enterprise and Sector Tables; Adjustment to the Accounting Standards System.

Banco de Portugal (2010). Central Balance-Sheet Studies, 1, November 2010 – Enterprise and Sector Tables.

Banco de Portugal (2008). Supplement to the Statistical Bulletin 1, 2008 – Simplified reporting: Inclusion of the Simplified Corporate Information in the statistics on non-financial corporations from the Central Balance-Sheet Database.

Gonçalves, H. A. (2011). "How can administrative databases help us to understand the funding behavior of non-financial corporations?", Banco de Portugal, Supplement to the Statistical Bulletin 1, 2011.

Lavrador, I. (2011). "Exploring the statistical potential of micro-databases", Banco de Portugal, Supplement to the Statistical Bulletin 1, 2011.

ECCBSO: a tool for the analysis of the non-financial sector of the economy (from individual to consolidated accounts)

Manuel Ortega[1]

ECCBSO – A source of accounting micro-data on non-financial corporations

The European Central Balance Sheet Offices (ECCBSO) is an informal body created in the mid-1980s that currently integrates the Central Balance Sheet Data Offices (whatever this term may mean!) of around 20 European countries that create and maintain around 50 databases of micro-data, mainly based on accounting, directly or indirectly provided by non-financial corporations and groups. Since the inception of the ECCBSO, the work by its members has provided for the possibility of obtaining direct data from the core of the production system (the non-financial corporations, NFC), using to this end the more basic and clear data produced by NFC, namely their financial statements. ***The availability of these databases represents an alternative source so as to better understand the situation of and developments at the NFC.***

The document "Products and Services of the European CBSOs" (see at www.eccbso.org) provides a full overview of the institutional facts of the European CBSOs (most of them belong to National Central Banks, others to National Statistical Institutions), their databases (principally from accounting data of individual and consolidated accounts, in different annual, half-yearly or quarterly periods), their main objectives (statistics, risk assessment, economic analysis, legal publication of accounting data, others) and products and services. According to Eurostat data (statistics on Business Demography, 2012), 25 countries of the EU had in 2009 slightly less than 9 million corporations; the European CBSOs have in their databases information on more than 6 million of these corporations.

Moreover, the ECCBSO itself has two international databases of aggregated data comprising the individual accounts of non-financial corporation (BACH, diffused in www.bach.banque-france.fr) and the consolidated accounts of listed European groups (ERICA database, used only internally by the members of the Committee).

The purpose of this article does not extend to an integrated overview of the ECCBSO tasks and that of its members. It will concentrate its analysis instead on presenting one of the challenges currently faced by NFC statistics and how CBSO data can be of help in this connection: the creation of financial soundness indicators or scoreboards to evaluate the health of the NFC institutional sector. ***The need to use and integrate macro indicators with statistical distributions of micro-data*** (to better understand the concentration and/or dispersion shown by the picture

[1] ECCBSO & Central Balance Sheet Data Office, Banco de España, manuel.ortega@bde.es.

provided by National Accounts and Financial Accounts of the Economy) ***could benefit from the use of micro-data available at European CBSOs.*** Moreover, addressing this matter will pose the challenge, prompted by globalisation, of correctly interpreting the individual accounts of each single company compared with the consolidated accounts of the group it belongs to.

Key performance indicators of NFC: the need for accounting micro-data

Concepts such as production, gross added value, compensation per employee, gross capital formation, net lending or net borrowing, net financial transactions, market price and others are usual in the statistical domain (they form part of National Accounts manuals). ***But they are far removed from normal language of business, i.e. the accounting standards***. Of course, transition tables can provide for switching from the available accounting data (with information elements such as net turnover, purchases, trade receivables, cash and cash equivalents, staff costs, investment in fixed assets, fair value accounting, and so on) to the equivalent concept in National Accounts guidelines. But there will in any case be concepts and ratios where the information available in European CBSOs will provide a unique and cheap (as it is currently available) source of information. Ratios related to profitability and indebtedness, which are necessary to analyse the resilience and confidence of the financial and economic structure of the NFC institutional sector, can benefit from the use of our databases. Moreover, studies on the financial stability of the NFC sector ***do not only need the aggregate picture provided by the Financial Accounts of the economy, but also the subtle details given by statistical distributions, along with breakdowns by size and activity available in CBSO's databases***. It is important to know how the cake is cut and who eats each portion, isn't it?

The challenge of integrating the pictures provided by individual and consolidated accounts

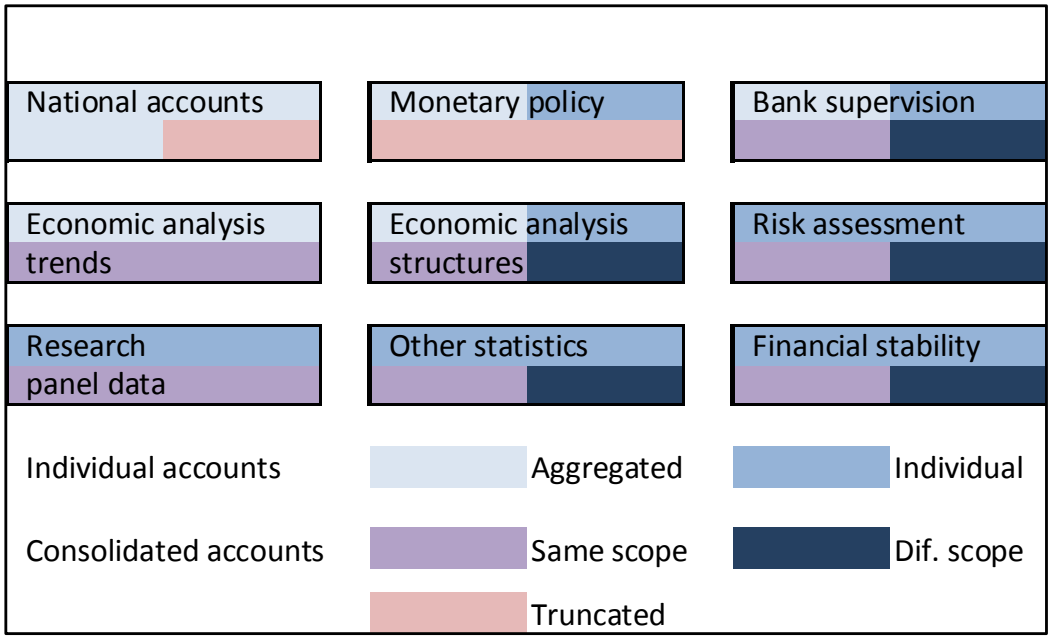
Although the use of individual accounts has been the main source of information for the European CBSOs, ***only recently have they faced the challenge of integrating the data provided by the consolidated accounts of groups in their economic and financial analysis***. Of course, some CBSOs have worked since the late 80s with consolidated accounts, mainly for risk assessment analysis. But it was in the late 90s when new needs arose as a result of globalisation: whereas ***previously, the mere aggregation of individual accounts offered a good picture of NFCs' behaviour, since then this process has had to deal with the "double accounting" problem***. Indeed, the aggregation of individual accounts doubles the profits of subsidiaries that belong to a parent company (which receives the dividends) when both are aggregated in a sample, as is also the case with the losses of the subsidiary, provisioned in the accounts of the parent. Intercompany loans (financial and commercial alike) should also be submitted to a process of consolidation, in order to create clear pictures of the financial structure of an aggregation of individual accounts. In short, the use of individual accounts, by

aggregation, to produce statistics and analysis on the behaviour of a set of companies, should be subject to certain adjustments, to eliminate the double accounting created. We might well ask:

- If we have the consolidated accounts of the group, what is the problem?
- Use it!

But here another problem arises: the scope (the perimeter) of the consolidation available is not what the statistician would like to have, but what the group decides to produce: a perimeter that includes subsidiaries abroad, on one hand, and one which changes continuously, every single year, making it difficult – if not impossible – to use consolidated accounts for the analysis of how the aggregates trend on a constant basis. For example, the increase in net turnover one year comprises its increase in companies that existed the previous year (organic growth) together with the rise brought about by the acquisition of new companies (corporate growth). In theory, the availability of truncated consolidated accounts (by region, or by sector of activity), and also restated annual accounts (comparing the group with the same perimeter in two consecutive years), could solve the problem.

Map of possible uses of individual/consolidated data



ECCBSO members are addressing within their working groups the issue of improving knowledge on how to deal with both kinds of data (individual and consolidated). The panel attached shows, from the experience of the different CBSOs belonging to the committee, a map of uses of both accounts, when it is necessary to use the micro-data and when the aggregation of individual data is enough for the analysis. For example, risk assessment needs individual accounts, treated on an individual basis (company by company), but also consolidated information on the group the company belongs to, irrespective of the changes in the scope (perimeter) of consolidation.

In an unsettled environment, globalisation challenges require statisticians to improve their data sources, reducing at the same time the reporting burden on companies. This new miracle of extracting so much for so many from so little can only be achieved using all the tools at hand. The Central Balance Sheet Data Offices databases can be one more tool on the list.

Credit Information System – the use of micro-data in credit statistics

Renato Baldini Jr[1]

Introduction

The Credit Information System (*"Sistema de Informações de Crédito"* in Portuguese; SCR) is a data bank maintained and administered by the *Banco Central do Brasil* (BCB), with registers of all loans (plus leasing and other credit-like operations) from national financial institutions to enterprises and individuals. Each single credit operation from borrowers whose responsibilities against the financial system are equal or superior to R$ 1,000 (c. USD 500) is individually registered on SCR by financial institutions on a monthly basis. The remaining operations are also included on SCR but in aggregated form.

The SCR was implemented in 1997 with information on credit responsibilities above R$ 50,000, a threshold that was gradually reduced to R$ 5,000. In April 2012, the limit was further lowered to R$ 1,000 in a milestone that was regarded as the move to SCR Phase 2.

The recent growth of the Brazilian credit market

The Brazilian credit market has fully developed only after the consolidation of macroeconomic stability, even though progress has been stronger on the loans market than on the securities market. Having started from a very low base, loans have boomed over the last ten years – total loans to GDP increased from 26% in December 2002 to 53.8% in December 2012 – basically following economic growth and some important institutional advances which allowed for the reduction of interest rates and the expansion of credit supply.

Against this background the SCR has evolved with the main objective of providing the BCB's banking supervision with detailed information about credit operations of the institutions of the national financial system. SCR data allows the banking supervision and the banking system to assess the size and the profile of the financial debt of each borrower (enterprises or individuals), as well as credit-related losses of each financial institution and of the whole banking system.

The access of financial institutions to information on SCR about potential borrowers' credit history and the size of their aggregated financial responsibilities is crucial for the sustainable development of the credit market as it helps market participants to avoid information asymmetries and adverse selection, contributing for the adequate pricing of credit operations. By helping financial institutions to avoid lending further resources to over indebted borrowers and to lower interest

[1] Department of Economics, Banco Central do Brasil, renato.baldini@bcb.gov.br.

rates for historically good payers, SCR also fosters competition on the credit market and facilitates the decline of average interest rates and banking spreads.

Compilation of statistics from SCR data

SCR covers a wide range of information about credit operations, financial institutions and borrowers. Such information can be used for the compilation of a number of macroeconomic and financial stability statistics. The move to SCR Phase 2 resulted in a significant improvement on the coverage of several statistical datasets extracted from SCR especially because of the expansion of the number of borrowers with detailed information registered on the system.

In December, 2011, there were 150.6 million credit operations individually registered in SCR. After the move to SCR 2 the number of registered operations has been increased to 475 million (June 2012). Such data covered information on credits borrowed by 61 million individuals and 3.5 million enterprises.

Information available on SCR has been used for the compilation of regional credit statistics which provide a very precise geographical distribution of credit in Brazil. The compilation of this set of statistics is based on the borrower's postal code which is informed by lenders on SCR. Previous regional statistics were based on information about the location of the lending institution, which tended to result in significantly distorted information, as the location of financial institutions are highly concentrated at major Brazilian financial centers like São Paulo and Rio de Janeiro. Publishing of regional credit statistics based on SCR began in 2007, with available data breakdowns for the five regions and the 27 Brazilian states. The postal code data of credit operations allows for even further breakdowns, at municipalities, neighborhoods and streets level, although the compilation has not advanced to such detail.

A very important indicator based on SCR is the Real Estate Collateral Value Index ("Índice de Valorização de Garantia de Imóveis Residenciais Financiados" in Portuguese; IVG-R). This index measures the long term trend of housing prices based on the appraisal values of residential units financed in the biggest eleven metropolitan regions. The IVG-R was published for the first time at the BCB's Financial Stability Report in March 2013.

Housing prices indexes have been regarded as fundamental to assess the sustainability of housing markets. Their relevance has been particularly highlighted after the U.S. subprime crisis. Accordingly, the compilation of housing prices indicators has been proposed as one of the twenty recommendations of the G-20 Data Gaps Initiative, conducted by the International Monetary Fund and the Financial Stability Board, with support by the BIS, the World Bank, OECD and Eurostat on the aftermath of the 2008 crisis.

As a major instrument of the banking supervision, the SCR has been used for the compilation of a large number of financial stability indicators. Some of them have proved to be crucial to assess the prevailing conditions on the credit market as well as to inform policy decisions. Examples of the application of such indicators are related to the macroprudential measures implemented in the end of 2010 and afterwards. The rapid growth of consumer credit in Brazil has boosted domestic demand and significantly contributed to the acceleration of the economic activity over the last ten years. Along with those positive effects came the buildup of some

specific risks, particularly in the auto loans' market. BCB then introduced measures aiming to curb the expansion of auto loans with long terms and very high Loan to Value ratios (LTV). LTV statistics based on SCR proved to be very important before and after the implementation of those measures, as well as vintage data for non performing loans (NPL), which were intensively used to gauge the evolution of NPLs since 2010.

Information available on SCR is also used for the compilation of statistics on the distribution of loans according to the size of the operations and by their remaining maturity. SCR data also contribute for the compilation of public finance statistics by providing data on the banking debt of government enterprises.

A number of other interesting and important statistical datasets are still being developed based on SCR data. Information on the turnover of enterprises and the income of individuals are going to allow for the compilation of credit statistics with breakdowns by the companies' size and the individuals' income classes. The outcome is expected to shed light for example on the patterns of indebtedness of different income classes and their behavior with regards to the payment of financial responsibilities. Such behavior can also be crossed with information about the number of years of the borrower's relationship with the financial institution, which allows for the assessment of implications of the increased access of lower income classes and smaller enterprises to the credit market.

Another important development front is the compilation of statistics about the distribution of credit across economic sectors. Currently published statistics only provide breakdowns by public or private sector and by households, housing, industry, agriculture, commerce and other services. New statistics being developed from SCR data will be able to open information according to the economic activity code of each enterprise, which will allow for the opening of distributional data in very deep detail.

Concluding remarks

The SCR is a very powerful instrument not only for the sake of banking supervision but also for the compilation of macroeconomic statistics and financial stability indicators. Important information is being extracted from this system contributing for a large number of analyses of the fast growing credit market. SCR's sheer size has posed some operational difficulties for data compilers but gradually new indicators have been produced opening many different ways to expand SCR's utilization as a major source of credit statistics.

A flexible approach to credit statistics: the use of the Portuguese Central Credit Register for statistical compilation

Paula Casimiro[1]

> "*Information is the oil of the 21st century, and analytics is the combustion engine*"
>
> Peter Sondergaard, senior vice-president at Gartner

1. Some initial considerations on data needs for the compilation of statistics

The possible benefits of an information system should be measured against the its value-added to the concerned users. In the specific case of monetary and financial statistics. producing high-quality and timely data has been a key responsibility of National Central Banks (NCBs) and ensuring that these statistics remain adequate implies keeping pace with the developments at the level of the financial markets, assessing the statistical impact of financial innovation in an early stage and introducing the necessary amendments in timely manner. Ideally, these goals should be accomplished without overburdening the reporting agents.

In the recent years, new challenges to the statistical function of NCBs were driven by the rapid financial innovation – with a proliferation of new and increasingly sophisticated financial products in parallel with the appearance of new types of institutions as well as new and expanded roles for existing ones – and by an increased interconnectivity among financial institutions and markets. The international financial crisis that erupted at the end of the last decade exposed some information gaps of the current statistical framework, both at national and international level, in particular for the purposes of financial stability analysis.

New and more demanding data requirements and the need to keep the reporting burden for respondents at a reasonable minimum, gave a new impulsion to the issue of further exploring the statistical potential of micro-data, namely data available at the level of administrative individual registers. In general, the statistical use of micro-data is technically easy to implement, has relatively low costs associated and can provide a good coverage of the reference population. In return, the use of administrative data for statistical purposes can deliver significant reductions in respondent burden, high data quality and, above all, an enhanced capacity to respond to new information requests from the users.

The experience of *Banco de Portugal* in managing highly granular datasets, covering different areas of the economy and the financial markets – namely, securities issues and holdings, a credit register, a central balance sheet database

[1] Statistics Department, Banco de Portugal, pcasimiro@bportugal.pt.

and, more recently, micro data on interest rates and amounts of new loans granted by banks to non-financial corporations – shows that micro-databases and item-by-item reporting can help NCBs to overcome a number of shortcomings related to the more traditional, aggregated, data collecting systems.

The paper is organised as follows: the next section presents a brief overview of the Portuguese Central Credit Register (CCR); section three illustrates how CCR data is being used in the context of the compilation of credit statistics; and, section four, presents some examples of the combined use of CCR data with other micro-databases for the compilation of more valuable sets of statistical data. Lastly, section five concludes.

2. Brief overview of the Portuguese Central Credit Register (CCR)

The Portuguese CCR is an administrative database established in 1978 (the first legal act dates from 1967), at the time covering only the credit liabilities of non-financial corporations (households were included later, in 1993). Although the use of CCR data for the compilation of statistics was authorized since 1996, the responsibility for the management of the database and all related services was assigned to the Statistics Department in 1999. Since then, a number of developments were introduced aiming at improving its coverage and usability, namely the establishment of a bilateral exchange of individual credit data among the seven European countries that signed the MoU[2] (in 2005) or the inclusion of the potential credit liabilities of personal guarantors (in 2007). A major development occurred in January 2009, with the implementation of a new information system that introduced additional breakdowns at the level of credit data and a greater efficiency in identifying private individuals. More recently, the CCR coverage was extended to include new reporting institutions (essentially non-financial corporations that buy credit portfolios from the resident financial sector); a new analytical data system for data analysis and exploration was developed; additional details were included to allow for the individual identification of loans used as collateral in Eurosystem financing operations; and borrowers can now easily access their credit reports online, via the *Banco de Portugal* website.

Picture 1 provides an overview of the coverage of the Portuguese CCR in terms of reporting institutions and data reported, whereas Picture 2 depicts the information flows between the CCR and the major external stakeholders. Some figures are worth mentioning: the reporting threshold is one of lowest for this type of registers (only 50 Euros); 204 institutions are currently providing information on a monthly basis; more than 285 thousand non-financial corporations and around 5.8 million private individuals are registered as borrowers in the database and 23 million records on credit data are reported, on average, every month. It should also be mentioned that the Portuguese CCR maintains for several years the first

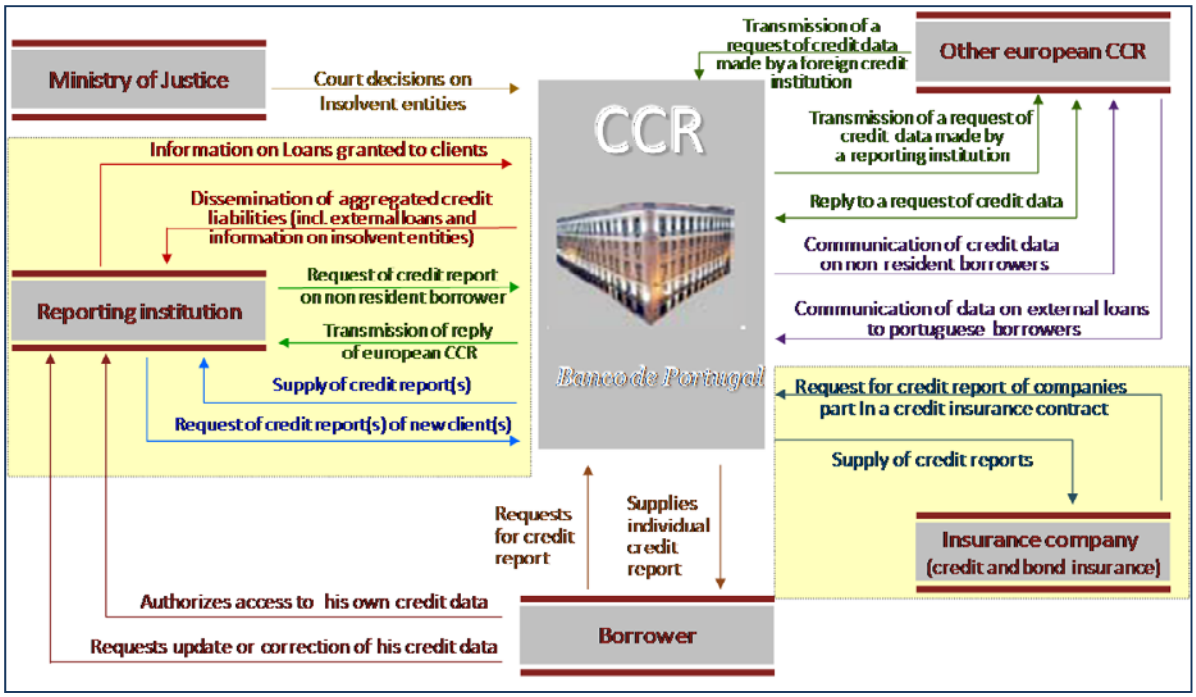
[2] *Memorandum of Understanding on the exchange of information among national central credit registers for the purpose of passing it on to reporting institutions*. Besides Portugal, the other signatory countries were Austria, Belgium, France, Germany, Italy and Spain. In 2010, the Czech Republic and Romania also joined.

place in the ranking of the World Bank's "Doing Business" annual report in terms of its coverage of the population.

Reporting Institutions and data reported to the Portuguese CCR — Picture 1

Reporting Institutions	Data reported
Banks, savings banks and mutual agricultural credit banks (MFIs)	Borrowers ID (for residents, the tax payer number is used)
Other non-monetary financial institutions that grant credit	Credit drawn (amounts outstanding at the end of the month)
Public agencies that grant credit	Credit undrawn (irrevocable credit commitments)
Non-financial corporations buying loans from the resident financial sector	Personal guarantees (potential credit liability)
	Type/purpose of the loan
	Collateral (type and value)
Insurance companies undertaking credit and bond insurance can access CCR data, although they do not report.	Periodic repayments (for some types of loans granted to private individuals)
	Original and residual maturities
	Credit defaults and write-offs
	Specific flags for *Banco de Portugal* internal use of the data (e.g. securitised loans, loans used as collateral in Eurosystem financing operations)
	Inter-bank loans, securities and financial derivatives are excluded

Information flows in the Portuguese CCR — Picture 2



3. The use of CCR data for the compilation of credit statistics

The main objective of the CCR is to contribute for the financial sector stability by assisting financial institutions in assessing the credit risk of their current or new credit clients. However, the prevailing CCR legal framework already foresees that data can be used in the context of specific functions of the Central Bank, such as supervision, research, financial stability analysis, monetary policy and statistical compilation.

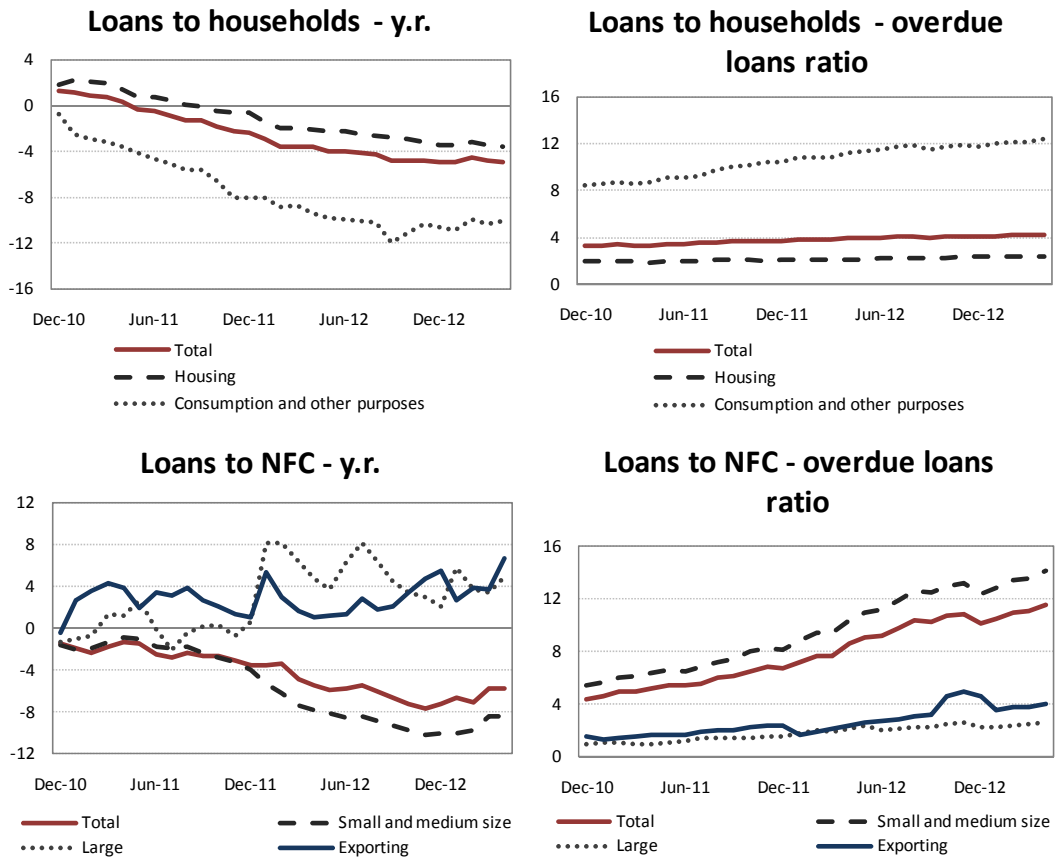
In the domain of banking supervision and regulation, CCR has been used in the assessment of credit risk and concentration of risk exposures, both at micro and macro level, and for improvement of on-site inspection practices. Economic research has been using CCR micro data for several research papers and analysis, e.g., structural analysis, frequently combining this data with other micro data sources, like the central balance sheet database. Finally, within the monetary policy framework, CCR has been used in the identification of loans used as collateral in Eurosystem financing operations.

The use of CCR data in the context of monetary and financial statistics (MFS) has been facilitated given the fact that both domains share the same data source, i.e. the same reporting institutions, the content of the reported information is the similar since the CCR covers a complete range of credit liabilities, they both have identical reporting frequency and timeliness and, last but not least, both the CCR and MFS are integrated in the same division in the Statistics Department.

The use of CCR data for statistical purposes has allowed, inter alia, an improvement in the quality of monetary financial institutions (MFIs) and other financial institutions (OFIs) balance sheet statistics (e.g., greater accuracy in the MFIs' classification of the institutional sector of the counterparties receiving credit), a better assessment of credit developments and the conception of new statistical products, without imposing additional reporting requirements and burden on respondents (e.g. CCR data was particularly useful to meet the ECB requirements on securitisation without having to increase the reporting burden on Financial Vehicle Corporations).

Statistics on loans granted by the resident financial sector to non-financial corporations and households compiled based on the CCR are regularly published on the Statistical Bulletin and on the BP*stat* | *Statistics Online*, the interactive dissemination database available at *Banco de Portugal* website, and include information on outstanding amounts of loans, default ratios and number of borrowers with default loans. Available breakdowns comprise the purpose of the loan, institutional sector, economic activity, region or size of the company. Picture 3 exemplifies some of the indicators currently published.

Loans granted by the resident financial system Picture 3



4. The integration of CCR data with other micro-databases

Economic research and financial stability analysis have been extensively using and combining micro data coming from different databases managed by the Statistics Department. CCR data is used in internal models (e.g. for estimating the probabilities of default under different macroeconomic scenarios) and in several scholarly research papers, frequently combined with other micro datasets, notably with the central balance sheet database.

In the first half of 2011, with the deepening of the Portuguese sovereign debt crisis, Portugal requested for joint EC/ECB/IMF financial assistance, which terms and conditions were laid down in an Economic Adjustment Programme. Under this Programme, a balanced deleveraging of the banking sector was deemed critical, necessarily resulting in a reduction of credit available for the economic agents, namely non-financial corporations and households. Given the elevated indebtedness level of these sectors and their high dependence on the resident financial sector for financing, concerns arose around their funding pressures and debt refinancing activities. On the one hand, it is important to safeguard adequate financing of those more dynamic sectors to spur growth and, on the other, it is no

less important to monitor the evolution of the non-performing loans, which put additional pressure on the resident banks' balance sheet.

Several statistical products have been developed to meet the huge increase in user demand stemming from the need to adequately monitor all developments on the financing conditions, both from the creditor and the debtor perspectives. Some of these outputs address ad-hoc requests of data, whereas other became more permanent and gave origin to new statistics published on a regular basis (e.g. new chapter in the Statistical Bulletin on the non-financial sector indebtedness).

Some of these new statistical products combine data coming from several statistical domains – monetary and financial statistics, securities statistics, balance of payments and international investment position and financial accounts – and from different micro datasets – central credit register, central balance sheet database and securities issues and holdings – making use of a common infrastructure for reference data, in particular a business register.

5. Conclusion

Banco de Portugal has been making use of the advantages entailed by the use of micro databases and item-by-item reporting: reducing respondents' burden, enhancing the quality control and cross checking of elementary data and improving responsiveness to ad hoc information requests from users. Pursuing this path it has taken advantage of the centralised management of these databases.

Notwithstanding the Portuguese Central Credit Register has been created with the objective of providing the participating institutions with relevant information to better access the risk associated with a specific credit contract or borrower, it holds also nowadays a significant potential for other purposes, namely compilation of statistics, economic research and financial stability analysis.

The data reported to the CCR has gained relevance with the current indebtedness situation of the Portuguese economy combined with the pressing need of economic agents in all sectors to deleverage their activity, including the banking sector. CCR data combined with other micro databases (namely securities holdings and issues and corporate balance-sheet data) has been a key factor in meeting all the data demands in the context of the economic and financial assistance programme.

For an efficient use of micro data and to fully explore its value for statistical compilation and analysis two major requirements do exist, namely when combining different micro data-sources: the need for a comprehensive and up-to-date business register and flexible analytical tools to explore the information.

References

Almeida, A., D'Aguiar, L. and Casimiro, P. (2011), Promoting enhanced responsiveness to users' data needs: the experience of the Banco de Portugal in exploring the statistical potential of micro-databases, 58th World Statistics Congress of the International Statistical Institute, Dublin, August 2011.

Banco de Portugal (2005), Supplement 1/2005 to the Statistical Bulletin, A New Source for Monetary and Financial Statistics: the Central Credit Register, April 2005.

Improving the breakdown of the Central Credit Register data by category of enterprises

Jean-Pierre Villetelle[1]

Departing from legal units, and applying the definition of enterprise complying with the 1993 EC Regulation and a 2008 French Law changes markedly the estimates of the relative shares of each category of firms in the economy. For instance, in terms of employment and Value Added in the private nonfarm business sector, the Large Enterprises (new definition) play in France a role twice bigger as previously estimated on the basis of the legal units with more than 5,000 employees (old definition). Conversely, the share of SMEs is notably smaller.[2] The aim of this note is to present the motivations and the results of the same analysis about the Central Credit Register data. It shows that the outstanding amounts of credit are much more concentrated on SMEs than previously estimated on the basis of legal units.

In most cases, statistical data on firms are collected from administrative sources. Therefore, they essentially relate to legal units, that is, to units recognized primarily for legal reasons (by fiscal authorities, commercial courts, social security administrations, control authorities, etc.). In some cases – and numerically, in most cases – the legal unit indeed corresponds to what is usually called an "enterprise". This however reflects essentially the case of the smallest entities. As soon as the entrepreneurs or the owners of an enterprise choose to structure this one in several legal units (for organisational, legal or fiscal reasons), there is no one-to-one correspondence anymore between the notion of enterprise and the one of legal unit. This is especially obvious if we take the case of holding companies which, by nature own other companies' shares and do not produce goods or services themselves.

In this respect, the Council Regulation (EEC) No 696/93 of 15 March 1993 "on the statistical units for the observation and analysis of the production system in the Community" stated that:[3]

> "The enterprise is the smallest combination of legal units that is an organizational unit producing goods or services, which benefits from a certain degree of autonomy in decision-making, especially for the allocation of its current resources. An enterprise carries out one or more activities at one or more locations. An enterprise may be a sole legal unit." (Section III.A)

The explanatory note of the Council Regulation clarifies that:

[1] Companies Observatory, Banque de France, jean-pierre.villetelle@banque-france.fr.

[2] Béguin J.-M., Hecquet V., Lemasson J., direction des Statistiques d'entreprises (2012), "Un tissu productif plus concentré qu'il ne semblait. Nouvelle définition et nouvelles catégories d'entreprises", Insee Première n° 1399 - Mars 2012; http://www.insee.fr/fr/themes/document.asp?ref_id=ip1399.

[3] http://eur-lex.europa.eu/LexUriServ/LexUriServ.do?uri=CELEX:31993R0696:EN:HTML.

"The enterprise thus defined is an economic entity which can therefore, under certain circumstances, correspond to a grouping of several legal units. Some legal units, in fact, perform activities exclusively for other legal units and their existence can only be explained by administrative factors (e.g. tax reasons), without them being of any economic significance. A large proportion of the legal units with no persons employed also belongs to this category. In many cases, the activities of these legal units should be seen as ancillary activities of the parent legal unit they serve, to which they belong and to which they must be attached to form an enterprise used for economic analysis."

This definition did not spread much in practice, but in the case of France, a law passed in 2008 and its related Statistical Decree define now the practicalities for using this definition, as long as the data for it are available.[4, 5]

In addition, this Decree introduces definitions for the size categories of enterprises, on the basis of three criteria that must be used simultaneously – and not alternatively: employment, turnover and balance sheet total. In between the usual Small/Medium-sized Enterprises (SMEs) and Large Enterprises (LEs), a new category is introduced: Intermediate-Sized Enterprises (ISEs) – for the French "*Entreprise de Taille intermédiaire* (ETI)". Finally, inside the SMEs, a subcategory of Micro-Enterprises is defined. Table 1 summarises the various combinations of criteria, and the outcomes as regards the size categories.

Employment remains the main determinant of each size category. But an enterprise of a given category according to the employment criterion only, can be upgraded to a larger category, when turnover and balance sheet total are taken into account.

[4] The Law for Modernizing the Economy (Loi de Modernisation de l'économie (LME), art. 51. http://www.legifrance.gouv.fr/affichTexte.do?cidTexte=JORFTEXT000019283050.

[5] Décret n° 2008-1354 du 18 décembre 2008 relatif aux critères permettant de déterminer la catégorie d'appartenance d'une entreprise pour les besoins de l'analyse statistique et économique, http://www.legifrance.gouv.fr/affichTexte.do?cidTexte=JORFTEXT000019961059.

Size categories according to three criteria: Employment, Turnover and Balance Sheet Total

Table 1

Turnover	Balance Sheet Total	Employment (number of employees)			
		< 10	10 to 249	250 to 4 999	> 5 000
< 2 €M	-	micro-enterprises	SMEs	ISEs	LEs
[2; 50[€M	< 2 €M				
	> 2 €M				
[50;1,500[€M	< 2 €M	micro-enterprises			
	[2; 43] €M				
	> 43 €M				
> 1,500 €M	< 2 €M	micro-enterprises	SMEs		
	[2; 43] €M				
	[43 ; 2,000] €M				
	> 2,000 €M				

(€M = million euros)

SMEs: Small/Medium-sized Enterprises

ISEs: Intermediate-Sized Enterprises

LEs: Large Enterprises (LEs)

Applying this decomposition to the Central Credit Register Data means:[6] first, from this set of legal units, identify all the related legal units – even if not registered themselves in the Central Credit Register – to build the enterprises as defined by the Statistical Decree; second, define their size category; third, evaluate the outstanding amount of credit for each of them, on the basis of the one of each constituting legal unit. *A priori*, this would be relatively easy since, on top the Credit Register data, all the information needed (financial linkages (*i.e.* the structure of ownership), employment, turnover and balance sheet total) are included in the French tax form. Linking these two sets of information, when available, is also easy with French data, on the basis of the SIREN number.[7]

[6] In the case of France, all the legal units with an outstanding amount of credit above € 25,000.

[7] In France, the SIRENE register (*Système informatisé du répertoire national des entreprises et des établissements*) lists the civil status of all the legal units. When registered for the first time, a number is given to each legal unit – the SIREN number – and in all subsequent events in the life of the legal unit, the SIREN number is reported.

The problem is actually the availability of information. The tax forms are indeed available from various data sources, but with various degrees of timeliness, quality and completeness. Banque de France collects some of them, as part of the FIBEN database.[8] In that case, they are largely checked and possibly corrected by the regional branches, regularly in contact with the firms themselves. But only fiscal forms for firms with turnover over € 750,000 are collected this way. For the smaller entities, it is possible to purchase the missing data to private data providers. Since, even this way, some data are still missing, we ultimately rely on the NSI, which implies an additional delay since the NSI itself gets this information only after the fiscal administration. We hence have some kind of pecking order, favouring both the quality and the timeliness of the data in using first our own data sources, second the private data providers and third the NSI.

This data limitation makes impossible to strictly adopt a breakdown in just three categories only: SMEs, ISEs and LEs. First, because before they fill in any fiscal form, it is impossible to have information of newly created legal units (either new independent legal units or new entities belonging to existing enterprises). It is hence impossible to determine the category of these entities before at least one year. But it can also be the case that the information is simply missing. Hence, no information at all can relate to a new entity, or a gap in our information system. Second, some information about the financial linkages may be missing as well, which can be a case of independent legal unit, or missing information. This is especially damaging regarding Real Estate Activities, because we cannot discriminate between property investment companies with business activities (for which we expect financial linkages with the entities they serve) and housing associations, with non-business activities (for which no financial linkages are expected since they act for themselves, the members of a same family for instance).

For the first time, in June 2013, with the data for April, we were able to publish results complying with the 2008-1354 Statistical Decree (Table 2). Despite the existence of two additional categories ("Other Real Estate companies" and "Unreferenced enterprises") we can focus on main categories SMEs (not distinguishing micro enterprises), ISEs and LEs.[9] The former category "holding companies" disappears as well. The category involving Real Estate Companies is markedly reduced. This table shows in particular that the outstanding amount of credit to SMEs is increased from € 210 Billion to € 360 Billion. Broadly speaking, from the initial estimate, € 190 Billion stay in the category SMEs (90 % of the initial amount). An additional amount of € 170 Billion come from categories previously referenced as: Real Estate activities (€ 110 Billion), Holding Companies (€ 40 Billion) and Larger companies (€ 20 Billion).

Conversely, the outstanding amount of credit to LEs is reduced from € 125 Billion to € 90 Billion. Only € 40 Billion of the initial estimate stays in this category (1/3rd). The rest is distributed to entities now seen as of smaller size, especially ISEs (€ 60 Billion). In addition, € 30 Billion coming from Holding companies and € 20 Billion coming from Real Estate activities are now referenced as credit to LEs.

[8] http://www.banque-france.fr/fileadmin/.../gb/i133gb.pdf.

[9] The distinction between "independent SMEs" and "SMEs belonging to a group" is not relevant anymore, since, made of a single legal unit, or being a combination of legal units, an enterprise is necessarily independent (SMEs being no exception).

The outstanding amount of credit of the new category ISE (€ 215 Billion) comes from the previously defined outstanding amount of credit of real Estate Activities (40 %), LEs (30 %) and 20 % of Holding companies, the remaining 10 % coming from the other categories, especially the SMEs.

For the subset of enterprises that belong both to the Central Credit Register and to the FIBEN database maintained at Banque de France, these estimates correspond for SMEs to around 95 % of their bank debts, as recorded in the fiscal forms. This ratio falls to 80 % for ISEs and to 60 % for LEs. For bigger entities, non-resident credit institutions take hence a significant part in total bank debts.[10]

The new definition has also an impact on the breakdown by sector (not developed here) with a reduction of the size of sectors like Holding activities and services to companies, to the benefit of the industrial and trade sectors.[11]

[10] Direction des Entreprises - Observatoire des entreprises (2010) "La situation des entreprises en 2009: baisse de l'activité et attentisme en période de crise", Bulletin de la Banque de France, n° 182, pp. 18-19, 4e trimester, http://www.banque-france.fr/fileadmin/..._etu_1.pdf.

[11] See: "Loans by type of companies", Banque de France website, http://www.banque-france.fr/.../detail/loans-by-type-of-companies.html.

Central Credit Register Data – France, April 2013

Table 2

A – Breakdown by categorises: Legal units

Outstanding amounts, Billion euros

	Number of	Drawn Credit		Drawn and Undrawn Credit	
	Enterprises	Outstanding amount	Annual change %	Outstanding amount	Annual change %
	April 2013	April 2013	April 2013 / April 2012	April 2013	April 2013 / April 2012
Independent SMEs	-	151.8	-0.1	168.2	-0.5
SMEs belonging to a group	-	64.2	3.9	75.3	3.3
Sub-Total (SMEs)	-	215.9	1.1	243.4	0.7
LEs	-	124.2	0.5	197.4	0.4
Holding Companies	-	109.9	-5.5	210.8	-6.2
Real estate activities	-	356.5	2.9	389.1	2.9
Total	-	806.5	0.8	1 040.8	-0.1

B – Breakdown by categorises: Enterprises

Outstanding amounts, Billion euros

	Number of	Drawn Credit		Drawn and Undrawn Credit	
	Enterprises	Outstanding amount	Annual change %	Outstanding amount	Annual change %
	April 2013	April 2013	April 2013 / April 2012	April 2013	April 2013 / April 2012
SMEs	1 080 887	357.6	0.7	405.1	-0.5
ISEs	4 802	215.1	-2.5	263.4	-1.9
LEs	219	85.5	-6.4	213.4	-5.0
Sub-Total (SMEs, ISEs and LEs)	1 085 908	658.3	-1.3	881.9	-2.0
Other Real Estate companies	471 201	137.3	3.7	143.4	2.8
Unreferenced enterprises	36 998	13.0	NS	18.9	NS
Total	1 594 107	808.6	1.1	1 044.3	0.3

Using credit registry data to assess the risks in the financial system – the *Banco de Cabo Verde* experience

Ana Semedo[1]

The current economic and financial crisis, now labelled "the Great Recession" by the IMF, started in 2007, largely as a result of increasing sophistication of the financial sector, which was not matched by improvements in regulatory and supervisory mechanisms.

Since then, and in order to cope with the dire consequences of a major financial crisis, such as the related huge social costs, world authorities have devoted greater attention to strengthening financial sector soundness. Further, authorities have gone to great lengths to prevent and mitigate systemic crises – often, as in the case of Cyprus, with devastating consequences in terms of economic agents' confidence.

Due to its strong dependence on the Euro Zone, the Cape Verdean economy has been severely affected by the European crisis, with FDI dropping (since 2009) and, in more recent years, lower levels of exports of goods, emigrant remittances and foreign aid. The reduction in these important external flows has hampered economic growth, which has been sluggish since 2009. The combined effect of low growth and counter-cyclical budget policies has been a surge in domestic and external debt.[2]

Bank asset quality has been hurt to a certain extent by the adverse external environment and by the deteriorating local economic context. Nonperforming loans (NPL) almost doubled in the last three years, and currently stand at 14 percent of total bank portfolio. Credit concentration risk with regards to both economic sectors and borrowers is also high, thus increasing the vulnerability of both banks and non-banks.

Considering the devastating consequences of a banking crisis, in a context of limited fiscal space for the public sector to assume the role of lender of last resort, as well as the need to develop the financial sector in line with the country's potential for economic development, Banco de Cabo Verde moved to set up a task force to study the reasons behind the recent surge in NPL in the banking sector.[3]

The task force was given a mandate to:

a. Identify the causes behind the increase in NPL by analyzing to which extent the deterioration in the local and/or international context, public policies, social causes (divorce, for instance) and failures in risk appraisal and moral hazard, … have contributed to bank asset deterioration;

1 Research and Statistics Department, Banco de Cabo Verde, asemedo@bcv.cv.

2 Budget policy from 2009 onwards also aimed at taking advantage of available concessional financing for infrastructure.

3 The focus on the banking sector owes to the fact that the securities market is incipient. In addition, banks are major holders of corporate bonds.

b. Forecast future trends in terms of NPL based on the causes and the risks identified;

c. Assess bank provisions and capital adequacy relatively to existing risks and provide recommendations for strengthening financial and macroeconomic stability.

The task force's mandate inspired this presentation because it was necessary to use (or at least attempt to use) information from BCV's Credit Information Registry (*Central de Riscos de Crédito – CRC*) to undertake the first of the three tasks assigned to the group.

The first step was to define information needs related to the assigned task. With regards to the corporate sector, the task force agreed on a number of issues pertaining to the company – name, location, domestic and foreign ownership, public ownership, and area of operation. Information from the banking side was also addressed, namely amount of financing to defaulting borrowers, loan dates, maturity and purpose (i.e., working capital, short term financing or capital investment) for each loan, guarantees, amount of NPL, date of last repayment and number of past due instalments.

With regards to individuals, the group aimed to collect information on fiscal status, address, birth date, gender, marital status, monthly/annual income, employer, whether emigrant or not, as well as information on loans on a per bank basis. In terms of loan purpose the group sought to identify cases of loans for housing purchase and construction, construction for rent/sale, vehicle acquisition and student loans.[4] The group further sought to identify risks related to the different financial instruments used in each case, namely promissory notes, credit cards, salary advances, etc.

The task was hindered by a number of constraints faced by the CRC in collecting information from the banks. Current bank reporting does not meet information needs identified by the group. For instance, with regards to borrower identification, the CRC only collects data on name, area of activity and municipality where the company's headquarters or the individual is located.[5]

The group also faced constraints relating to loan classification. CRC data only includes information on whether the loan was given to a single or to multiple borrowers, on loan instrument and maturity, guarantees, and status (performing, nonperforming, under litigation or written off).

To deal with the above constraints, the group sent the banks a questionnaire to collect data on the 100 largest defaulting borrowers, corporate and individuals.

The Task Force strongly believes that a CRC should meet the supervisory needs in an economically, financially and socially more complex and demanding environment, as well as provide a database with information to allow studies to improve knowledge about the country's economic and financial context, thus facilitating policymaking for economic development. To achieve these objectives, the Task Forces recommends:

[4] The perception is that most cases of nonperformance arose out of one of these circumstances.

[5] For individuals, CRC data also includes birth date and parents' names.

a. An assessment of CRC data collection, validation and analysis process, to be performed by a team of bank supervision, statistics and research experts;

b. Design and implementation of an Action Plan to maximize the usefulness of the CRC in terms of bank risk assessment and statistical needs;

c. Improved dissemination of the upgraded CRC to the banks so they may improve their risk appraisal framework, in line with a risk-based supervisory approach.

References

Banco de Cabo Verde. 2012. "*Relatório de Política Monetária Novembro 2012*". Praia.

Banco de Cabo Verde. 2012. "*Relatório Anual 2011*". Praia.

Banco de Cabo Verde. 2012. "*Relatório de Estabilidade Financeira 2011*". Praia.

Banco de Cabo Verde. 2012. "*Indicadores Económicos e Financeiros Dezembro 2012*". Praia.

International Monetary Fund. 2012. "*World Economic Outlook: Coping with High Debt and Sluggish Growth*". Washington, DC.

International Monetary Fund. 2012. "*Global Financial Stability Report: Restoring Confidence and Progressing on Reforms*". Washington, DC.

World Bank. 2012. "Global Financial Report 2013: Rethinking the Role of the State in Finance". Washington, DC.

Application of micro-data for systemic risk assessment and policy formulation

Karen Lee Fong Ling[1]

1. Introduction

In general, macro trends are useful as it indicates collective behaviours and can be used to develop macroprudential policies on a broad-based level. However, macro data, while helpful, may mask or differ from certain micro trends. This could be due to the cross-subsidisation effect when aggregating micro data. For example, potential financial stress within certain income groups in the household sector may be artificially suppressed by other income groups that have larger financial buffers, creating the inaccurate conclusion that the entire household sector is sound. On the contrary, granular information could reflect more specific individual behaviours. This allows for a more detailed and in-depth assessment on systemic risks, which facilitate the development of more targeted macroprudential policies.

Acknowledging the importance of granular data for systemic risk assessments, the Central Bank of Malaysia has, in recent years, intensified efforts to enhance data capture and incorporate micro data in its macro level assessments of the household sector.

2. Types and sources of micro data for household sector assessment

Micro data for systemic risk assessments on the household sector can be categorised into three broad categories:

i. Credit-related data;

ii. Financial assets / Income data; and

iii. Behavioural data.

I. Credit-related Data

Credit information is vital for the assessment of both institutional and systemic risks. Credit information could be sourced from both regulated and non-regulated entities, and can either be centralised or segregated.

In Malaysia, the bulk of financing (excluding financing from the financial market) extended to the economy is captured in the public credit registry, or more commonly known as CCRIS (Central Credit Reference Information System). CCRIS is

[1] Financial Surveillance Department, Bank Negara Malaysia, karenlee@bnm.gov.my.

a major source of credit data for the Central Bank's financial stability monitoring and risk assessment activities. This centralised database, administered by the Central Bank, collates transactional credit information of all borrowers (individuals, businesses and other borrowers) of financial institutions regulated by the Central Bank and major non-bank financial institutions, and disseminates such information back to the participating financial institutions in the form of a credit report. Each financing application, regardless of amount, received by the participating financial institutions is captured in CCRIS and updated at least on a monthly basis. CCRIS contains both positive and negative credit information of any borrowings. It also captures approved and rejected financing applications, repayment behaviour, and borrower demographics.

Within the Central Bank, data in CCRIS is analysed extensively for systemic risk assessments and micro supervision using a range of tools, from simple standard reports to more complex query tools with data mining capability.

II. Financial Assets / Income Data (of households)

Financial assets include liquid financial assets (deposits with financial institutions, holdings of unit trusts and equity shares), and illiquid financial assets (endowment policies and savings in pension funds). Income refers to recurring monthly remuneration received by an individual. Financial assets provide an indication of an individual's wealth capacity to match his/her financial liabilities, while income data assists in the assessment of the ability of an individual to meet monthly debt obligations.

III. Behavioural Data

Behavioural information gives insight into the motivation and psychological drive behind borrowers' attitude towards credit. While credit and financial / income data quantifies the capacity and ability of a borrower to attain and repay financing, behavioural data captures the willingness of a borrower to do the same. Such qualitative characteristics are difficult to capture purely from quantitative data alone.

3. Methodology

The Central Bank utilises a combination of all three data categories above to ensure timely identification of risks to financial stability arising from the household sector, and subsequently to formulate targeted policy measures. There is no one fixed method to arrive at such assessments, but the continuous process below depicts the general method that the Central Bank uses in its assessment of the household sector:

i. Identify data needs and potential sources of data

ii. Source / collect data – preferably at the granular level

iii. Match data from various databases

iv. Aggregate data

v. Assessment and monitoring of data needs

I. Identify Data Needs

Identifying data that is needed for an assessment is an important part of systemic risk assessments. The lack of data could impede the quality of an assessment. This would primarily involve reviewing existing data to assess whether it is sufficient to facilitate meaningful assessment. If necessary, new data will need to be sourced or proxied.

In Malaysia's case, there is minimal gap in credit-related data, as a huge part of micro credit data is already captured in CCRIS. Efforts are underway to improve collection of income data, at the granular level. Income data is needed to, among others, compute debt-to-income ratios. This is important as overly leveraged individuals in the household sector with significant exposures to the financial system could pose a risk to social stability, if not financial stability.

II. Source Data

As mentioned above, credit-related data is primarily sourced from CCRIS. Unlike credit information, financial information in Malaysia is not centralised. Financial assets and income data are sourced from both public and private sources, including from financial institutions, the government / national tax authority, the capital market, and private pension funds.

The government conducts a periodic survey on households' income and expenditure to get a sense of households' spending patterns, consumption behaviour and attitude towards credit. To ensure that the result from the survey is representative of the entire population, the survey is done on households from different geographical locations, educational backgrounds, work experiences, and income levels. The Central Bank also conducts an annual survey to financial institutions and major non-bank financial institutions to get the credit exposure of these institutions by income groups. Additionally, work is also underway to develop a consumer sentiment survey to gauge an individual's current and future outlook on the economy that may affect consumption.

III. Match Data

More often than not, data that is captured in separate databases are analysed in isolation. This may mask system level risks, which could have been detected from a consolidated assessment on a combination of data from different databases. The Central Bank is in the process of embarking on an exercise to match credit data with income data of each individual borrower. Using the national identification number as a unique identifier, credit data in CCRIS is to be matched with income data from various sources in a black box. To preserve the privacy and identity of the individual, only depersonalised information (sans identity) will be retrieved from the black box into a consolidated server for assessment. Given the size and magnitude of the information, analyses are to be done using a query tool that is sophisticated enough to allow data mining and assessments in the system itself.

In the interim, the Central Bank matches the different data types, for example credit and deposit or income information from various databases at the macro level to gauge, at the aggregate, the available financial buffers of different income groups.

IV. Aggregate Data

The micro credit data in CCRIS can be aggregated by various permutations, e.g. by financial stability indicators, by demographic or by sector (Table 1). It is important to ensure that data obtained is comparable, if data from different sources are to be aggregated.

Aggregation of CCRIS data — Table 1

Aggregation by	Examples
Financial stability indicators	Delinquency ratio, leverage level, repayment ratio
Demography	Geographical location, age groups
Sector	Financial institutions, non-bank financial institutions

V. Assessment and Monitoring

Another element in the process is continuous assessment and monitoring of existing data. This is important to ensure that data applied in assessment of risks remains relevant. If there is a need for new data, then the whole process repeats itself.

4. The Malaysian Experience

There are various types of systemic risk assessments that can be conducted with micro data. Among others, the information can be used to profile risk areas, identify concentration of risks, conduct trend analyses and develop econometric models. Based on the results of the assessments, targeted macroprudential policies may be formulated to address specific risk areas. Subsequently, the information would also help measure whether the policies that have been implemented meet intended objectives.

In the case of Malaysia, credit data that has been matched with income data, even at the aggregate, gives an indication of the leverage level of the different income groups. Although the entire household sector's balance sheet appears healthy when assessed in isolation, the consolidated assessment on data from different databases affirmed that lower income individuals had higher leverage levels when compared to those in the higher income category. Individuals in the lower income category were also found to have larger personal financing exposures with non-bank financial institutions, some of which are not regulated by the Central Bank. Based on these assessments, the authorities were able to develop targeted policy measures to help individuals in the lower income category cope with the higher cost of living. This included expanding the country's debt management and counseling services, which currently covers only regulated financial institutions, to include borrowers from non-bank financial institutions. Additionally, the government also provided financial assistance to targeted household sub-segments to ease financial burden.

Micro data can also be used to monitor the effectiveness of policy measures that have been implemented. For example, transactional data from CCRIS was used to implement the loan-to-value (LTV) limit to curb speculative activity and ensure a

sustainable property market in Malaysia. Subsequently, the data in CCRIS was used to measure the effectiveness of the same macroprudential measure. Following the introduction of a 70% LTV limit on the third housing loan onwards in 2010, the number of individuals with such loans grew at a much slower pace of 2.9% compared to 14.9% prior to implementation.

5. Conclusion

This recent crisis has reinforced the importance of data, especially granular level data, to assess systemic risks. A lack of granularity in data may impede in-depth assessments on macro-level risks to the financial sector. Based on Malaysia's experience, the availability of granular information has enabled the Central Bank to:

i. Conduct more in-depth risks assessments;

ii. Timely identification of risks to financial stability;

iii. Formulate targeted policy measures; and

iv. Measure effectiveness of macroprudential policies.

The process of capturing granular information and enhancing existing information databases is not without challenges. The relevant authorities would have to consider the available infrastructure and readiness of institutions to meet reporting requirements. Legal frameworks may also need to be established to ensure more efficient and effective sharing of information. There is also the need to obtain buy-in from all relevant stakeholders. Investment in infrastructure, may be costly and time consuming, but could avail authorities to consolidated data on a regular basis after the initial outlay. Alternatively, manual surveys could be cheaper and produce quicker results, but data will be less frequent, with potential break in the series. The other key challenge is capacity development. This includes having resources with the knowledge and skills to maintain quality of data, and to conduct data mining, interpretation and analytics.

Admittedly, having granular data alone may not help stop another crisis from occurring. Understanding the data, continuous surveillance and conducting consolidated assessments from various databases are also important. Collectively, this will, hopefully, help authorities in systemic risks assessments and formulation of macroprudential policies to help evade or minimise the impact of future crises.

References

Abdul Ghani, N. (2010). *Household Indebtedness and its Implications for Financial Stability in Malaysia*, SEACEN, Kuala Lumpur, Pages 67–88.

Bank Negara Malaysia (2009). *Macroprudential Approach to Financial Surveillance and Regulation in Malaysia*, Financial Stability and Payment Systems Report, Pages 48–52.

Bank Negara Malaysia (2013). *Central Credit Reference Information System (CCRIS)*, http://creditbureau.bnm.gov.my/index.php?ch=4&pg=6&ac=3&lang=en.

Micro-data based BoP data quality management

Zsuzsanna Fekete and Péter Bánhegyi[1]

Introduction

Hungary joined the European Union in 2004. Simultaneously with EU-membership the Intrastat survey system was introduced to capture the foreign trade transactions with the European Union member countries. Just after these changes the trade balance of Hungary improved significantly: exports increased, while it was not accompanied by a similar imports growth, although Hungary is a very open country. At the same time in the balance of payments statistics net errors and omissions increased and turned to a negative trend. This phenomenon highlighted the importance of systematic data quality management and the macro and micro level checking of data consistencies between financial and non financial accounts. There are different practices in the European Union as regards tasks and legal possibilities of public institutions in data collecting and releasing, the most common way, however, seems to be that data for financial and non financial accounts are mainly collected by NCB's and NSI's, respectively. At the same time balance of payments statistics are mainly published by NCB's, so there are common fields in non financial accounts and balance of payments, mainly foreign trade in goods and services. As a result, NSI's collect and (partly) NCB's release their data. The overall need for harmonization of financial and non financial accounts, and particularly the consistency of balance of payments, requires strong cooperation between these institutions. This paper shows via an interesting example of VAT registrations how the different aspects of globalization affect this cooperation.

Globalization: the case of VAT registrations

One of the main visible features of the globalization is the operation of multinational enterprises. They establish affiliates and branches in different countries throughout the world, using legal possibilities for their own purposes. National statistical authorities collect data on their operations only in the resident country although their transactions cannot be understood without knowledge of their overall activities.

This makes a challenge for statistical authorities, both at national and international level, as regards ensuring that the main goal of (economic) statistics, i.e. describing and showing (economic) processes in well-chosen numbers and indices, will continue to be carried out. At international level as regards our recent topic it mainly results in methodological harmonization between financial and non financial accounts. The work, of course, does not stop at this point: issues of

[1] Statistics Department, Magyar Nemzeti Bank, sisaknezs@mnb.hu, banhegyip@mnb.hu. With cooperation of Péter Szabó (Hungarian Central Statistical Office).

globalization make it necessary to clear more exactly the important issues of application of the new statistical standards related to topics as global production and sale chains or off-shore activities (see *e.g.* the activity of Task Force on Global Production).

At national level these developments should drive to better use of available data and, as a consequence, the closer coordination between NCB's and NSI's. The requirement is that their data should be more reconciled in order to decrease inconsistencies between financial and non financial accounts at micro level and especially in balance of payments.

In the case of Hungary the important field of the cooperation was the compilation of goods and services data in the balance of payments upon data collected by the HCSO. Methodologically the difference between them and settlements based data on goods and services (collected by the MNB) should be mainly trade credits. After the Hungary's joining the EU this calculated trade credits moved in one way, became larger step by step and so unrealistic. To make balance of payments data more transparent, in 2005 the MNB moved the main part of these trade credits to net errors and omissions. Reclassifying turnover in trade credits into statistical errors and its size automatically led to a significant increase in the gap between net financing liabilities calculated from the bottom up and from the top down in the balance of payments. This discrepancy was observable continuously in the MNB's press releases and published balance of payments data.

At the same time a special cooperation has started with the HCSO for detecting the source of this phenomenon. It came to light soon that this issue concerns both the national accounts and the balance of payments. It was and still is a common interest to handle this problem adequately and to start micro level quality exercise.

An important source of company level discrepancy was found: the appearance of VAT registration in Hungary. After joining the EU, the community law made it easy for non resident firms to establish in Hungary registrations only for VAT purposes, without physical presence. The turnover through these registrations does not belong to Hungary but since the intra EU foreign trade is based on the Intrastat system and its register links to VAT register (at national level), these VAT registrations will be data providers for NSI's while there is no settlement transaction reported by them since they belong to non resident enterprises and physically they do not operate in the resident economy.

The practical problem is with them that the turnover of these registrations usually has (large) positive balance while it does not form a part of the resident GDP.

In Hungary in the first step the HCSO made a macro level estimation for this difference and it was released first in 2008 in the balance of payments and national accounts. This, however, raised the broader question of the accessibility of the MNB to foreign trade goods and services data collected by the HCSO at company level.

From 2008 onwards the data check at company level could be extended to the majority of financial accounts and related incomes. The main principle is that at least for the larger companies the balance of their turnover in the reported balance of payments data follow some kind of pattern, in an optimal case almost zero balance (if *e.g.* they have important activity in Hungary but their all assets and liabilities are held *vis-à-vis* non residents). In many cases, however, the goods and services data collected by the HCSO would be missing from this check if there were no legal accession for MNB to them.

In 2009 the MNB has received legal access to these individual data collected by the HCSO.

Specially in the case of VAT registrations it became clear that a macro level estimation is not enough to handle this issue since there are different relative balances throughout these registrations and at least some (in optimal case all) of them should be estimated at individual level. The relevance and importance of these issues altogether in Hungary is shown in the following table.

Trade balance revisions and its effect on NEO in Balance of Payments — Table 1

	2004	2005	2006	2007	2008	2009	2010	2011
% of GDP	–0.5	–1.3	–1.9	–2.1	–2.1	–2.7	–3.5	–4.7

The work is going on as regards further refinement of the estimation related to VAT registrations. The main result, however, along with the important reduction of statistical errors in balance of payments and national accounts, is that from 2012 onwards the experts of the MNB and the HCSO analyze together regularly the data for large enterprises or enterprise groups in order to make the quality of data in different macroeconomic statistics better.[2] It should, however, be told that the original problem is not totally solved, there are still errors and omissions remaining after this adjustment, although it was undoubtedly an important step forward.

At last, but not at least, this issue raises the question whether foreign trade data themselves, at least macro level, should be also adjusted in national publications because of VAT registrations in order to be more appropriate for users' needs. Since foreign trade data are used in addition to balance of payments and national accounts purposes among others for short term indicator for GDP, users other than official statisticians may address this issue to compilers of foreign trade data.

Concluding remarks

The reduction of inconsistencies in macroeconomic statistics due to VAT registrations may help users of these statistics, for using micro data and switching different databases will be in the future a commonly applied method of statistical analysis serving monetary and fiscal policy. The need for more sophisticated use of statistical data highlights the responsibility of statisticians of producing and at the same time the use the comparable and well documented methodology of data.

[2] Moreover, experts of the MNB released the first statistical research based on comprehensive micro level dataset (see Mariann Endrész, Győző Gyöngyösi, Péter Harasztosi, Currency mismatch and the sub-prime crisis: firm level stylized facts from Hungary; MNB Working Papers 8, 2012).This paper investigates currency mismatch in the Hungarian corporate sector. The analysis is based on the use of both foreign trade statistics and balance of payments data at company level.

References

United Nations Economic Commission for Europe (2011) "The Impact of Globalization on National Accounts", United Nations, New York and Geneva, 2011.

Bánhegyi P. (2011) "Real or Mock Surplus? Uncertainties in Measuring Goods and Current Account Balance", 58th World Statistical Congress, Dublin (Session IPS110), 2011.

Sisakné Fekete Zs.(2011) "Treatment of VAT registration in BOP and NA, Uncertainties in measuring goods and services in different statistics" OECD Working Party on International Trade in Goods and Trade in Services Statistics, Paris, 7 November 2011.

Ahmad N. (2012) "Global Value Chains OECD Work on Measuring Trade in Value-Added and Beyond", STD/CSTAT/WPNA (2012)17.

Johnson R. C., Noguiera G. (2012) "Accounting for Intermediates: Production Sharing and Trade in Value Added", *Journal of International Economics*, 86, 224–236.

Bánhegyi P., Szabó P., Sisakné Fekete Zs. (2013) "Adjustment of Domestic Value Added – Trade Balance Example" The 59th World Statistics Congress Hong Kong, China, 25–30 August 2013.

The use of micro-data to enhance the compilation and analysis of BoP statistics: the case of South Africa

Stefaans Walters[1]

Introduction

Sound macroeconomic policy formulation depends heavily on the availability of timely, reliable and accurate statistics. Traditionally, estimates of balance-of-payments aggregates are based on data obtained from international trade statistics, information acquired from international transaction reporting systems, statistics collected via enterprise surveys, administered records and/or a combination of the aforementioned. Since statistical discrepancies arise in most cases from the incomplete coverage of transactions, misclassifications, different timing and asymmetric valuations, statistical agencies or central banks can never afford to become complacent when it comes to the quality of available data sources. In order to enhance the compilation and analysis of official balance-of-payments statistics, data collection systems have to be reviewed and expanded on an ongoing basis to ensure that potential data gaps are identified and addressed in a timely and efficient manner. At the same time new or alternative data sources could provide additional checks and balances, a precondition for the compilation of reliable and accurate balance-of-payments statistics.

The South African Reserve Bank (the Bank) is officially responsible for compiling balance-of-payments statistics in South Africa. This brief presentation will focus mainly on the Bank's practical experience in using micro data to supplement existing data sources in recent years. More specifically, attention will be paid to the use of micro data in the computation of international trade in services, primary income payments to the rest of the world and the compilation of inward direct investment statistics.

International Transactions Reporting System (ITRS)

In South Africa, like in many other countries, the International Transactions Reporting System (ITRS) is a by-product of past (and existing) foreign-exchange control regulations. In terms of the system, Authorised Dealers in foreign exchange (*i.e.* resident banks and bureaux de change – Authorised Dealers with limited authority) report the settlement of transactions between South African residents and non-resident parties to the Bank on a daily basis (indirect reporting). In addition, the system also makes provision for the direct recording of foreign-exchange transactions between major domestic companies and their overseas counterparties for which prior approval was obtained from the Bank (direct reporting). In view of the fact that South African residents (*i.e.* individuals or

[1] Research Department, South African Reserve Bank, Stefaans.Walters@resbank.co.za.

enterprises) are still subject to exchange control regulations, all international settlements have to be made via local Authorised Dealers in foreign exchange.

In accordance with official government policy to liberalise exchange control regulations gradually – the Bank took a decision in 1999 to broaden the responsibilities of the then Exchange Control Department (now known as the Financial Surveillance Department) in the Bank to collect a more comprehensive set of data from reporting entities. What was significant was the fact that the detailed set of information that had to be collected from Authorised Dealers in foreign exchange was better aligned to meet the requirements of balance of payments manuals to assist compilers in the computation of balance-of-payments statistics. At the same time, a decision was taken to replace the previous system which captured transactions related to the purchase and sale of foreign currency manually to an electronic (same source) reporting system. International experience has shown that with the liberalisation of exchange controls, a reporting and compliance vacuum could easily develop if existing systems were not timeously aligned to serve as a complementary source of information for balance-of-payments purposes.

After extensive international research, the Electronic Data Interchange for Administration, Commerce and Transport (EDIFACT) was found to be the most suitable international standard of a paperless process to report balance-of-payments transactions. EDIFACT, as an electronic data interchange standard (EDI), has the advantage that data are extracted from the accounting systems of Reporting Entities, thereby eliminating the manual and dual capturing of transactions while at the same time speeding up the reporting process to the Bank. Hence, the Bank adopted the EDIFACT standard (incorporating the balance-of-payments Customer Transaction Report Message (BOPCUS)). In 2001, the electronic Cross Border Foreign Exchange Transaction Reporting System (ITRS) was introduced, which replaced the previous-paper driven system, as a compulsory reporting system to be used by all reporting entities based on the EDIFACT international standard. The reporting entities report – by means of BOPCUS, in Extensible Mark-up Language format (XML) – all the details of payments made to foreign parties by South African residents as well as payments received by South African residents from foreign entities, irrespective of the value of the underlying transaction. The ITRS was further enhanced by the introduction of a simplified BOPCUS version 2 in 2004 to accommodate additional reporting requirements of primarily balance-of-payments compilers.

With the subsequent enhancements to the reporting system, the Bank managed to create a world-class, cost-effective reporting system which provides

- accurate (same source) and up-to-date information that is validated via the Bank's Data Exchange Architecture (SARBDEX);
- comprehensive detail of each and every transaction, including the names, surnames (of individuals), addresses, identity numbers, company registration numbers, telephone numbers and the physical addresses of transactors;
- the timeous receipt of information – value date plus two days;
- reliable and stable 24-hour submission time where all files are logged and acknowledged;
- guaranteed confidentiality by using Internet encryption, SARBDEX senders validation and SWIFTNET;

- the elimination of time-consuming verification and balancing of transaction records against separate returns by using an automated balancing module;
- the virtual elimination of paper in the system; and
- legislation to enforce the reporting process.

In keeping with most recent developments in an ever-changing financial environment, it is envisaged that further enhancements to the technical specifications of the system will be implemented by the end of 2013, requiring an additional costly investment in this regard.

International trade in services

The comprehensive measurement of international trade in services is complex, especially if balance-of-payments compilers have to rely on sample surveys which are often, due to insufficient coverage, not representative. Unlike enterprises importing and/or exporting merchandise on a regular basis, transactions in international trade in services undertaken by domestic entities may occur on a fairly infrequent basis. Although data recorded by an ITRS reflect the settlement of transactions in services between residents and non-residents, the accurate classification of transactions remains a major challenge. By analysing the value of transactions per reporting category, for example, in descending order, specific enterprises could, in the first instance, be identified alongside the verification of the appropriate classification of transactions. Second, by questioning certain transactions, confirmation could also be obtained that different types of transactions were not lumped together. Once verified, information obtained for specific categories in the ITRS could be used to construct mini-sample surveys, allowing compilers to collect more detailed information from a group of enterprises and/or individuals in accordance with the methodological guidelines contained in various manuals.

Primary income payments to the rest of the world

The computation of dividends declared by listed and unlisted South African companies to resident and/or non-resident shareholders is done in South Africa by making use of different methodologies. For companies that enjoy a primary listing on the Johannesburg Stock Exchange, information pertaining to the issued share capital and the dividends payable per share are available publicly. Based on the distribution of shareholding between resident and non-resident investors, that is, information obtained from share registrars and other sources, the value of dividends payable to non-resident shareholders can be determined.

Dividend declarations by unlisted domestic companies are more difficult to account for in balance-of-payments statistics because in many cases unlisted companies do not have an established dividend policy. In addition, dividend declarations by this group of companies may not necessarily reflect underlying economic fundamentals or be aligned with the underlying phases of the business cycle of an economy. Information obtained from an ITRS, however, allows balance-of-payments compilers to identify unlisted domestic companies (fully or partially

owned by foreign companies) which, on an irregular basis, pay dividends to foreign shareholders. The system also affords balance-of-payments compilers the opportunity to verify the transaction details.

Inward direct investment statistics

The Bank has been collecting data on inward and outward direct investment statistics by using sample surveys since 1956 – quarterly surveys which often contain preliminary estimates are reconciled on an annual basis with audited balance-sheet information furnished to the Bank in annual sample surveys. To ensure that samples remain representative, Censuses of Foreign Transactions, Liabilities and Assets are conducted at regular intervals, such as every five to ten years. For intervening years, sample surveys are then utilised to determine capital flows, opening and closing stock balances and valuation adjustments. The success of maintaining representative sample surveys depends largely on the ability and capacity of compilers to keep track of changes in the economic landscape on a continuing basis. In achieving this objective, the Bank's Balance of Payments Division has recently embarked on an exercise whereby it tracks, on a micro basis, transactions in mergers and acquisitions, new inward direct investment transactions and disinvestment transactions. By focusing on individual group structures, before drilling down into transaction details, a vast amount of time and effort is saved in identifying the appropriate companies to be surveyed for various purposes.

Conclusion

The management of statistical processes to produce quality statistics remains a difficult and challenging task. In this regard, the more active use of micro data has in recent years made a significant contribution towards enhancing the compilation and analysis of balance-of-payments statistics in South Africa. Policy makers in need of timely and frequent statistics have also benefitted tremendously from the availability of micro-data.

How granular should data be to compile BoP statistics?

Carla Marques[1]

Abstract

The Balance of Payments (BoP) and International Investment Position (IIP) summarize the economic relationships between residents of a certain economy and non-residents, thus providing an integrated framework for the analysis of an economy's international economic relationships. It is a complex reality which is reflected in a variety of different users and requirements and a demanding compilation system.

The micro-databases, and specially the integration of different micro-databases, might contribute to favourable balance the equilibrium between costs and quality, resilience and robustness, quality management and analysis. However, in terms of the final required output, each economy needs to disclose aggregated information about its economic relationships with other countries or economies. So, one can argue the need of collecting micro data, with all its inherit costs, when the required output is at an aggregated level.

The aim of this paper is thus to contribute to the discussion on the level of granularity within the BoP and IIP statistical compilation system. The paper is organized as follows: in the first section is presented a definition of micro-data. Section 2 refers to the Portuguese experience with the micro-data compilation. Section 3 discusses the advantages and disadvantages of providing micro-data. Finally, some final remarks are presented.

[1] Statistics Department, Banco de Portugal, csmarques@bportugal.pt.

1. How to define micro-data?

The term "micro-data" denominates a concept which can be used in different situations and applied to different purposes. In statistical terms, micro-data are defined by The Organization for Economic Cooperation and Development (OECD) as being "an observation data collected on an individual object – statistical unit".[2]

The first step to define micro-data is the identification of the relevant "statistical unit". Within BoP/IIP "statistical unit" can be considered in different ways: the resident non-financial corporations, the economic and financial transactions, for example. In each case, there must be a key that unequivocally distinguishes each one of those units for which a set of characteristics can be stored in a database, which would configure a statistical micro-database. Ultimately, if there are different micro-databases for the same statistical unit, each one containing different characteristics, a more complete micro-database could be created under the condition that the identifying key is the same in those databases.

Multiple data sources are used in the compilation of the Portuguese BoP/IIP statistics. Many of them can be seen as micro-databases and for those under the responsibility of the Statistics Department of Banco de Portugal the statistical unit is the same, namely resident entities, identified by their fiscal number. This is an advantage in the sense that it supports an integrated approach of the statistics produced. Also, there has been an effort to create shared reference tables, containing all the relevant dimensions (not only in order to classify the entities, but also in terms of institutional sectors, instruments, maturity,...), thus responding to the needs of the different units of the Statistics Department.

2. The Portuguese experience: direct reporting for BoP/IIP and integration with other systems

As it was previously mentioned, the compilation of the Portuguese BoP statistics' is derived from several data sources. Some of them are under the responsibility of the BoP Unit while others are under the responsibility of other Units of the Statistics Department or even by other departments of Banco de Portugal; finally, other data sets are managed by other entities (*e.g.,* Statistics Portugal, Eurostat or the International Monetary Fund).

These data sources have different levels of granularity. The data collected by Banco de Portugal have a higher level of detail whereas data produced by other entities are transmitted to the BoP Unit on an aggregated basis, being the quality of these data mainly ensured by the providers. Under this framework, the inherent information granularity depends on the data sources and it is possible to find BoP/IIP items compiled with data with different levels of detail.

Currently, Banco de Portugal is changing from an International Transaction Reporting System (ITRS) approach to a direct approach, where the resident

[2] Economic Commission for Europe of the United Nations (UNECE), "Terminology on Statistical Metadata", Conference of European Statisticians Statistical Standards and Studies, No. 53, Geneva, 2000.

companies report information on their operations with the rest of the world, on a monthly basis. The quality control designed to ensure the robustness of the statistics will be based on bank settlements, taking advantage of the fact that banks will continue to report information on international settlements on an operation-by-operation, client-by-client basis (although without the statistical classification of these operations).

In both cases (direct reporting and bank settlements) the result is a micro-database where the key identifier is the fiscal number of the resident entity. The fiscal number is maintained throughout the different phases of statistical production (namely acquisition, production and exploration and, in the near future, in the dissemination phase). Namely, the integrations of these two databases serves various purpose: in the acquisition phase, the settlements concerning each individual entity are made available for that entity as a support for the submission of the data; in the production phase, to support the identification of non-responses, the estimation and the quality control; in the exploration phase to integrate other statistical data, and; in the dissemination phase to provide each entity with an approximation of its individual BoP. Additionally, the dimension of our economy is such that it is possible that a few operations have a great impact in a specific item of the BoP. Therefore, with this level of detail, the analysts have knowledge about the entities performing such operations, complementing it with other information. However, the analyst must be aware that there is a trade-off between the advantage of performing a micro level analysis and the risk of losing the focus of the macro perspective.

Due to the characteristics of some financial instruments, it was necessary to create specific surveys, particularly for Financial Derivatives and Direct Investment. In the case of Financial Derivatives, the survey is reported on an operation-by-operation and investor-by-investor basis. The information currently collected through the above mentioned survey goes far beyond that required for BoP purposes. The survey also asks information about the type of derivative, its purpose (*e.g.* risk hedging, speculation or arbitrage) and the notional amount, which are essential features to understand this instrument and its inherent potential risk. Also for the reporting institutions, it is easier to report detailed data, as long as the details are readily available within their IT system. The complexity underlying this instrument and the high level of risk involved in its negotiation makes it possible to anticipate that, in a short time, there will be a growing interest about data characterizing this reality. In fact, some specific requests by analysts were answered without the need to ask the reporting entity for further data.

In the case of Direct Investment, there is an annual survey (*Inquérito sobre o Investimento Internacional* or ISII) where the statistical unit is each share of each company, identified by the Fiscal number, thus originating a very complete micro-database, including qualitative information. This survey also provides the necessary information to design group structures and to ensure that the direct investment operations are considered in the statistics in a correct, complete and balanced fashion.

The fact that the database created from ISII is at the enterprise level, identified by the Fiscal Number, allows its connection to other databases, namely administrative databases, such as the *Informação Empresarial Simplificada* (IES), which contains accounting information. By linking this two databases it is possible to have more comprehensive information, without overburdening the reporting

entities (for example, within ISII no information is asked about own funds or the percentage of foreign capital, which are obtained in IES).

Other internal micro databases are used. Securities Statistics Integrated System (*Sistema Integrado de Estatísticas de Títulos* – SIET) is a database with data on a security-by-security and investor-by-investor basis, used to compile Portfolio Investment. The integration of this database within the BOP/IIP compilation system has numerous advantages: in terms of quality control (e.g. by crosschecking with the direct reporting and by insuring that there is no double counting between direct investment and portfolio investment); in terms of production of both portfolio assets and liabilities, for which the micro data by ISIN code is essential to apply the residual approach. Besides the statistical compilation perspective, there are analytical advantages of having information with this level of detail. For example, it revealed to be essential to comply with additional data requests by the IMF, under the Economic and Financial Assistance Program.

Resident banks' balance sheets are another example of a database under the responsibility of other unit of the Statistics Department of Banco de Portugal that is used, namely for the compilation of Banks' Other Investment. Keeping in mind the financial instrument Loans, and regardless of the institutional sector, if each country had a Central Credit Register with a harmonized level of detail and with a key identifier (like the future Legal Entity Identifier), it would be possible to obtain mirror comprehensive data about Other Investment liabilities.

Other data, produced by entities other than Banco de Portugal and appropriated on an aggregated level, are used. Statistics Portugal provides international trade statistics, used for the compilation of the Goods account, and also tourist activities' statistics, used for the compilation of the Travel account. These pieces of information are provided to the BoP Unit without any reference to the units to which they refer to, thus limiting the analytical value and the integrations with other data sources (also in terms of quality control).

3. Advantages and disadvantages of micro-data

The use of micro-data is advantageous as it allows the production of different outputs using the same information. In this respect, it is easier to fulfil new methodological requirements or additional requests without having to ask the reporting entities for more data, giving the compilation systems the flexibility and resilience to adapt to a developing environment. Additionally, If there are several databases with different information about a specific statistical unit (a company, for example), it is possible to achieve a comprehensive database for those units, as long as there is a common key identifier. This data exchange possibility makes it unnecessary to have all the detail about a statistical unit in a single database. However, it is necessary to ensure that the different micro-databases share not only the same key identifier but also the same conceptual references (e.g. financial instruments and institutional sectors). Also in terms of the human resources structures there is a demand for cooperation and coordination.

In addition, the quality control process is more exhaustive but also more demanding in terms of human and IT resources. Also, it might exist a "micro-database trap" risk, due to a tendency to focus on the micro-data and lose sight of the macro perspective.

In cases where there are integrated micro-databases, there should be a clear responsibility delimitation in terms of the statistics produced in order to avoid the risk of conflicts, and a reinforcement of the management procedures.

On the reporting entities side, it is always possible to identify a burden associated to the need to respond to statistical requirements, and it is not clear whether it is easier to report micro-data or aggregated data. Nonetheless, in order to lower that burden, the compiler should make an effort to approximate the requirements to the information available within the reporting entities, taking up the cost of bridging the gap between that data and the statistical BOP/IIP requirements. The integration of different micro-databases might also contribute in the sense that no information is asked twice.

4. Final remarks

The virtues of using statistical micro-databases, and integrating different databases, are easily understood, with aspects like flexibility, the enhancement of analytical capabilities and a better understanding of the realities portrayed at the top of the list. In the Portuguese BOP/IIP case, the recognition of such advantages has sustained the development of a new business intelligence architecture based on micro-data across the different phases of the statistical production (acquisition, production, exploration and dissemination). However, as there are costs involved in terms of IT requirements, human resources and also in terms of burden for the entities providing the information, it is necessary to maintain a balance between the required information and the costs involved.

Challenges in moving to micro-data based external sector statistics – Japan's case

Hidetoshi Takeda[1]

1. Increasing and fast-changing needs over data on cross-border activities

Globalization, rapid changes in global economy, and recent financial crises increased needs for detailed, comprehensive, and timely external sector data. Required data tend to be more and more granular, and counterpart information is considered as vital (e.g., G20 Data Gap Initiative). Also, these needs change quickly reflecting changes in relevant situations.

Recent changes of international standards and requirements reflect these data needs (BPM6, BD4, MSITS, CDIS, and enhanced CPIS).

2. Source data

Source data need to be more granular than before to satisfy the granularity of the required data, and to capture flexibility to react fast-changing user needs.

These granularity of source data can be achieved by (1) more granular data reporting, or/and (2) accessing existing micro data that are collected for non-statistical purposes.

1. More granular data reporting (e.g., security-by-security reporting system for portfolio investment)
   - Need to consider reporting burden, though granularity of data to be reported does not necessarily results increase in reporting burden.
   - In many cases, need to change statistical regulations. Therefore, needs time and administrative procedures.
2. Access to existing micro-data collected for non-statistical purposes (various administrative records, such as tax records, and supervisory/regulatory data).
   - Need to prepare legal framework for the data sharing.
   - The data sharing needs to be understood and accepted by general public.

[1] International Department, Bank of Japan, hidetoshi.takeda@boj.or.jp.

3. Japan's current situation

1. Data reporting system
   - Mixture of ITRS and direct reporting. Direct reporting gather mostly aggregated data. Micro data can be obtained via ITRS reports.
   - No s-by-s reporting system for portfolio investment.
2. Access to existing micro-data collected for non-statistical purposes
   - Very limited. Public institutions usually very reluctant to share their data with other institutions for statistical purpose.
   - General public also do not expect the administrative data related to theirs are used by other institutions beyond the primary purpose.
3. Some backgrounds
   - Highly de-centralized statistical system
     - Most public institutions compile statistics, but they have been little experiences to cope with each other. Most of these statistics are compiled for the compiling institutions themselves.
     - Legal base for these statistics are prepared only for the compiling institutions. Confidentiality requirements for each of such statistics prevent data sharing with other institutions.
   - BOJ, which compiles most of external sector statistics, is under the same situation as other public institutions.
     - No access to most of administrative records, including tax records, and supervisory/regulatory data.

4. Way forward

Unrealistic to make total review within a short period. Maybe 10-year-project.

Without control tower of overall statistical system, challenges only by external sector compiler may face difficulties.

Strategies for the foreseeable future could be;

1. Make use of ITRS
   - Micro-data on the transactions of individual company can be obtained via ITRS report. A database of such information would be useful.
2. Introduction of s-by-s reporting system for portfolio investment.
   - Big challenge, but can put the target at the time of introducing the BPM7.
   - Start study on pros and cons, communication/education of users now.

The Bulgarian National Bank's experience with loan-by-loan data collection for the compilation of external statistics

Emil Dimitrov[1]

Abstract

The purpose of this paper and the related presentation is to prove that microdata are very important for statistics. However, statistical data (including that based on microdata) needs first a common methodology, so as to be compiled with the expected reliability and quality and to be comparable among countries.

[1] Statistics Department, Bulgarian National Bank, dimitrov.e@bnbank.org.

Introduction

It is a general rule that reliable statistics are a cornerstone for evidence-based policy making. In a period of global financial crises and increased uncertainty, the general public is confronted with a loss of confidence in the economy, and the users look more carefully at the published data. Data quality is especially important in areas linked to the package on economic governance (the so-called "Six-pack"), particularly the indicators in the Macroeconomic Imbalances Procedure (MIP) Scoreboard. Under the Alert Mechanism, the quality of data becomes very important at both national and European levels: users at the national level are very much interested in knowing whether data among countries are statistically comparable; users at the European level want scoreboard data among countries to be comparable, as this increases the reliability of the Alert Mechanism Report (AMR) and the whole MIP mechanism. Moreover, international investors and rating agencies are willing to use AMR's reports for their analysis if they consider statistical data sufficiently reliable.

The European Statistics Code of Practice and data quality

The European Statistics Code of Practice (CoP) is based on 15 Principles covering the institutional environment, the statistical production processes and the output of statistics. Among them there are key principles that, from the producers of data viewpoint, are very important to assure high quality data – "Principle 7. Sound Methodology" and "Principle 8. Appropriate Statistical Procedures".

Principle 7, in practice, has requirements to both levels of statistics – European (Eurostat and European Central Bank) and national (statisticians from the NSIs and from the NCBs). According to the indicators of this Principle 7, both Eurostat and ECB have to make sure that there are European and international standards, manuals and guidelines for the compilation of the respective statistics and, in case something is missing, to take the necessary steps to develop them. On the national level, statisticians should implement the methodologies prepared by the international organizations, including ECB and Eurostat.

Observing the requirements of Principles 7 and Principle 8 is a precondition for meeting the following Principles of the Code of Practice that are related to the user needs: "Principle 12. Accuracy and Reliability", "Principle 13. Timeliness and Punctuality" and "Principle 14. Coherence and Comparability". The latter principle clearly requires common methodology, standards and definitions for statistical data to be comparable among countries, which is a key precondition for the preparation of AMR.

External statistics and sound methodology

External statistics has two indicators in the MIP Scoreboard: "Current account balance" and "Net international investment position" (both as a percent of GDP).

The compilation of both indicators is based on a solid methodological base – SNA 2008[2] (SNA 1993), ESA 2010[3] (ESA 1995) and BPM6[4] (BPM5). In addition, external statistics also have to comply with "External Debt Statistics: A Guide for Compilers and Users" (2003), as well as four other methodological guides and manuals.

On a national level Art. 42 of the Law on the Bulgarian National Bank (BNB), Currency Law, Ordinance No. 27 of the BNB on the Balance of Payments Statistics (BoP) and other documents underpin the strong methodological requirements. Therefore, external statistics fully meet the requirements of the key CoP "Principle 7. Sound Methodology".

Establishment of a loan-by-loan micro-database

The loan-by-loan (l-b-l) external statistics micro-database was developed by BNB in the mid-90's for ESA 95 sectors S.11 (Non-financial corporations), S.12 (Financial corporations) and S.13 (General Government), for two purposes: first, for the compilation of external statistics – BoP, external debt and, later on, international investment position (IIP); and second, for BoP and external debt projections. At that period Bulgaria had to develop its external statistics from scratch. During the same period Bulgaria had several IMF stand-by arrangements that needed external data.

The establishment of an l-b-l external statistics database has been fundamentally driven by the lack of almost any regular statistical information in this field. This was behind an important decision – the database would cover the whole population, as there was no registry with information regarding resident/non-resident financing.

The main steps for the development of the external statistics followed the common rules for creating any statistics. First, we had to develop a methodology based on the requirements of BPM5. The methodology was later amended, following the publication of the "External Debt Statistics: A Guide for Compilers and Users" (2003). Second, we amended the Currency Law by establishing the legal obligation for any company, bank and non-bank corporation, to register at the BNB every loan received/granted from/to non-residents (above a certain threshold). After the adoption by the Governing Council of the BNB of Ordinance No. 27 on the

[2] The System of National Accounts, 2008 (2008 SNA) is a statistical framework that provides a comprehensive, consistent and flexible set of macroeconomic accounts for policy making, analysis and research purposes. It has been produced and is released under the auspices of the United Nations, the European Commission, the Organisation for Economic Co-operation and Development, the International Monetary Fund and the World Bank Group. 2008 SNA is an updated version of the System of National Accounts, 1993 (1993 SNA).

[3] ESA 2010 stands for European System of Accounts 2010. Following the 1993 SNA revision, there was a need to revise the European System of Accounts (ESA 95) set up by Regulation (EC) No 2223/96 in order to take into account these new developments so that that system constitutes a version of 2008 SNA adapted to the structures of the Member States' economies, and so that the data of the European Union are comparable with those compiled by its main international partners.

[4] BPM6 stands for Balance of Payments and International Investment Position Manual, 6th edition. It updates the 5th edition published in 1993 (BPM5), providing guidance to IMF member countries on the compilation of balance of payments and international investment position data. The production of BPM6 was conducted in parallel with the update of the OECD Benchmark Definition of Foreign Direct Investment and the 2008 SNA, to maintain and enhance consistency among these manuals.

Balance of Payments Statistics (including the design of reporting forms and instructions), the "L-b-l Registration Form" and the "Reporting Form" were designed. The next step was to develop the procedures for data checks (quality checks) and data compilation. At a later stage we started to develop an IT system for data processing and management of the micro-database.

Advantages of a loan-by-loan external statistics micro-database.

The micro-database reduces the respondent's burden as the required information is in practice extracted directly from the accounting reports of the companies. It allows cross checking of input data as well as enhanced quality control. The micro-database provides better data interpretation and responsiveness to *ad hoc* data requests. The Interlink of various micro-databases may disclose additional information that would contribute for improved data quality as well as for better external debt and financial stability analysis.

Conclusions

The recent developments on Europe set additional requirements for data quality. Before the crisis relatively more attention was paid to the quality of the indicators presenting the development of the European Union. However, the current financial crisis made changes in these priorities. Policy-makers in Europe, national policy-makers and the public now look more carefully to the national data as country comparability became more important. This also was enforced by the establishment of the MIP, part of which is the Annual Scoreboard Indicators and AMR.

The demand for high quality country data increased significantly. Micro-data are a precondition for high quality statistical indicators. However, the production of reliable data needs also sound methodology. Attempts to create legislation without paying attention to the need of sound methodology (for every indicator), avoiding the responsibility of Eurostat for the development of such methodologies, emphasizing only the responsibility of the national statisticians for quality of data and – last, but not least – asking for inventories instead of preparing clear procedures for data quality, would not deliver the results required by the Council Decisions – improving the underlying statistics and ensuring data comparability among countries.

References

Casimiro, Paula. "The Portuguese CCR, Main Features and Use for Statistical Purposes". Presentation for the meeting of the WG Monetary and Financial Statistics, Brussels, 16–17 May 2013.

Council of the European Union, Economic and Financial Affairs, 3122nd Council meeting. Press Release. Brussels, 8 November 2011, p. 11.

European Commission. "Proposal for a Regulation of the European Parliament and of the Council on the provision and quality of statistics for the macroeconomic imbalances procedure", COM (2013) 342 final. Brussels, 7 June 2013.

Eurostat. "European Statistics Code of Practice", adopted by the European Statistical System Committee. 28 September 2011.

Lavrador, Isabel. "Exploring the Statistical Potential of Micro-databases". Banco de Portugal, Statistics Department, July 2010.

Pop, Valentina. "EU statistics still incoherent due to national differences", EU Observer, 8 May 2012.

Understanding money markets via surveys

Peter Neudorfer[1]

Introduction

Euro area statistics, conducted and published by the European Central Bank (ECB) in cooperation with other Eurosystem members, are in general based on national aggregates that are compiled by National Central Banks (NCBs) on basis of their own sources. Consequently the ECB fulfils its statistical function on basis of data collection processes, which do not imply any direct contact to any reporting agents. This principle has proven as an efficient and effective method for all statistical needs for monetary policy purposes. The large number of reporters in the euro area is one rational behind handling the data collection via the intermediate step of national compilers. In addition the reporting requirements in the area of monetary statistics are following (international) standards, which implies that their scope and details are quite stable over time.

In an increasing (although still small) number of cases DG Statistics in the ECB has adopted an alternative data collection approach in order to fulfil specific analytical user needs, that reach beyond traditional monetary statistics. More precisely the ECB also started collecting and processing micro data that are directly retrieved from individual economic agents. This data may for instance represent the micro basis of information that NCBs so far only provide to the ECB in form of aggregates but the data may also result from the design and operation of specific surveys by the ECB addressing individual reporting agents. A prominent example is the "Euro Money Market Survey" (MMS), which is already in operation for more than a decade. As in similar endeavours[2] the basic idea is not necessarily to cover a complete reporting population in order to derive an exhaustive measure of economic activities. Instead the information goal is rather capturing the essential qualitative developments that characterises the behaviour of the economic agents selected for the survey. Features like also covering non-traditional statistical elements (e.g. off-balance sheet activities) or the sole focus on qualitative assessments illustrate why surveys are characterized by a high flexibility in selecting the sample of reporting agents. In theory the latter could always optimally match the eventual information goal. The challenge is, however, to retrieve ex-ante the necessary knowledge how to select the sample.

As illustrated below the MMS exemplifies this typical trade-off between the flexibility of meeting a specific information target on the one side and providing resilient quantitative indictors that can be extrapolated to aggregate assessment on the other.

[1] Directorate General Statistics, European Central Bank, peter.neudorfer@ecb.europa.eu.

[2] *E.g.* the "*Bank Landing Survey*" (BLS), the "*Survey on the Access to Finance of SMEs in the euro area*" (SAFE), the "*Survey on Correspondent Banking in euro*" [see ECB website for reference] or the (non-published) "*Market Intelligence Survey*" (MIS). Not all initiatives are (directly) operated or supported by DG-S.

Main features of the "Euro Money Market Survey"

The aim of the (so far annual) MMS is to collect information on the euro interbank money market activities during the 2nd quarter of a year, thereby only including transactions denominated in euro and explicitly excluding transactions of banks with (non-bank) customers, inter-group flows and transactions with Central Banks. More precisely it records the turnover (broken down by maturity and market structure) in the:

- unsecured market;
- secured market (repos / reverse repos);
- OTC market (various types of swaps, FRA);
- short-term securities market.

Furthermore the survey also collects qualitative information on efficiency of market segments, changes in the liquidity and the breakdown of transaction by location of counterparty and trading systems.

In principle the MMS could be based on statistical sampling, e.g. at least for the euro area based on the frame of the Monetary Financial Institutions (MFIs) contributing to the monthly euro are Balance Sheet Item (BSI) statistics. The correlation between BSI inter-MFI positions and unsecured market activities as measured in the MMS would support this approach. Unfortunately the high variance of transactions other than unsecured market segments across the involved institutions has, at least so far, hampered the design of a statistical sample. In addition it became apparent that several banks outside the euro area and even beyond the EU play an important role in money market activities denominated in euro.

Consequently, the MMS reporting population depends on the voluntary participation of market participants,[3] a feature that the MMS shares with several similar initiatives. This implies a continuous change in the composition of the population of respondents, which in turn unfortunately also means that it is not possible to extrapolate data to the overall market volume or inversely to assess coverage rate of the MMS.

The attempt to survey all "relevant" players in the EU has led to a stepwise increase in the number of participating institutions ("total population") during the last decade: The 2012 round for instance comprised 172 credit institutions across Europe (all but one resident in the EU). In order to be able to compare findings with those of previous studies and to analyse long-term trends in the euro money market, a "constant panel" of 105 banks is applied for all money market studies dating back to 2002.

The "*Money Market Survey*" was launched in 2001 has since then been conducted annually (covering transactions during the 2nd quarter of a year). Every second year a comprehensive "*Money Market Study*" is being published. High level

[3] In the case of the MMS the participants a selected by NCBs on basis of their knowhow on the presence of individual banks on (local) markets. The response to the respective invitations is quite satisfying. The participating banks receive feedback on their share in the total results.

results are distributed via ECB press releases and aggregated results (indices) can be retrieved via the ECB's Statistical Data warehouse (SDW).

Way forward

The financial crisis during the last years has significantly increased the demand for information that allows a deeper understanding of various aspects of financial activities, which evidently also includes a better analysis of the money market. While the MMS is overall recognised as a unique data source on short term transactions, the need for more frequent – at least quarterly – results became evident. A respective pilot project, based on a subset of the annual participants and on a reduced questionnaire, started in Q3 2012.

From a methodological perspective plans exists to revisit the definition and management of (constant) panels, in particular to (better) control for shifts in business models, mergers, etc. among the survey participants. Furthermore the issue of statistical sampling and grossing up may be revisited.

Last but not least, as more surveys on specific financial markets topics may step by step broaden and deepen the collection of micro data, the question may be raised to which extent the linking of survey results would create synergies and thus increase the analytical power of the retrieved micro information.

References

ECB, Euro Money Market Study, December 2012.

ECB, Euro Money Market Survey, September 2012.

The use of payments data to improve monetary and financial analysis

Filipa Lima[1]

Introduction

Payments data can be relevant in a variety of domains. In the words of Gerdes *et al*. (2005), "*The information may aid in understanding the purposes for which different payment types are used, helping financial institutions, payments networks, service providers, and other payments organisations better understand and serve the public. Depository institutions can use the information to compare the relative use of payments with the relative use of payments at groups of similar depository institutions. Historical trends in the use of payments and information on patterns of substitution and replacement among payment types may aid in forecasting trends.*"

Based on Garcia (2011), the aim of this paper is to illustrate the use of payments data covering, namely, statistics' compilation, and monetary and financial analysis.

Characteristics of payments data

The Blue Book, a publication issued by the European Central Bank/Eurosystem, contains the description of the payment systems and instruments/ means of payment in every EU country. The Bank for International Settlements publishes the Red Book, which is an annual publication that provides data on payments and payment, clearing and settlement systems in the CPSS (Committee on Payment and Settlement Systems) countries.

Banco de Portugal publishes payment systems statistics covering the aggregate interbank operations processed in Portugal, either via clearing, or in gross, in the different payment instruments available: cheques, credit transfers, direct debits, bills of exchange and Multibanco (Portuguese ATM and POS network). Indicators of the overall settlement system transactions are presented by type of system, indicating quantity and value.

In general these data have the advantage of being very timely and cost-effective.

[1] Statistics Department, Banco de Portugal, slima@bportugal.pt.

I am thankful for the comments and suggestions provided by Luís D'Aguiar and Paula Menezes.

Balance of payments

In Portugal, card payments have been used for the estimation of the travel item in the Balance of Payments (BoP). As Gonçalves and Matos (2012) describe, Portugal has a favourable institutional environment in what concerns payment cards, since there is only one large company involved in processing payment cards data, SIBS – Forward Payment Solutions. SIBS and the other payments institutions report to Banco de Portugal, on a monthly basis, information related with the payments system. This report is managed by the Payments Systems Department which shares the data with the Statistics Department to use it in the compilation of travel statistics. Data include the number and value of operations performed in Portuguese automated teller machines (ATMs) and points of sale (POS) with cards issued abroad, and performed abroad with cards issued by Portuguese entities, as well as the characteristics associated with the cards and the type of channel used. This provides, on a monthly basis, a significant measure of travel expenditure both for credits and debits. In the course of 2013 a new report is expected to start, and it will provide more complete and detailed information, with an increased number of variables used to characterise the operations and a better coverage of payments' networks.

Financial integration

Payments data can also be used to measure the evolution of monetary and financial services as illustrated by Cadete de Matos and D'Aguiar (2009). The authors describe the "boom in retail financial services". The most striking one is Multibanco, a sophisticated network shared by every bank operating in the economy that fully integrates ATM and POS terminals, where, in addition to cash withdrawals and balance inquires, customers can carry out a wide range of operations through ATM (currently more than 60 different services, of which 30 also available at POS terminals), *inter alia*: money transfers (both to other customers in the same bank and to other banks), payments for utilities bills, payments to the State and the Social Security, mobile phone top-ups, transport ticketing and event booking and ticketing.

Short-term forecasting

Forecasters are always in search of new indicators that are related with the macroeconomic variable of interest and available earlier. With an application for Portugal, Esteves (2009) proposes the use of the ATM/POS data as an indicator to estimate private consumption. According to Banco de Portugal (2011), "*Speed of access to relevant information for assessing the economic situation at any point of time is highly useful to the extent that macroeconomic forecasts and policy decisions are highly affected by the quality of this assessment. It is, therefore, important to track the evolution of available indicators related with the variable whose growth is to be estimated. In this context, the use of ATM and POS data appears to be a natural option for now casting private consumption. These data on electronic withdrawals and*

payments are related with consumption expenditure and are rapidly available – usually in the first few days following the end of the month."

More recently, on a methodological review of the short-term forecasting for the Portuguese economy, Esteves and Rua (2012) also highlight, in the case of private consumption, the data referring to ATM/POS terminals.

Other purposes

On a different vein, concerning the use of ATM data, it is worth mentioning the work of Attanasio and Japelli (2002) who estimate the demand for currency using a data set that provides detailed information on the management of cash balances, interest rates on alternative assets, and the adoption of new technology. They find substantial differences in the equations for ATM cardholders and non-holders: indeed, the demand for currency of those who choose to have an ATM card is considerably more elastic to the interest rate than that of the households that do not hold such a card. The evidence they obtain from average balances is also confirmed by their study of withdrawals, trips to the bank, and types of payments for income.

Occasionally, specific surveys can be conducted which can provide additional information on payment behaviour. In this regard, Jonker (2005) provides a valuable literature review, from which the picture emerging is that the use of electronic payment cards (debit and e-purse) is negatively related with age and positively related with the educational level of consumers. Furthermore, women seem to use more different payment instruments than men. Cash is regarded as a universally accepted, but relatively unsafe means of payment, whereas the debit card is considered to be modern, easy to use and practical. This evidence is also found in Banco de Portugal (2007): "*The survey focusing on consumers produces findings in line with studies in other countries (the U.S., Belgium and the Netherlands) to the effect that the use of electronic payment instruments (payment cards) is directly related to income and education and inversely to age*".

Final remarks

The implementation of the Single Euro Payments Area (SEPA) constitutes a major window opportunity for NCBs to benefit from harmonised payments' data across Europe; SEPA is an initiative of the European banking industry that will make all electronic payments across the euro area – *e.g.* by credit card, debit card, bank transfer or direct debit – as easy as domestic payments within one country are now.

The use of payments data confirms the typical advantages of using administrative data: reduced burden on respondents and more cost-effective compilation of statistics. Similarly, these data have the advantage of being very timely and cost-effective. NCBs and NSIs should be aware of the potential usage of these data: it is an excellent example of how response burden and costs can be reduced while, at the same time, timeliness and quality of statistics can be improved. This potential should be explored by both NCBs and NSIs.

References

Attanasio, O.P., Guiso, L. and Jappelli, T. (2002). "The demand for money, financial innovation, and the welfare cost of inflation: an analysis with household data." Journal of Political Economy, 110 (2) pp. 317–351.

Banco de Portugal (2007). "Retail Payment Instruments in Portugal: Costs and Benefits", July 2007.

Banco de Portugal (2011). "Box 5.2 The use of ATM/POS data to estimate the evolution of private consumption", Annual Report 2010, Banco de Portugal, 173–176.

Cadete de Matos, J. and D'Aguiar, L. (2009). "Measuring the evolution of monetary and financial services in Portugal", 57th Session of the International Statistics Institute (ISI) / Durban, 16–22 August 2009.

Esteves, P. S. (2009). "Are ATM/POS data relevant when nowcasting private consumption?", Banco de Portugal, Working Papers 25 | 2009.

Esteves, P. S. and Rua, A. (2012). "Short-term forecasting for the Portuguese economy: a methodological overview", Banco de Portugal, Economic Bulletin, Autumn 2012.

Garcia, A. (2011). "The role of Payment System Statistics on the evaluation of economic activity", Committee on Monetary, Financial and Balance of Payments statistics, February 2011.

Gerdes, G.R., Liu, M.X., Parke, D.W., Walton II, J.K. (2005). "Trends in the use of payment instruments in the United States". Federal Reserve Bulletin 91, 180–201.

Gonçalves, P. and Matos, B, (2012). "Integrating data sources and frameworks for compiling travel", Working Party on International Trade in Goods and Trade in Services Statistics, Organisation for Economic Co-operation and Development, October 2012.

Jonker, N. (2005). "Payment Instruments as Perceived by Consumers – a Public Survey", DNB Working Paper, No. 53/September 2005.

Is micro-data reporting the future for interest rate statistics?

Patrícia Antunes[1], Inês Correia[2] and Rita Lameira[3]

1. Introduction

The availability and integrated management of micro-databases at the national central banks is of key importance, as it allows improving flexibility and reducing reporting burden and data overlaps.

In the field of Monetary and Financial Statistics (MFS) this gained particular significance with the recent financial turmoil and the decision to implement the Single Supervisory Mechanism. This paper presents a recent experience of *Banco de Portugal* (BdP) in launching a reporting scheme to collect data on new loans granted to Non-financial corporations (NFC) on an operation-by-operation basis.

The paper is organised as follows: in the next section, there is a brief reference to the MFS path; in the third section it is described the new micro-data set; in the fourth section some examples of analyses enabled by micro-data are commented; and finally, the last section presents the main conclusions.

2. The MFS information

According to its Organic Law[4] and whilst Statistical Authority in the framework of the National Statistical System[5] as well as member of the Eurosystem, BdP is responsible for the compilation of MFS. In this context, the Portuguese Central Bank is entitled of computing and disseminating information on the balance sheets of Banco de Portugal (BdP) and Other Monetary Financial Institutions (OMFI) and data on the interest rates applied by OMFI on euro denominated deposits and loans vis-à-vis Households and NFC resident in the Euro Area. This information is considered of utmost importance for the evaluation of the Portuguese financial system and of the economy in general. More particularly, the statistics on interest rates are essential to appraise the transmission mechanism of monetary policy and to assess the convergence of euro-area financial markets. In addition, it is a key instrument to provide insights on the financing conditions and on the deposits remunerations, to evaluate the financial margin of the banking sector and to improve the knowledge on the debt service effort of the non-financial sector.

1 Statistics Department, Banco de Portugal, paantunes@bportugal.pt.

2 Statistics Department, Banco de Portugal, ipcorreia@bportugal.pt.

3 Statistics Department, Banco de Portugal, rlameira@bportugal.pt.

4 Law 5/98 of 31 January 1998.

5 Law 22/2008 of 13 May 2008.

Following the publication of Regulation ECB/2001/18, transposed to the national context with the Instruction 19/2002 of BdP, in January 2003 the Portuguese Central Bank started to compile and disseminate interest rate statistics harmonised and comparable across the Euro Area. Within this legal frame, the interest rates data were collected from a sample of institutions (less than twenty), considered representative of the universe and which was revisited, and if necessary revised, at the end of each year. This Instruction defined the collection of information on new business and outstanding amounts and on the corresponding interest rates, with details by institutional sector, instrument, initial period of rate fixation and purpose of the loans.

Subsequent to the issuance of Regulation ECB/2009/7 and Instruction 12/2010, and in parallel with the implementation of the additional requirements associated with new operations, concerning mainly a further breakdown by initial period of rate fixation, by loan volume and whether it is totally collateralized, BdP started collecting interest rates data from the universe of the resident OMFI (around 70 institutions) instead of maintaining the sample approach.

3. The new micro-data on interest rates

In 2012, in the framework of the financial crisis and of the Financial Assistance Programme to Portugal, BdP considered necessary, for the purpose of financial stability analysis, to encompass more detailed information on the evolution of banks' credit to NFC. With this objective, it has launched Instruction 20/2012, which set up a new monthly reporting scheme for individual data on interest rates and amounts of new euro denominated loans, including overdrafts, granted by OMFI to NFC resident in the Euro Area. The new report entered into force in July, for the reference period June 2012, and provides monthly details, on a loan-by-loan basis, about the NFC fiscal number, date of the operation, contractual maturity, initial period of rate fixation, amount, annualized agreed rate, existence of collateral securing the total amount granted, nature of the loan ("totally new", "renegotiated" or "automatically renewed and other"[6]) and residency of the NFC ("Portugal" or "other Euro Area country"). It establishes that the micro-data report is only required to OMFI that reach the threshold of 50 million euro of new loans to NFC, in the aggregated report of each month.

The implementation of this new report was quite effortless for the majority of the MFI, as it requires basically the submission of the individual information underlying the aggregated report. Furthermore, if in the future the aggregated data on new operations were entirely replaced by a micro-data reporting system, it would certainly reduce the reporting burden of all OMFI and increase the quality of the data available in BdP. Moreover, the possibility of using the NFC fiscal number, which is available for all corporations resident in Portugal, to link this micro-data with the information available in other statistical databases, allows a multiplicity of analysis, namely by NACE classification, firm size (micro, small, medium or large), private vs. state-owned and exporting vs. non-exporting corporation. As mentioned in a recent study by Santos (2013), the specific features of each loan – not only the

[6] The "automatically renewed and other" loans are not considered under the ECB concept of new business.

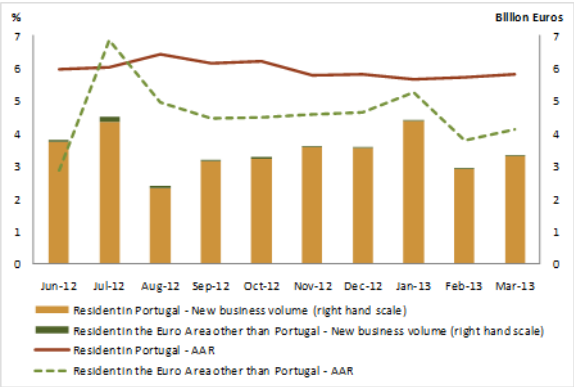
characteristics of the operation itself, but also the characteristics of the lending OMFI and of the borrowing NFC – are determinant for the interest rate applied, and these features should be taken into consideration when comparing the interest rates for different countries. Therefore, the availability of similar databases with individual information in other countries would certainly improve international comparisons.

4. The potentialities of micro-data – some examples

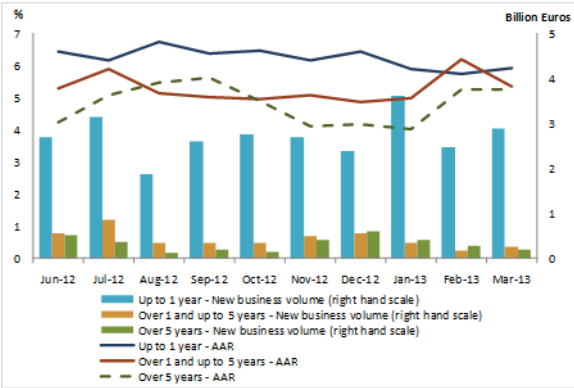
The new micro-data set comprises a monthly average of twelve institutions (with around 140 000 records), among which only ten OMFI were required to send information every month (approximately 130 000 records). Therefore, for the sake of consistency, the following analysis considers this constant sample of 10 institutions, which represents around 85% of the total monthly new loans, including overdrafts, reported by all resident OMFI in the aggregated tables. The micro-data constant sample comprises a monthly average of 44 000 new loans, excluding overdrafts, which corresponds to a total amount of 3.5 billion Euros. Regarding overdrafts, there are over 85 000 monthly operations, which stand for an average amount of 10 billion Euros.

As illustrated in Figure 1, over 95% of the loans to NFC, excluding overdrafts, are granted to residents in Portugal, even though the interest rates are, on average, higher for this segment. Considering only the operations with resident NFC, the majority of the loans (around 80%) are short term (up to one year) and the differential between long term (above one year) and short term interest rates remains negative for the whole period (figure 2). These analyses by resident / non-resident and by contractual maturity are only enabled by the current available micro-data, as the aggregated report only comprises the total Euro Area operations and the breakdown of loans by initial period of rate fixation.

New loans, excluding overdrafts, granted to resident and non-resident NFC — Figure 1



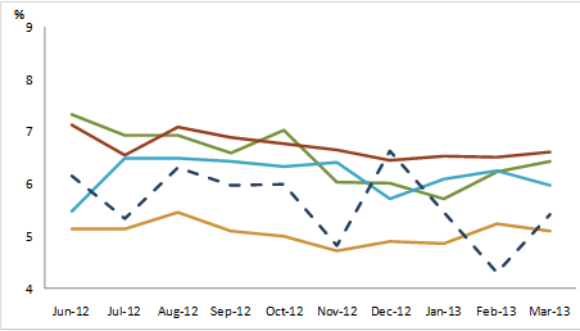
New loans, excluding overdrafts, granted to resident NFC, by maturity — Figure 2



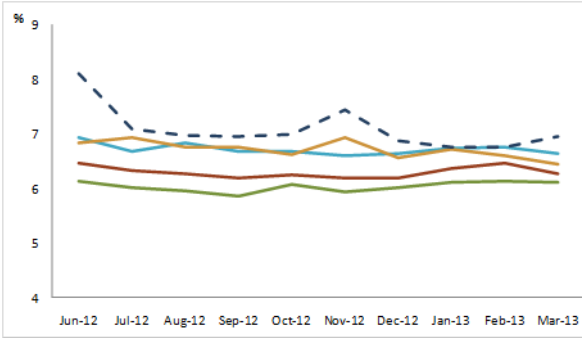
Interest rates on new loans granted to resident state-owned and private NFC,[1] by firm size

Figure 3

New loans, excluding overdrafts



Overdrafts



Private corporations - Micro
Private corporations - Small-sized
Private corporations - Medium-sized
Private corporations - Large
State-owned enterprises

[1] Excluding private non-financial holdings.

Furthermore, with the aggregated data, the loan volume was the only indicator accessible to distinguish the operations with small and medium-sized enterprises (SME) from the operations with large enterprises. In addition, the breakdown between private and state-owned enterprises was totally unavailable. Currently, linking the micro-data with other data-bases, it is possible to disaggregate the loans by state-owned and private NFC and, for the last ones, additionally by firm size. A closer look at figure 3 suggests that credit conditions are on average tighter for private SME than for private large enterprises. The opposite seems to be valid for overdrafts, as can be seen in figure 4. In what concerns the state-owned enterprises, the interest rates on the new loans do not present a regular behaviour, while in the case of overdrafts the interest rates are usually higher than the ones applied for private corporations.

5. Concluding remarks

In conclusion, according to BdP experience, the micro-data reporting presents several advantages as it allows to reduce the reporting burden, to improve the quality of the data, to enhance the economic analysis by taking advantage of the integrated management of different databases and to improve the capacity to respond to ad hoc information requests.

This type of solution should be encouraged in other countries and in other statistical fields, as for example the deposit operations and the corresponding interest rates.

References

Antunes, A., R. Martinho, (2012), "Access to credit by non-financial firms", Financial Stability Report of Banco de Portugal, May 2012, Pages 159–176.

Santos, C., (2013). "Bank interest rates on new loans granted to non-financial firms – an initial assessment of a new set of microeconomic information", Financial Stability Report of Banco de Portugal, May 2013, Pages 129–137 (in the Portuguese version).

The use of central bank databases for macroprudential analysis

Laura Vajanne[1]

Introduction

The framework for the monetary and banking statistics of the European System of Central Banks during the preparations for the EMU and during the initial years of EMU's existence focused on delivering statistics needed to conduct monetary policy. The approach was very aggregative, and main efforts were directed towards reliable and harmonised statistics for the whole euro area, even statistics at a country level were used with caution. Now, with experience of financial crisis for almost six years, Eurosystem has been forced to place much higher priority on financial stability concerns, which has had a direct impact on data requirements. Before the crisis data needs for financial stability purposes were mostly satisfied by supervisory data collected and analysed in different forums. Now the setting has changed and many euro area central banks are involved e.g. in macroprudential policy decision making. Many central banks have also increased their role in prudential supervision.

Implementation the macroprudential and/or prudential policy decision making into central banks normal procedures has increased demand for data relevant in support for financial stability analysis. The global financial crises highlighted the need for better data for the monitoring and managing of the build-up of risks to financial stability. The work to fulfil data gaps identified in this area is performed and coordinated at international level by G20 and Financial Stability Board, as well as by the ECB, European Systemic Risk Board and the European Supervisory Authorities.[2]

There are many advantages to be reached when using individual banking information in addition to aggregate level information, when central banks are monitoring financial stability or conducting macroprudential tasks. There is for instance a risk that aggregated series might hide a serious development in the same part of the banking industry. Institutional data is also essential when discussing the development of systemically important institutions required from the future Single Supervisory Mechanism. In many countries, the banking sector is also quite concentrated, and a quick scanning through the development of the most important banks already gives a relevant picture of the whole MFI-sector.

1 Financial Stability and Statistics, Bank of Finland, laura.vajanne@bof.fi.

2 A recent ECB Occasional Paper (2013) illustrates the work carried out by the European bodies.

What kind of data for macroprudential analysis

The Bank of Finland has started a process producing a common quarterly risk assessment together with the Finnish Supervisory Authority for macroprudential policy decision making. An important element in this new process is to optimize the use of all relevant information collected by the Bank of Finland and to combine information collected by the Finnish Supervisory Authority.

Referring to the databases of the MFIs, the aim is to use data collected from systemically important banks operating in Finland. We have chosen four to the six largest banks to be monitored at the institutional level in addition to aggregate statistics produced monthly. As the banking level data is confidential, the access rights to this part of the risk assessment are highly restricted.

In addition to using the balance sheet and interest rates data of the banks for macroprudential policy purposes, information from the Centralized Securities Data Base and later from Security Holding Statistics will be included in the package for monitoring financial stability purposes. Also the enhanced RIAD will include for instance information on eligible assets and close links.

When monitoring more granular data, we must be much more selective than in the case of aggregated data. There is a danger to have too much detailed information. The so called instrument approach used in macroprudential analysis can be helpful when selecting indicators to be followed since choosing instruments gives us hints what indicators should be followed. The instrument based approach starts from the instruments available for policy makers and the indicators are used to indicate that these chosen instruments are within the given boundaries. The main focus is in systemic risk aspects.[3]

At the Bank of Finland, we are monitoring

a) excessive credit growth and leverage of systemic important banks
   - measures of credit cycle (growth rate of loans; stock and new),
   - interest rates,
   - housing loans,
   - loans to non-financial corporations,
   - leverage ratios (loans to assets)
   - measures of sectoral concentration

b) excessive maturity mismatch and market illiquidity
   - liquidity coverage ratio,(liquid assets to total assets or short term liabilities),
   - net stable funding ratio (loans and other long-term assets to long term funding)
   - Loan-to-deposit ratios
   - Lending spreads

[3] See *e.g.* Committee on the Global Financial System (CGFS) report No 48/December 2012.

Conclusions

One of the unambiguous lessons from the crisis is that dangerous imbalances can be built underneath a seemingly peaceful macroeconomic surface. Inflation can be stable, output can appear to be at potential, but things may still not be quite right. Sectoral booms may lead to an unsustainable composition of output – for example, too much housing investment. Or financial risks may build up because of the way real activity is funded, for instance, excessively leveraged financial institutions, excess household indebtedness, excess maturity mismatches in the banking system or turn to off-balance-sheet products with large tail risks.

Despite progress over recent years, the development and implementation of macroprudential policies are still at an early stage. In the area of systemic risk monitoring, efforts have focused on closing data gaps and on developing better indicators and models to assess systemic risk. It is clearly seen that we need to adapt with more flexibility to changing requirements in data collections, both macro and micro data should be available, and collection systems should be able to respond also to ad hoc needs. Central banks have to improve on the communication of their statistics both internally and externally. And the issue of coordination and cooperation is crucial, within the central bank itself and at the national and international level. Cooperation between central banks and supervisors is of key importance and needs to be strengthened. Limitations on data sharing between the two organisations need to be reduced or removed, partly in order to be able to respond to some international statistical initiatives.

References

CGFS (2012). "Operationalising the selection and application of macroprudential instruments", CGFS Papers No 48, December 2012.

Israël, Jean-Marc, Sandars, P., Schubert, A. & Fisher, B. (2013). "Statistics and Indiactors for Financial Stability Analysis and Policy ", ECB Occasional Paper Series No 145 / April 2013.

Summary of the parallel sessions (A-F) and the panel discussion

Summary and conclusion of session A

Chair: Enrico d'Onofrio

Credit registers are enjoying something of a "*renaissance*". All four experiences report a host of initiatives relating to this tool, which began its history several decades ago. Some countries are trying to increase its value by harmonizing the information and using the register for purposes well beyond the original one of helping banks to manage credit relations; others are introducing innovative methods and technologies; and finally, one country is currently setting up a credit register.

Within the European System of Central Banks (ESCB), the first steps towards upgrading the role of credit registers were taken in 2005 with the introduction of reciprocal data exchanges. This process continued in the following years with in-depth analyses to understand how the registers' content could be enhanced and adapted to the statistical needs of the euro area and the European Union (EU), to alleviate the statistical reporting burden and increase transparency. More recently, the process has picked up speed with concrete recommendations to improve national central bank's granular data on lenders, borrowers and credit attributes, as well as methodological issues (ESCB Task Force on Credit Registers), and with measures to support ongoing macro- and micro-supervisory tasks with a view to the forthcoming Single Supervisory Mechanism (SSM).

The experiences of Armenia and Malaysia, although they began in different periods, are now consolidated and offer interesting insights into the management and the technical/organizational solutions adopted. The high frequency of information updating in the case of Armenia (three business days) and the manner of providing services in Malaysia (in 2011 a mobile customer service coach was set up to provide greater access to the public) are impressive.

Malaysia has also described an original method of coordinating a public central credit register managed by the central bank with private credit bureaus.

Ireland – after a first experience of constructing a database on micro credit under the pressure of the economic crisis – is now grappling with a project for an information system on credit that is due to become operational by 2015. The country is aware of the need to involve statisticians in this central credit register project, as well as in decisions regarding the type of data that will maximize its usefulness. Its primary objective, however, will be to support banks and supervisory activities.

Although they differ as to state of progress and approach, all these experiences seem to have the same goal: the development of the central credit register from a single-purpose information system to a multi-purpose one. If this is the objective, we must be aware that the path will not be easy and it will present many challenges.

As statisticians a clear vision is needed to impose a holistic approach to meeting information needs, and ultimately to manage the evolution of the central credit register as part of an integrated statistical system.

The old compartmentalized approach has now run its course and there is broad agreement that "integration" is the only paradigm capable of maximizing the value of statistical databases.

Suitable organizational solutions and consistent methodological tools and technologies are needed to support the new approach. But a new culture is also necessary, one that regards information sharing as "natural" and "normal". Only in this way will it be possible to overcome the legal, juridical and operational restrictions that characterize these systems today.

Summary and conclusion of session B

Chair: Jesmond Pulé

The session acknowledged that many central banks and other national decision-making bodies nowadays recognize the potential offered by item-by-item data collection systems. One of these solutions is the security-by-security (s-b-s) data collection system with variants such as the s-b-s investor-by-investor model, issues of s-b-s and holdings of s-b-s statistics. These micro-databases are increasingly being used for statistical purposes in the area of monetary and financial statistics, balance of payments/international investment position and financial accounts statistics and for economic analysis and research. The *Banco de Portugal*'s presentation even asserted that a level of excellence in statistical processing can be achieved with the use of such databases.

The advantages of s-b-s data are various, the most important being (a) the higher level of quality (than when statistics are transmitted by respondents on an aggregated basis), (b) reduction of reporting burden on respondents, (c) flexibility of usage, (d) prevention of data redundancy, (e) a more efficient data quality management and (f) quicker response to requests. With regards to the latter point, in fact, the Lehman Brothers bankruptcy case showed how quickly central banks can provide the required granular information on an *ad hoc* basis provided that they have at their disposal s-b-s databases. In the Portuguese case, the existence of the micro-data databases was also a key success factor for an efficient and effective response by the *Banco de Portugal* to a myriad of new data requests for monitoring the development of the financial assistance program to Portugal and the preparation of the quarterly review missions. Driven by the financial crisis of the 1990s–2000s and the most recent one, the availability of s-b-s data was also seen as crucial for monetary policy purposes. This crisis spelled out demand for such data as to help detect the imbalances that can potentially trigger instability in the financial system.

Admittedly the costs on central banks have increased since the bulk of the work has shifted from reporting agents to central banks and this was not necessarily coupled with an increase in financial and human resources. There are still certain inherent problems to reach the desired quality of the granular data, especially if the s-b-s information will one day be the potential solution for prudential supervisory requirements. Each country has its own *modus operandi* and is in different stages of developing its s-b-s systems. Moreover, countries have different operational and administrative arrangements, and sources of data. Furthermore, while some countries rely completely on the granular data transmitted by respondents, others rely on data procured from commercial data providers. Within the euro area, rather

than investing in their own databases, many countries opted to rely on the ESCB's Centralized Securities Database (CSDB) which is fed by data purchased from a number of commercial data providers coupled with input from national central banks. The latter also carry out related data quality management. Iterative processes could potentially enhance the CSDB to the quality required for prudential supervisory and macroeconomic policy purposes. However, even though this database does not have the necessary "book-keeping standard", it is still better than any individual alternative especially if "health warnings" are pre-announced. Further cooperation, especially with countries having international financial centres, would further improve the quality of the CSDB data.

Other central banks, such as the Czech National Bank, involved the supervisory body from the outset of the project to be able to satisfy the latter's needs. This co-operation also led to considerable synergies, including the avoidance of duplication in data collection. Their solution also lends itself to flexibility in the filling of unforeseen data gaps. Further plans are envisaged by the Czech National Bank in using the full potential of such a database.

In Thailand, concerted effort by the Bank of Thailand, Thailand's Ministry of Finance, other national authorities and *ad hoc* committees was the key success factor to launch the financial instrument information centre. This co-operation may even be broadened to be able to adopt common security codes, which at present are still lacking. Despite the challenges involved, Thailand's effort in this regard was still considered to be worth the investment.

As s-b-s databases improve, investment in the quality of such databases should also be coupled with enhanced staff profiles as the dimension of the work involved will change. Despite processes being shifted from reporting entities to the compilers and despite technological, administrative and operational challenges which may take some years to achieve, there seems to be unanimous agreement that s-b-s databases are still considered to be worth all the investment. Furthermore, there also seems to be a consensus towards a strategy to enhance the quality of such data for use in monetary policy and prudential supervisory policy-making. Another challenge which seems to provide further opportunities in the not too distant future is the integration of s-b-s solutions to other micro-databases, such as loan-by-loan and central credit register databases, which may even incorporate deposit-by-deposit databases.

Summary and conclusion of session C

Chair: Timur Hülagü

The session on Central Balance Sheet Databases and Business Registers exhibited the importance of business registers and central balance sheets in measuring activities of non-financial companies. It also elaborated on "whys" and "hows" on the critical features of such databases and efficient ways of dealing with them.

Participants in the session succeeded to highlight the use and importance of these databases. More information will lead to better and more correct analysis, which in turn will allow for right policy decisions. In analyzing and estimating how firms react to monetary policy decisions one needs to control for firm specific characteristics and unobserved characteristics, which is only feasible through micro-data. Aggregated data gives a picture about the whole economy or some

sector's dynamics but conclusions based purely on these have the potential to be incomplete.

To reach an accurate picture of true financial situations of firms, at the first place, it is recommended to reach accounting documents of firms. This will lead to perform impulse response analyses of monetary policy, as well as identifying risks to financial stability. It will also enable a consolidation of data to avoid the double accounting problem.

Compiling business registers and central balance sheets will also reduce the reporting burden of companies. Without them, and without an efficient system that includes greater inter-institutional cooperation and sharing, companies will continue reporting to several authorities.[1] Such inefficiency also decreases the belief in statistics by many parties.

Presentations also discussed "how" to reach a better system. The need of an easy and efficient way to get data is clear. In general, the information is not easily available. Data is scattered around several different databases managed by different authorities. Here comes the importance of a great inter-institutional cooperation. Sharing data between them is also crucial. An ultimate goal can be sharing the information to everyone, as *Banco de Portugal* is seeking currently. A legal base to support such sharing and cooperation is unavoidable. To reach this status, such a system should be approved publicly with the understanding of "its importance", its "need before falling in the next crisis" and its condition as a "public good".

At the end of the day, an efficient system which allows the compilation of granular data is needed. Then, it is possible to reach distributions rather than simply aggregates, and consolidated figures as well.

Data should also be globally standardized. The world has changed and we need internationally comparable data and data capable of being integrated. Not for company data yet, but for finance, an initiative called Global Legal Entity Identifier System has been developed and being implemented under a G-20 charter. This will enable to collect real-time, granular and globally standardized financial data. In a world where a crisis can occur within small amount of time, action has to be taken much quicker than before.

And then this brings us to the point of timeliness. It may not be easy to compile firm data at a higher frequency. Quarterly data can be submitted with a lag, compilation and analysis will take some time too. This may be too late to take the right action. We need to consider ways to achieve more timely data as much as we consider other issues.

Summary and conclusion of session D

Chair: Jean-Marc Israël

In this session Brazilian and Portuguese presentations showed mature and impressive credit registers in their respective countries. The thresholds are very low,

[1] This is painful for them, we received many phone calls from them at those bad times and, believe me, they don't call for just saying "hi".

€344 for the former, €50 for the latter – at the explicit request of credit institutions, which leads to an almost complete coverage. Important services are developed to give an appropriate feedback to the credit institutions in assessing the creditworthiness of their (potential) clients/borrowers. The credit registers are used for supervisory analyses but also for a wide range of other purposes, in particular for statistics (to check other datasets, to add new or more detailed breakdowns) or economic or financial stability analysis.

France presented a classification of enterprises by size with three different criteria (turn-over, total balance sheet assets and number of employees). It was concluded that this work helps to clarify what is at play and, in turn, may have a significant impact on the results. In this case, it showed that the share of credit granted to small- and medium-sized enterprises (SMEs) in France was higher than usually shown.

Cabo Verde and Malaysia showed two examples of the use of credit registers data. In the former case, the breakdowns currently available were not sufficient to analyze policy-relevant developments in non-performing loans. A task force investigated in greater depth what happened and delivered conclusions and recommendations on how to improve the c register for supervisory and other policy purposes. In the latter case (in Malaysia), data drawn from the credit register enabled to identify specific developments in the credit to households for housing purposes, analyze and model what was going on and take a very focused, countercyclical policy decision to avoid a bubble in the making. Results, also measured by more recent developments in the credit register dataset, show that the impact was fast and in-depth, also mitigating the underlying systemic risk.

The presentations and discussions turned clear that linked to the increasing complexity of economic and financial developments, as well as their higher volatility in the wake of the crisis, macro-economic statistics are no longer sufficient to assess credit developments. Also, measuring the distribution/breakdowns of credit – by borrowers' categories (size, activity) or by types of loans (funding need, maturity, collateralized or not) – has become of utmost importance. Granular data are, in turn, welcome by reporting agents as they enable a feedback loop for their own decision-making on credit, their internal reporting, as well as benchmarking; reporting burden is also minimized, after a one-off cost to implement in automated systems, as changes in *e.g.* statistical classifications do not affect the stability of the reporting requirements.

As the criteria for the breakdowns and distribution metrics (of credit) may change, only granular data allow for making the right aggregation at the right moment in time.[2] This is a prerequisite for conducting in-depth timely analysis and preparing policy-making – also using more accurate (macro-prudential) tools, and measuring their impact over time.

Credit registers should not be seen in isolation. They are part of an information system for central banking and supervisory purposes. A key complementary element is the business register that enables to better assess who the counterparts (borrowers) are. Other databases may also be important such as on securities

[2] Knowing that the time to fruition for large projects that entail IT developments are long, often more than 2 years, with risks that the resulting dataset does not fit the needs anymore, only granular data allow getting the relevant information on time.

(issued and held), central balance sheet offices, etc. A link is needed between these datasets via common identifiers (of lenders and borrowers; possibly also loans).

Summary and conclusion of session E

Chair: Jacques Fournier

In the session it was showed that micro-data play a major role for balance of payments compilation and analysis, as outlined in particular by the Bulgarian and Portuguese central banks. Allowing detailed cross-checking, they are particularly adapted to the balance of payments that reflect a large diversity of operations between multiple counterparts and sectors. Moreover, micro-data are necessary to answer rapidly to new questions raised by international organizations, as experienced during the crisis where financial support was often conditional to quick responses.

Still, micro-data collection is not widespread yet and may be difficult to achieve. Legal and costs barriers for accessing administrative databases can be very difficult for central banks to overcome, as explained by the Bank of Japan. Indeed, the efficient use of micro-data requires close cooperation between all concerned authorities, as underscored by the South African Reserve Bank. IT hurdles to an efficient management of complex databases have also to be kept in mind.

While these difficulties should not be underestimated, the Magyar Nemzeti Bank and the Central Bank of the Russian Federation shed light on another very important merit of micro-databases, namely the detection of fictitious transactions. Indeed, a sensible analysis of raw data, coupled with a matching of balance of payments and national accounts (or even tax information), can lead to measuring the shadow economy. In this perspective, data mirroring between countries can be very helpful.

A demanding but promising tool, micro-databases appear very useful for a number of statistical and macroeconomic purposes. However, heed should be paid to avoiding excessive costs and delivering adequate data quality.

Summary and conclusion of session F

Chair: Gregor Bajtay

The session brought together different initiatives to obtain additional data and/or use the existing data for a different purpose. In this context, the ECB, for which, in general, published statistics are based on national aggregates compiled by national central banks (NCBs), started collecting and processing micro-data that are directly retrieved from individual economic agents. The target is to capture the essential qualitative developments underlying the behaviour of economic agents selected for the survey. The challenge of these surveys, which in general are characterized by high flexibility in the selectivity of sample reporting entities, is to retrieve *ex ante* the necessary knowledge of how to select the sample and to adjust the surveys' frequency to the need of the level of understanding of various aspects researched.

Portugal's presentation illustrates the use of payments data covering statistics' compilation and monetary and financial analysis. Payments data are relevant in a variety of domains and the implementation of the Single Euro Payments Area (SEPA) brought the opportunity for NCBs to benefit from harmonized payments' data across Europe. These data are being utilized, for instance, in the estimation of the travel item of the balance of payments, to measure the evolution of monetary and financial services in the financial integration, and to estimate private consumption in short-term forecasting.

Another experience presented by Portugal was the launching of a reporting scheme to collect data on new loans granted to non-financial corporations (NFCs) on an operation-by-operation basis. Initially collecting interest rates data, afterwards it launched a new monthly reporting scheme that encompasses more detailed information on the evolution of bank's credit to NFCs. The micro-data reporting presented several advantages as the reduction of the reporting burden, improvement of data quality, enhancement of the economic analysis by taking advantage of the integrated management of different databases, and better capacity to respond to *ad hoc* information requests. This type of solution could be applied to other statistical fields, as the deposit operations and the corresponding interest rates, for example.

The presentation from Finland exposed the joint work from the Bank of Finland and the Finnish Supervisory Authority in producing a common quarterly risk assessment for macroprudential policy decision-making. An important element was the optimization of the use of all relevant information collected by both authorities, keeping in mind the importance of being selective in the eligible data and on the danger of having too much detailed information. Some features called their attention: (i) central banks have to improve on the communication of their statistics both internally and externally; (ii) the issue of coordination and cooperation is crucial, within the central bank itself and at the national and international level; (iii) cooperation between central banks and supervisors is of key importance and needs to be strengthened; and (iv) limitations on data sharing between the two organizations need to be reduced or removed, partly in order to be able to respond to some international statistical initiatives.

The session confirmed the trend of recent years: the conversion of utilization of aggregated data into individual data. The challenge has become to achieve a common link between separated databases. This process may have an impact on changing the structure of data collection and also increase the initial cost of providing IT solutions. On the other hand, it can reduce the reporting burden, improve data quality, provide for a more efficient compilation of statistics and, last but not least, render more timely data.

Summary of the panel discussion

Katherine Hennings

The breakout sessions covered presentations about an ample set of statistical databases developed to support policy analysis, supervision and research. The financial crises that emerged in 2007 illustrated that analysts, policy makers and regulators missed something. As a consequence, efforts were intensified to fill the

data gaps and to reinterpret the linkages between financial markets participants and the linkages between different components of financial markets.

One aspect that emerged was the relevance of granular data. The ease to collect basic granular statistics and the flexibility it allows to respond to different questions reduce the cost burden and permit the assistance of multiple users. In this context we could see the revival of central credit registers and the huge effort to build security-by-security databases. Also, other initiatives were enhanced, such as the central balance sheet databases, business registers, cross-border capital flows registers and other data repositories for statistical purposes.

The compilation of granular data allows information to be locally tailored in order to attend different users. Databases of single basic data serve to support micro-supervision analysis and enable bank supervisors to assess credit risk in supervised institutions. In addition, they also support economic analysis from central banks and financial markets analysts, as well as the research community. Therefore, micro-databases serve to a diversity of clients: central bank Board members that use economic analysis to take policy decisions, central bank analysts, supervisors, financial markets analysts, and academia members.

The experiences presented showed not only different areas of central banks working together, but frequently the joint work of a diversity of authorities. These efforts made clear the upmost importance of sharing data. It is recognized that the appropriate legal steps for a regular exchange of data serving to analytical and policy purposes is needed and that, sometimes, changes in the legal framework must be provided. This issue appeared as one of the challenges in the process of building a common database.

Considering that the data collected should attend to different purposes and users, the *praxis* raises other challenges, like the definition of a core set of information to meet main users and the definition of the attributes of statistics asked. This implies the harmonization of concepts and definitions, as well as the convergence of data coverage and data contents.

Apart from being robust and reliable, statistics compiled should be available within appropriate timeliness for timely decision-making processes. The security aspect and the feedback to respondents should also be taken into account. A clear specification of data to be collected is important for improving the quality of information received. Transparency and communication, both inside central banks and outside them, are crucial. It is important not only to make it clear the type of information required, but also explain the reason and present the results. In this context, tools and the appropriate language should be found.

Another aspect that emerged from the presentations and discussions was the increased cost for central banks to build the technological tools required. Since databases needed to store micro-data tend to be bigger than those required to store aggregate statistics, and given the fact that the use of micro-data is expected to serve different purposes, which are likely to change over time, the staff dealing with them must have special skills and training. This means a non-negligible increase in costs associated with IT and human resources for central banks. It has been proved that such costs are worth it but they ought to be well understood by the Board members. The buy-in by senior management is important, and, as suggested by a speaker, it should be granted and maintained.

Experiences also indicated that there is room for improvement. The credit registers could further explore other borrowing sectors and business deals –

household are not adequately covered, as well as inter-financial corporation lending, loans to the government sector, and the activity of "private credit bureaus". It was also discussed the importance of extending the coverage of information sources to comply with non-bank financial institutions (hire-purchase, leasing and credit card issuing companies; government agencies (which provide credit for specific purposes), issuance companies and payment system operators.

Security-by-security (s-b-s) databases are being developed in a number of economies and their coverage is still being discussed. In the building of an s-b-s database, it is crucial to define the attributes of the debt securities to be collected. In the European Central Bank (ECB) model, each security is identified by its International Securities Identification Number (ISIN) code and defined by more than 50 individual attributes. Since geographical coverage could differ, it should be also defined. Efforts and benefits are being evaluated in face of the costs – the ampler and deeper the database coverage, the higher the costs.

Notwithstanding all these challenges, the flexibility of granular data provided analysts at central banks and supervisory institutions with a previously inaccessible picture of the state of financial markets and of the funding of the economies, and helped to develop models to estimate the developments. Therefore, granular data is a source of information that is being increasingly used in combination with the available aggregate data.

Summary and conclusions of session on "Micro-databases Integration and Business Intelligence"

Chair: Miguel de Castro Neto[1]

Session D.2 "Micro-databases Integration and Business Intelligence" covered the entire lifecycle of business intelligence platforms in the context of central banks' micro-databases integration, including issues from collecting and processing data to the delivery of information to support decision-making to internal and external stake-holders.

The session program included five presentations detailed below that presented the different realities and challenges faced in distinct countries and environments (Germany, South Africa, Romania, Portugal and Mexico) in what concerns the evolution that is being made in adopting a business intelligence approach in the management of micro-databases and the creation of added value to the conversion of data into information in an increasingly analytical environment.

In the first presentation entitled "House of Micro-data and Research Data and Service Centre at the *Deutsche Bundesbank* – A draft concept", Ulf von Kalckreuth (Statistics Department, *Deutsch Bundesbank*) summarized the state of discussion in the Statistics Department of the *Deutsch Bundesbank* on setting up micro data based information systems. According to von Kalckreuth, currently the available production systems are not geared to micro level data analysis and changing them will be a time and resource consuming endeavour. The presented suggested approach features a short term enhancement of using existing micro-level data, jointly with a longer term reform of the infra-structure. Finally it was referred that the short term enhancement will be greatly helped by the fact that the Bundesbank has a long-standing tradition of providing micro-data to internal and external researchers on-site.

"Business intelligence tools and micro-data related to the South African equities market: Application and experience" was the title of the second presentation, where Zeph Nhleko (Research Department, South African Reserve Bank) addressed the specific thematic of equities market statistics referring that fully understanding the meaning of equities market statistics has become an increasingly difficult task in recent years. According to him, following the financial crisis many well-established relationships and causalities in the market have been weakened somewhat and this created the necessity for monetary policy-makers to analyse equities statistics in a much more pronounced granular manner. In this context the use of business intelligence tools and micro-data sets has never been more relevant and during the presentation it was possible to understand how such tools have been applied to extensive micro-data sets in the South African equities market environment.

The third presentation, made by Andreea Varjoghe (Statistics Department, National Bank of Romania) on "Data management – Running systems and ongoing projects in the National Bank of Romania", consisted in an overview of the main projects underway at the National Bank of Romania in this context. As the author

[1] Deputy Director, ISEGI, Universidade Nova de Lisboa.

referred, National Bank of Romania's present information systems consist in Intranet/Internet sites, communication/hardware/software infrastructure, dedicated systems (for accounting, for payment incidents, for balance of payments, etc.) and reporting systems: SIRBNR – Reporting Information System to the National Bank of Romani, which is the main reporting system to the National Bank of Romania and is in use from January 2007; and RAPDIR – Direct Reporting System, which was designed especially for reporters other than credit institutions.

This information systems' environment led to the development and implementation of a Data Warehouse – Business Intelligence (DW-BI) system in the National Bank of Romania, integrating data at micro and aggregate level from SIRBNR and RAPDIR databases, Excel/Word/PDF files and paper sources as well as from other systems and databases. Simultaneously, mechanisms and software tools were designed and implemented for loading and updating data in the DW-BI database on regular basis or on supervisor decision. According to the presenter, this approach and the business intelligence tools will allow users and key users to have *e.g.* better insight knowledge of the business areas, faster data access, *ad hoc* reports, dashboards, predictive analysis, drill-down analysis and what-if scenarios.

Maria do Carmo Aguiar (Statistics Department, *Banco de Portugal*) made the fourth presentation, on "Business intelligence in securities statistics: An evolutionary approach". According to Maria do Carmo Aguiar, securities statistics at *Banco de Portugal* have been around for a while. In fact, they were first released in the *Banco de Portugal*'s Annual Reports of the 1940s and their regular publication, with information on the issuance of short-term securities (*e.g.* Treasury bills and commercial paper) and long-term securities (shares and bonds), dates back to the 1980s. Monthly data on net issues and end-of-period positions, including totals and values broken down by type of security, maturity, currency, institutional sector of the issuer and respective classification of economic activity, were released for the first time in January 2005. As referred, a milestone in the history of *Banco de Portugal* securities statistics occurred in the late 1990s with the creation of the Securities Statistics Integrated System (SSIS). This system was developed with the purpose of gathering, in a single repository, all the information deemed necessary to comply with reporting requirements on securities.

It was referred that the SSIS stores securities issues and portfolios (holdings) data on "security-by-security" and "investor-by-investor" basis (comprising detailed data on stocks and transactions of debt securities, shares and other equity – financial derivatives are not included). The classification of securities relies on the existence of a reference database that, in the case of domestic securities, is maintained by *Banco de Portugal*; foreign securities are classified according to the Centralised Securities Database (CSDB) managed by the European System of Central Banks (ESCB).

With regard to securities holdings, Maria do Carmo Aguiar mentioned that detailed information is collected on the investments by residents in domestic and foreign securities, as well as on the holdings of non-resident investors in domestic securities. This information is reported by financial institutions, both as investors and custodians, and other resident entities.

An important aspect referred was that SSIS promotes consistency across statistics produced by *Banco de Portugal* and makes it possible to meet user needs, both at national and international level. In fact, SSIS information is used as an input for the compilation of a wide set of statistics produced at *Banco de Portugal*, from monetary and financial statistics to balance of payments and international

investment position and financial accounts. On the other hand, information from other statistical sources is used as an input for estimation of missing data and for data quality management purposes.

The initial SSIS architecture relies on two relational databases: the collected data is stored and validated in the "transactional database" (which is also used to estimate missing information); validated and "enriched" data are copied to the "exploration database" on a daily basis. In 2006, the analytical capabilities of the system were enhanced through the development of an analytical database, which enables user-friendly multidimensional analysis of the information.

Maria do Carmo Aguiar stressed out the fact that while multidimensional analysis of securities data has proved to be a powerful tool for rapid analysis and reporting, they foresee advantages in adding dimensions of analysis from other statistical systems, whose integration is not straightforward. Thus, more than ten years after its inception, it was launched in 2012, an Information Technology (IT) project aiming at developing a new SSIS within the Business Intelligence (BI) framework envisaged for the *Banco de Portugal* statistical systems.

Finally, Alejandro Gaytán González (Directorate of Financial System Information, *Banco de México*) made the fifth presentation, on "The use of micro-data in the financial system information model of *Banco de México*". According to the author, the current financial system information model of *Banco de México* was a strategic response to the challenges the 1994–1995 "Tequila Crisis" brought about. Although there was evidence on the building up of some unbalances (*e.g.*, large public debt denominated in foreign currency, increasing current account deficits and private credit booms), it was different to the traditional balance of payment crisis. Nonetheless, mismatches in bank balances and financial contagion mechanisms remained hidden.

In addition, in the mid-1990s the information model of the financial authorities was cumbersome and had several inefficiencies: (i) disperse, non-structured and duplicated information requirements from financial authorities; (ii) heavy information burden, *i.e.*, too many data requirements, some of them obsolete or inaccurate; and (iii) inadequate frequency, granularity, and opportunity of the information.

Thus, as Alejandro Gaytán González referred, information was insufficient to identify measure and monitor risks. Also strengthening the supervision and compliance of new regulation as well as the measurement and monitoring of risks posed heavy information challenges. In this context, it was said that some important strategic decisions were undertaken: (i) coordination and information sharing among financial authorities; (ii) redesign of data requests to reduce duplicates and to generate economies of scale; and, (iii) *Banco de México* develop a highly granular information model to ease future analysis and reduce information scarcity in turmoil times.

A Research Data and Service Centre (RDSC) at the Deutsche Bundesbank – a draft concept

Ulf von Kalckreuth[1]

Abstract

This paper summarises the state of discussion in the Statistics Department of the Deutsche Bundesbank on setting up micro data based information systems. Currently, the production systems are not geared to micro level data analysis. Changing them, will be a time and resource consuming endeavour. The suggested approach features a short term enhancement on the basis of existing infrastructure. A "House of Microdata" will be created, essentially a Statistical Data Warehouse, together with a Research Data and Service Centre. The pre-existing statistical processes will be left intact, minimising operational risk. This is at once a huge leap forward towards availability of micro data for analysis and research, and a step on the way to a longer term reform of the informational infra-structure and data-management. The enhancement will be greatly helped by the fact that the Deutsche Bundesbank has a long-standing tradition of providing micro-data to internal and external researchers on-site.

[1] Deutsche Bundesbank, Statistics Dept., Central Office, ulf.von-kalckreuth@bundesbank.de.

A Research Data and Service Centre (RDSC) at the Deutsche Bundesbank – a draft concept

This paper summarises the state of discussion in the Statistics Department of the Deutsche Bundesbank on setting up micro data based information systems. Currently, the production systems are not geared to micro level data analysis. Changing them, will be a time and resource consuming endeavour. The suggested approach features a short term enhancement on the basis of existing infrastructure. A "House of Microdata" will be created, essentially a Statistical Data Warehouse, together with a Research Data and Service Centre. The pre-existing statistical processes will be left intact, minimising operational risk. This is at once a huge leap forward towards availability of micro data for analysis and research, and a step on the way to a longer term reform of the informational infra-structure and data-management. The enhancement will be greatly helped by the fact that the Deutsche Bundesbank has a long-standing tradition of providing micro-data to internal and external researchers on-site.

1. Mandate

In line with the objectives set out in the Bundesbank's Strategy 2016, micro data are to be made available, in a new setting and "philosophy", for both analytical tasks and scientific research. This is partly in order to meet the increased data needs of the Bundesbank's departments and the ESCB resulting from the expansion in their range of duties and the new issues they face. The following problems need to be solved:

- Data management in the Statistics Department (and elsewhere in the Bundesbank) is not primarily designed to support micro-level analysis; it is usually based on process data used to create statistical aggregates or data for prudential or administrative purposes. Micro data are generally not available as closed data sets;
- Documentation for use by externals is often absent or incomplete;
- The use of micro data is restricted by data protection rules; different legislation applies in different contexts. The level of confidentiality has to be systematically assigned to user groups;
- The data are created through various processes and are not integrated. For any given financial institution, the Bundesbank has data from various sources, eg the bank balance sheet statistics, the borrowers statistics, the external position and prudential reports; it is very difficult to analyse these data simultaneously;
- For both internal and external researchers and analysts, working with unknown and/or complex micro data requires investing a lot of time and effort. It takes long to understand the content of a data set – including restrictions and sources of error – well enough to draw reliable conclusions regarding distributions.

The proposed solution is an interdepartmental project coordinated by the Statistics Department to set up a ***"House of Microdata" (HoM)*** and a ***Research Data and Service Centre (RDSC)***.

1. The "House of Microdata" will be the infrastructure component, receiving "clean copies" of process data from data production entities, formatted in such a way as to facilitate their joint use. The data clusters will be selective and geared to the use for analytic purposes. In the House of Microdata, the data can be retrieved by anybody within the Bank who holds the access rights. An existing, powerful platform for storage and retrieval of time series can be extended, and the well-established SDMX-standard will be used to create data models.
2. The Research Data and Service Centre will ***facilitate data access*** for researchers and analytics, as well as ***provide important services***:
   - compiling, documenting and archiving informative micro data sets,
   - enhancing existing data through record linkage,
   - supplying internal and external researchers and analysts with micro data sets in line with the relevant legislation,
   - methodological and content-related support and services for users,
   - analyses and customised processing of requests from internal analysts in cases where legal constraints prevent them from accessing the data or where it would be too time-consuming for them to process the data themselves,
   - supplementary methodological and descriptive research based on the data sets created by the unit.

   Looking ahead, potential is to be created for

   - receiving and supplying micro data from other departments and external sources,
   - exchanging research data within the ESCB and with other research data centres.

Strategically speaking, the production and provision of micro data sets fit for micro level analysis is a new and unique type of statistical product; the outcome of the statistical process is not the indicators themselves, but a ***new possibility for users to produce their own statistical results following their own criteria***. To support this process, the unit will provide documentation and advisory services, clarify organisational and legal issues, and create and operate a technical infrastructure which allows micro data to be processed within a governance structure that fulfils all the relevant legal requirements and the principle of "participate and contribute".

By contrast, in conducting its own analyses on behalf of others, the RDCS will essentially be providing one of the Statistics Department's classic products: it will aggregate the available individual data to derive moments and indicators for distributions. Users generally want information on distributions – such as risk measures or correlations – rather than individual data. However, the RDCS would provide information flexibly according to varying user requirements. This service is particularly suitable for meeting the growing data needs of other Bundesbank's departments and the ESCB. Where certain analyses come up time and again and can be standardised, the unit will also assume development functions for the entire department.

Given the RDCS's role in development and dissemination, it will need to carry out a certain amount of own research. The unit needs to compile research data sets

using process data. This is a constructive service which will have to be geared to the users' needs. Developing a research data set is, in itself, a form of research. It requires explorative work and examination of current issues and standpoints in the world of research. A minimum level of active research is one of the prerequisites set by the German Council for Social and Economic Data (RatSWD) for recognising an institution as a research data centre.

2. Developing the RDSC

Initially, the RDSC's work would focus mainly on data structures for which the Bank already has experience in micro-level scientific analysis. These data structures would take the form of scientific use files (SUFs), would be saved periodically (data freeze), archived in line with DOI standards and documented. User rights and restrictions would be clarified and an application procedure set up. Record linkage would be carried out on the basis of available characteristics. The statistical process routines used at present would remain unchanged; individual data would be selected periodically according to certain rules and – where necessary – manual record linkage would be carried out using available common identifiers. The trade-off for this would be a fairly long time lag between the reporting or recording of the original data sets and collective availability. This would be less of a problem for academic users than for analytical work. Users will also generally need to be provided with sample routines for generating aggregates both at the level of observational units and for sub-populations.

Setting up this unit will be greatly helped by the fact that the Research Centre of the Deutsche Bundesbank has a long-standing tradition providing micro-data for researchers on-site. In the past, micro data sets have been provided on special requests, concerning

- Corporate Balance Sheet Statistics,
- MFI Monthly Balance Sheet Statistics (BISTA) and profit and loss accounts,
- External position of banks,
- Statistics on Securities Investments (formerly Securities Deposit Statistics),
- German Panel on Household Finances (PHF),
- Micro-Database Direct Investment (MIDI),
- Several merged data sets.

It could be a suitable point of departure to fully develop these micro data for analysis. While supplying users with an initial menu of data sets, the unit would simultaneously start to develop more and better integrated micro data sets and well-organised statistical support for use by internal and external analysis.

3. The structure of an RDSC

Apart from a manager, an RDSC would require

a) ***experts for the existing data sets***, who would be responsible for the creation and development of microdata sets, documentation, data archiving and user support. They would act as a liaison point between users and the departments responsible for the data in question. Additionally, they would carry out special analyses on request and conduct a limited amount of research.

b) ***qualified staff for the technical provision of data*** by means of

- teleprocessing
- a secure on-site computer
- de facto or fully anonymised scientific use files
- monitoring data output
- (possibly) an exchange with other research data centres

c) ***qualified staff for merging and validating data,*** as well as supplementing data with additional information, such as sector identifiers.

4. Integrated micro data management

With an integrated approach to processing micro data, each characteristic should in principle only be recorded once – this characteristic is then added to the existing data on the respective observational unit and made available with all others. Such a state of integration largely ***eliminates statistical redundancy***, reduces the workload of the data providers and increases the volume of information that can be analysed simultaneously. Record linkage would be a by-product of data processing.

From an internal perspective, it is of particular importance that different units in the department, working with the same observational units (banks, enterprises, securities), are able to ***exchange data more easily***. This will facilitate more efficient, rapid and flexible processing of statistical data as well as enhancing quality control in the production of the data. Ideally, the goal is to make all statistical information ***available as both aggregated and granular data***, allowing them to be combined in various ways depending on the requirements of the task at hand. This would e.g. be very useful for examining financial risk issues from a new angle.

It will be easier to apply this to banking statistics than to firm-level data, for which there are currently no common indicators in the Bundesbank. An integrated approach to firm level data is severely affected by the fact that currently, for legal reasons, the Bundesbank has no access to the Federal Statistical Office's business register. Much progress in the integration of data has already been made in the area of securities. If need be, the processes used in the production of statistical data will have to be adapted or replaced. Ultimately, the reporting system will have to be overhauled.

The existing production systems answer diverse analytical and administrative questions by equally diverse and independent statistical processes, very often operating under different legal and institutional frameworks. Any step of further integration must be ***realised under the strict condition of continuous functioning*** within the pre-existing organisational and legal context.

A very important step towards integrated data management will be achieved by ***setting up the House of Microdata***, which is essentially a Statistical Data Warehouse for micro data. The SDMX based data models will be created in such a way as to facilitate the joint use of the data, and, where possible, there will be joint reference data for the micro units. This will enable the statistical data production units to use the finalised data of other units for consistency checking and outlier analysis. However, this is not tantamount to integrated data management on the process level. A full integrated data approach can be approached only gradually, and as a ***result of an evolutionary process***.

The associated work will involve many units and can by no means be carried out entirely (or predominantly) by a separate unit. However, the data repositories that will be created as a result of a reorganisation of data could be operated centrally for all other units. The RCDS's potential would be amplified by employing an integrated approach to data processing – particularly because the data which have already been integrated would no longer have to be merged.

Business intelligence tools and micro-data related to the South African equities market: application and experience

Zeph Nhleko[1]

Abstract

Fully understanding the meaning of equities market statistics has become an increasingly difficult task in recent years. Following the financial crisis many well-established relationships and causalities in the market have been weakened somewhat. This makes the necessity for monetary policy-makers to analyse equities statistics in a granular manner much more pronounced. The use of business intelligence tools and micro-data sets for this purpose has never been more relevant. This note seeks to outline how such tools have been applied to extensive micro-data sets in the South African equities market environment.

Keywords: micro-data, all-share price index, business intelligence

[1] Head, Capital Market and Flow of Funds, Research Department, South African Reserve Bank, zeph.nhleko@resbank.co.za. The views expressed are those of the author and do not necessarily reflect the views of the South African Reserve Bank.

1. Introduction

The purpose of this note is to show how business intelligence tools and micro-data may be combined to provide a view on developments in equity prices. Business intelligence tools in this context refer to software or applications that can be used in a central banking environment to analyse, query and report on capital market data. In particular, *I-Net Station* by I-Net Bridge (I-Net) and *Reuters Eikon* by Thompson Reuters (Reuters) are considered here. I-Net is a South African financial services company that specialises in providing economic data, financial market data and corporate market intelligence in South Africa. Reuters provides a synthesis of human intelligence, industry expertise and innovative technology to deliver insight to, among others, financial, risk, legal, tax and accounting markets. Micro-data on the other hand is used here to refer to individual listed company data on South Africa's stock exchange, the JSE Limited (JSE). Such data include volumes, values and prices at company level in the market.

2. Rationale for using equities market micro-data to assess equity price developments in South Africa

The focus of this note is on equities and, in particular, trying to understand the role of dual-listed shares on the overall all-share price index (Alsi) of the JSE. The Alsi is a free-float, market capitalization weighted index that tracks the performance of all companies listed on the JSE. It is the broadest and representative measure of the exchange's performance. There is a ***view***[2] that dual listed shares may reflect offshore sentiment or developments and the rand exchange rate movements – and may as such influence the overall share market in South Africa. This view finds further support from the fact that the weight of dual-listed shares on the Alsi is over 65 per cent. However the concept of dual listing can be confusing at times[3] and must be carefully interpreted – a couple of points are important in this regard.

Firstly, the extent of dual listing (in percentages) varies across companies and this may affect the level of impact such shares may have on the domestic market. *Secondly*, the standard dual listing status of a company does not take into account depository receipts, *i.e.*, negotiable financial instruments issued by banks to represent foreign companies' publicly listed shares. Although some companies listed on the JSE do not have secondary listings, they have depository receipt programs with banks offshore. Typically these banks list the shares on offshore exchanges to be traded by non-residents. *Thirdly*, some companies, even though they are only listed domestically, have foreign ownership in excess of 50 per cent, which should make them susceptible to offshore sentiment and developments. *Fourthly*, the objective for buying certain shares may actually have nothing to do with investor sentiment or economic factors for that matter, as such shares may be

[2] For example, (1) Internal discussions, (2) Uranium One prejudiced by JSE rules, http://www.miningmx.com/..._site_id=83 and (3) Arbitrage opportunities on dual-listed stocks http://www.moneyweb.co.za/...-duallisted-stocks.

[3] E-mail response to questions posed, Mark Randall, Manager – Indices: Post Trade Services, JSE Limited.

bought to be defensive or for growth purposes. *Lastly*, the proportion of revenue earned in South Africa will make shares sensitive to a certain degree to domestic factors.

However one assessment that can be done to check the validity of this ***view*** is to analyse the drivers of each share on the JSE. Another assessment is to compute separate indices for dual-listed shares (DLSi) and non-dual-listed shares (NDLSi) for South Africa. Similar to the Alsi, these indices should be weighted indices that include/exclude the dual-listed companies in order to capture the supposedly domestic economic and other influences on share prices. Companies whose dual listing is in Africa need not be excluded from the index as the impact of their share price movements would be minimal on the JSE. Based on current listing and for purposes of calculating these indices, "dual-listed" means listed at the JSE and outside Africa.

There are several ways of compiling indices, but the two most basic formulae for computing indices are the Paasche and Laspeyres[4] methodologies. The only real difference between them is that the Paasche index uses the number (n) of quantities and the Laspeyres index uses base period (0) quantities. The indices take the following forms respectively;

$$I_p = \frac{\Sigma(\Delta p_{t_n} \times \Delta q_{t_n})}{\Sigma(\Delta p_{t_0} \times \Delta q_{t_n})} \qquad \text{and} \qquad I_l = \frac{\Sigma(\Delta p_{t_n} \times \Delta q_{t_0})}{\Sigma(\Delta p_{t_0} \times \Delta q_{t_0})}$$

where p is the price level, t_0 is the base period and t_n is the number of periods for which the index is computed. Adapting the Laspeyres methodology and in line with JSE's index calculation technique, the proposed DLSi and NDLSi can take the form;

$$I = \sum_{i=1}^{N} \frac{(p_i \times s_i \times f_i)}{d}$$

where N is the number of securities in the index, p_i is the latest share price, s_i is the number of shares issued or traded, f_i is the investability factor (free float weighting) and d is the divisor representing all issued shares at the base period. The divisor changes with the issued shares in order to avoid distorting the index.

3. Business intelligence tools and the compilation of an alternative all share price index

The process of calculating the separate indices requires micro-data at individual company level, such as share prices and individual company volumes of shares traded on a daily basis. I-Net and Reuters allow for collecting these types of data over time. These tools allow company code names to be identified separately and

[4] Etienne Laspeyres, 1871 and Herman Paasche, 1874.

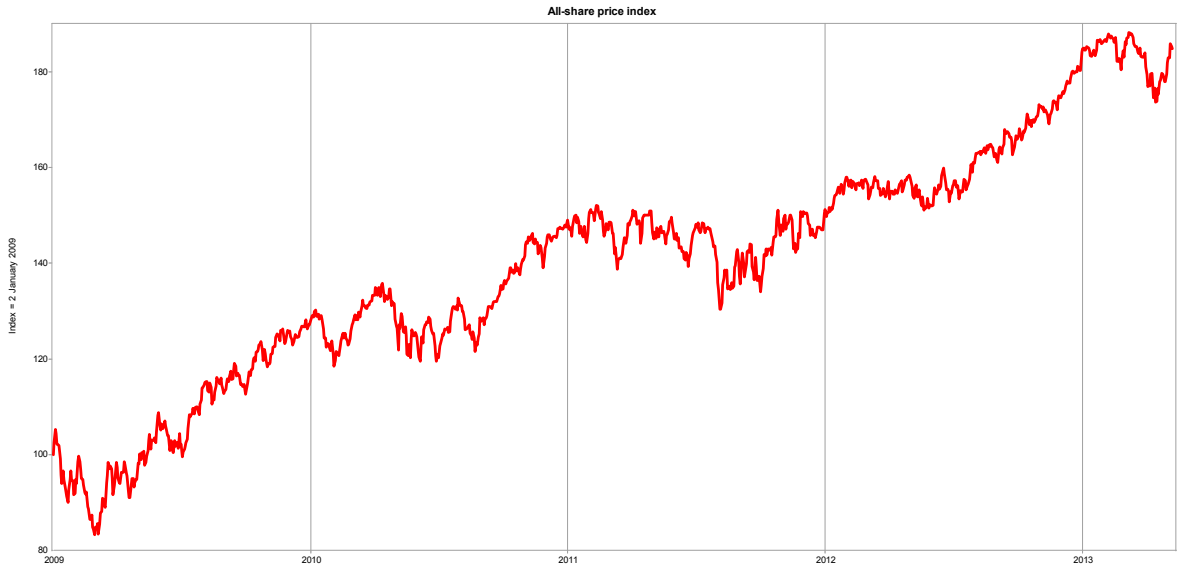
used to compile a matrix with share prices, volumes traded, market capitalisation and similar variables. Individual company code names can also be used to check and verify company registration/deregistration dates with the JSE – this is important information for maintaining the index, i.e., the process of adjusting the index divisor.

4. Conclusion

Figure 1 shows that, although fluctuations are more pronounced, the average prices of non-dual-listed shares increased by 96 per cent from February 2009 to April 2013, much faster than dual-listed share prices at 67 per cent. The overall Alsi increased by 87 per cent over this period. The differential between the Alsi, DLSi and NDLSi represent a confluence of domestic and offshore factors that affect domestic share prices. Therefore the view that dual-listed shares drive the Alsi more than other factors has no solid basis. Further analysis is required.

Share prices

Figure 1



References

Capital Market and Flow of Funds, 2012 (March): Cheapness of the South African equity market – a perspective on PE ratios, Unpublished.

Capital Market and Flow of Funds, 2012 (July): Performance of the All-share Price Index (Alsi) and some of its drivers, Unpublished.

Capital Market and Flow of Funds, 2012 (August): Is the compensation for holding shares vs. bonds negative?, Unpublished.

Chance, W.A. 1966. A note on the origins of index numbers, *The Review of Economics and Statistics*, 48 (1). February, pp. 108 – 110.

FTSE/JSE. 2004. Guide to calculation methods for the FTSE/JSE Africa index series.

FTSE/JSE. 2004. Guide to calculation methods for the UK series of the FTSE Actuaries share indices.

FTSE/JSE. 2012. Ground rules for the management of the FTSE/JSE Africa index series.

I-Net Bridge, 2013. http://www.inet.co.za/#about_us/company_prof.php. Accessed 10 April 2013.

S&P Indices. 2012. Index mathematics methodology

Thompson Reuters, 2013. http://thomsonreuters.com/about/. Accessed 10 April 2011.

Data management – running systems and ongoing projects in the National Bank of Romania

Andreea Varjoghe[1]

The National Bank of Romania (NBR) information system consists in Intranet/Internet sites, communication/Hardware/Software infrastructure, dedicated systems (for accounting, for payment incidents, for balance of payments, etc.) and reporting systems (SIRBNR, RAPDIR).

SIRBNR (Reporting Information System to the National Bank of Romania) is the main reporting system to the National Bank of Romania (NBR) in use from January 2007.

Credit institutions report data regarding their own activities or in the name of their clients to the NBR using this system that integrates all core business areas:

- Balance of payments
- Financial markets
- Monetary policy
- Monetary statistics
- Interest rates statistics
- Financial stability
- Prudential supervision, etc.

There are several important features that make this system a very efficient tool:

- Unique secure channel: SIRBNR portal (web page accessed with user/password authentication)
- Unique format: XML message
- Minimized number of connections
- Standard flow of messages (regarding preliminary, business and cross-check validation)
- No redundancy
- Centralized administration
- Nomenclature harmonization (standards enforcement)

Reported data are stored in a large database that integrates several micro-databases and aggregated indicators.

NBR users access SIRBNR applications for data processing (such as cross validation, aggregation, other calculations, data extraction to Excel/CSV files), data monitoring and final reports (based on business requirements, both for internal analyses and for reports to ECB and other international institutions).

[1] Statistics Department, National Bank of Romania, andreea.varjoghe@bnro.ro.

On the other hand Credit institutions have access to their own reported data and to aggregated indicators.

RAPDIR (Direct Reporting System) was designed especially for reporters other than Credit Institutions. This system is currently on *parallel-run* test session and will start running this year. It will be the main channel for collecting statistical data directly from reporters such as:

- Non-bank financial institutions
- Insurance companies
- Investment funds
- Natural persons
- Non-financial corporations etc.

The estimated number of reporters is 40.000. Due to this large number of reporting entities, the NBR's branches will be involved in the reporting process administration under the Statistics Department supervision.

Based on the SIRBNR experience, RAPDIR has many similar features and also a series of improvements (like more flexibility in the uploading process and in system administration):

- Secure channel (web page accessed with user/password authentication)
- Reporting format: online/offline XML files or online forms (users can also make on-line corrections if necessary)
- Automated flow of messages (regarding preliminary, business and cross-check validation)
- No redundancy
- Nomenclature harmonization (standards enforcement)

All statistical data collected through RAPDIR will be stored in several integrated micro-databases covering areas like international trade in services, foreign direct investment, foreign accounts transactions, non-monetary transactions, assets and liabilities of investment funds and non-bank financial institutions, foreign exchange transactions, portfolio investment etc.

There will be software tools in RAPDIR similar to those in SIRBNR that will allow NBR users to:

- Process data (calculations, extraction to Excel/CSV files etc.)
- Monitor transmissions
- Generate final reports for analysis and decision
- Prepare data sets to be reported to ECB and other international institutions

There are all premises for developing and implementing a Data Warehouse – Business Intelligence (DW-BI) system in the NBR:

- Diversity, high quality and huge amount of collected data
- Good performance of the SIRBNR reporting system
- Imminent start of the new reporting system (RAPDIR)
- Large business area (and number of reporters) covered by reporting systems

The new DW-BI system will integrate data at micro and aggregate level from SIRBNR and RAPDIR databases, Excel/Word/PDF files and paper sources as well as from other systems and databases. There will be designed and implemented mechanisms and software tools for loading and updating data in the DW-BI database on regular basis or on supervisor decision.

Business intelligence tools will allow users and key users to have better insight knowledge of the business areas, faster data access, ad-hoc reports, dashboards, predictive analysis, *drill down* analysis, *what-if* scenarios.

Business intelligence in securities statistics: an evolutionary approach

Maria do Carmo Aguiar[1]

"It is not the strongest or the most intelligent who will survive but those who can best manage change."

Charles Darwin

Storyboard

Securities statistics at Banco de Portugal (BdP) have been around for a while. In fact, they were first released in the BdP's Annual Reports of the 1940s and their regular publication, with information on the issuance of short-term securities (e.g. Treasury bills and commercial paper) and long-term securities (shares and bonds), dates back to the 1980s. Monthly data on net issues and end-of-period positions, including totals and values broken down by type of security, maturity, currency, institutional sector of the issuer and respective classification of economic activity, were released for the first time in January 2005.

A milestone in the history of BdP securities statistics occurred in the late 1990s with the creation of the Securities Statistics Integrated System (SSIS). This system was developed with the purpose of gathering, in a single repository, all the information deemed necessary to comply with reporting requirements on securities.

The SSIS stores securities issues and portfolios (holdings) data on "security-by-security" and "investor-by-investor" basis. It comprises detailed data on stocks and transactions of debt securities, shares and other equity (financial derivatives are not included). Both stocks and transactions are collected on a monthly basis. The classification of securities relies on the existence of a reference database which, in the case of domestic securities, is maintained by BdP; foreign securities are classified according to the Centralised Securities Database (CSDB) managed by the European System of Central Banks (ESCB).

In what respects to securities holdings, detailed information is collected on the investments by residents in domestic and foreign securities, as well as on the holdings of non-resident investors in domestic securities. This information is reported by financial institutions, both as investors and custodians, and other resident entities.

SSIS promotes consistency across statistics produced by BdP and makes it possible to meet user needs, both at national and international level. In fact, SSIS information is used as an input for the compilation of a wide set of statistics produced at BdP, from monetary and financial statistics to balance of payments and international investment position and financial accounts. On the other hand,

[1] Statistics Department, Banco de Portugal, maria.carmo.aguiar@bportugal.pt.

information from other statistical sources is used as an input for estimation of missing data and for data quality management purposes.

The initial SSIS architecture relies on two relational databases: the collected data is stored and validated in the "transactional database" (which is also used to estimate missing information); validated and "enriched" data are copied to the "exploration database" on a daily basis. In 2006, the analytical capabilities of the system were enhanced through the development of an analytical database, which enables user friendly multidimensional analysis of the information.

While multidimensional analysis of securities data has proved to be a powerful tool for rapid analysis and reporting, we foresee advantages in adding dimensions of analysis from other statistical systems, whose integration is not straightforward. Thus, more than ten years after its inception, it was time to reformulate SSIS. Launched in 2012, an Information Technology (IT) project aims at developing a new SSIS within the Business Intelligence (BI) framework envisaged for the BdP statistical systems. The following section describes the main features of this framework.

Behind the scenes

Evolution of methodologies, frameworks and statistical demands are the main reasons for the need to reformulate most of the information systems that are used in statistical production. Furthermore, the need to interconnect different statistical domains and cross check data has become a fundamental requisite.

Business intelligence (BI) is "a broad category of applications and technologies for gathering, storing, analyzing, and providing access to data to help enterprise users make better business decisions". Simply put, BI is about getting the right information to the right people at the right time, so that they can make good decisions that improve organisational performance.

In the beginning of 2008, the IT Department and the Statistics Department of BdP initiated a study with the purpose of defining a BI framework, to be used as a reference in all future IT developments in the statistical area. This framework is built upon three pillars: a data warehouse, centralised reference tables and a common IT platform. The data warehouse guarantees a central access point to every statistical data, independent of the input source or the production process; the centralised reference database provides common reference data and enables cross linking information from different sources and systems; the consistent usage of a common technological infrastructure across the multiple information systems makes it easier to integrate and reuse components and promotes data access efficiency and transparency to final users.

The stocktaking exercise on the relationships between existing statistical systems showed a high level of interdependency, with multiple sets of data being exchanged over the network, each of which poses a need for data extraction, cleaning, transformation and integration. A data warehouse was proposed to deal with this overhead of multiple data exchanges. Each new or renewed system would contribute to, and use data from, a single centrally managed database. The goal was to achieve a situation where there is a single version of trusted data, and little or no data is directly exchanged between statistical systems.

The second pillar of the BI framework is the Master Data Management (MDM), which collects, consolidates, stores and delivers reference data (e.g. countries, currencies, financial sectors,...) that are used across the systems. A register of financial institutions has been kept for long, and there was an effort, still ongoing, to streamline the process of gathering data from several sources and consolidating it into an historical register of all resident companies. The MDM assigns each register an internal surrogate key that never changes; the mapping between the internal key and the corresponding external business or natural keys (different data sources often report different keys for the same data entity) are kept in the MDM system. This is in fact an extremely important component of the framework, as it gives consistent meaning to the facts stored in the data warehouse, and makes much easier the crossing of fact data from different systems.

In what respects to the third pillar, the IT platform, in the last quarter of 2007, the IT Department and the Statistics Department jointly carried out a Proof of Concept (PoC) for the selection of the BI Platform to be used in the forthcoming IT projects of the statistics business area. After a preliminary market analysis, and given the history of BI tools usage at BdP in the recent years, the short list was reduced to two solution providers: Microsoft and SAS. The results of this PoC showed that none of the two platforms could deliver a high level of functionality across the whole spectrum of business needs, but they were in fact complementary. The final choice was then a combination of the two solutions.

Meanwhile, BI tools and concepts have evolved and the focus nowadays is much more on the so called data discovery tools, which enable more powerful analysis of data from multiple sources, which do not necessarily imply a centralised data warehouse. Against this background, a new PoC is currently taking place at BdP in order to re-evaluate the best BI platform solution to accomplish our targets.

To be continued...

In a nutshell, statistical systems comprise four phases: Acquisition, Processing, Exploration and Dissemination. Each phase has a comprehensive set of core tasks and processes. Redesigning the SSIS according to the BI framework implies redefining these processes, along with adopting the common IT infrastructure. The data model that supports SSIS needs to be redesigned in order to separate data acquisition and data processing repositories, which currently are the same. Internal reference data will be replaced with MDM references, ensuring harmonisation with other statistical systems. When available, input from other systems will be provided via the data warehouse; this will be an incremental process. Multidimensional analysis, already being performed, will benefit from data of other systems, available via the data warehouse. Consequently, efficiency gains are expected, along with improved data analysis.

BI researchers state that "organisations which have the greatest success with BI travel an evolutionary path, starting with basic data and analytical tools and transitioning to increasingly more sophisticated capabilities". On the other hand, market leaders in BI are constantly providing new tools to address the main issues in this field: integration, information delivery and analysis.

In statistics, like in many other areas, there is the need for continuous improvement and innovation. A stepwise approach is not only wise but the most realistic to be followed.

References

Aguiar, M (2007). "Integrated Statistical Systems: evolution or revolution?", 56th ISI Session (Irving Fisher Committee, IFC Bulletin N. 28), Lisbon, August 2007.

Aguiar, M & Martins, C (2011). "Adding Business Intelligence to statistical systems – the experience of *Banco de Portugal*", Eurostat Conference on "New Techniques and Technologies for Statistics". Brussels, February 2011.

IBM & Gartner (2011). "A step-by-step approach to successful Business Intelligence".

Nelson, Greg (2007). "Introduction to the SAS® 9 Business Intelligence Platform: A Tutorial", SAS Global Forum 2007.

The use of micro-data in the financial system information model of *Banco de México*

Alejandro Gaytán González[1]

Introduction

The current financial system information model of *Banco de México* was a strategic response to the challenges the 1994–1995 "Tequila Crisis" brought about. Although there was evidence on the building up of some unbalances (*e.g.*, large public debt denominated in foreign currency, increasing current account deficits, private credit booms), it was different to the traditional balance of payment crisis. Nonetheless mismatches in bank balances and financial contagion mechanisms remained hidden.

In addition, in the mid-1990s the information model of the financial authorities was cumbersome and had several inefficiencies: i) disperse, non-structured and duplicated information requirements from financial authorities; ii) heavy information burden, *i.e.*, too many data requirements, some of them obsolete or inaccurate; and iii) inadequate frequency, granularity, and opportunity of the information.

Thus, information was insufficient to identify measure and monitor risks. Strengthening the supervision and compliance of new regulation as well as the measurement and monitoring of risks posed heavy information challenges. In this context, some important strategic decisions were undertaken: i) coordination and information sharing among financial authorities; ii) redesign of data requests to reduce duplicates and to generate economies of scale; and, iii) *Banco de México* develop a highly granular information model to ease future analysis and reduce information scarcity in turmoil times.

The Information Model at *Banco de México*

The information model in its collection of data integrates information collected by *Banco de México* and other authorities and sources. Currently, two types of data models converge into a single model: aggregated information and micro-data (transactional as well as other highly detailed information).

In general, the model has transactional information on all transactions of financial intermediaries in the different markets (see Figure 1):

- Daily: FX transactions, detailed transactions in Debt Securities (Spot sales/purchases, Repos, and lending) and Derivatives (OTC and exchange traded), Interbank Loans and Time Deposits, all transactions in Credit and Debit Card Switches.
- Monthly: Equity holdings and Commercial Loans and Mortgages.

[1] Directorate of Financial System Information, *Banco de México*, agaytan@banxico.org.mx.

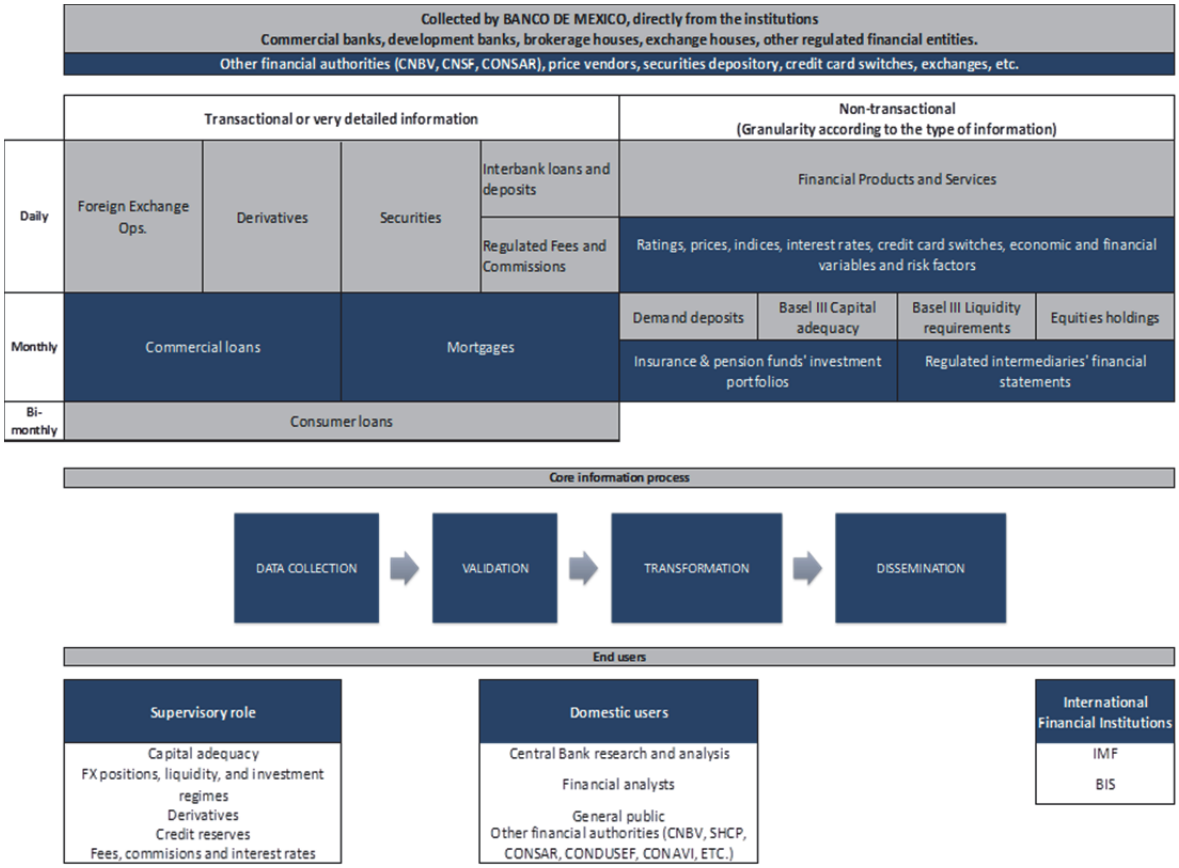
– Bimonthly: Consumer Credit (information card by card of Credit Cards and loan by loan in other consumer loans).

– Continuous: Registry of Regulated Fees and Commissions.

Non transactional information on the other hand includes daily information in Foreign Claims and Liabilities, information provided by data vendors and other sources, Demand Deposits, Basel III Capital Requirements and Liquidity Coverage, Equity Holdings and other intermediaries' financial information.

The Information Model at *Banco de México* – The current Banco de Mexico's information model, core information process and main end users

Figure 1



Some of the main advantages of this information model are:

– Since exhaustive attributes of all transactions are warehoused, it is possible to replicate a large set of aggregates.

– It helps to ensure the quality and consistency of information through comprehensive cross validation of information within the institution and between institutions.

– It allows fulfilling different information requests by users without additional requirements, reducing the reporting burden to financial intermediaries, *e.g.*, sectorial decomposition of banks' assets and liabilities, and statistics on markets activity.

– Entities' aggregated information may mask the build-up of risk; however, a comprehensive model could be reached through the high granularity information collected that allows for a deeper analysis of the entities.

- Financial supervisors can identify relevant operations and measure their marginal contribution.
- Counterparty identification of all transactions allows building very detailed models of risk and network analysis.

The "core information process" is divided in four different processes:

1) Data collection
2) Validation
3) Transformation
4) Dissemination.

As mentioned before, the process of validation is exhaustive. The "transformation" process provides the value added and it is here where the data are contextualized, aggregated, consolidated or utilized as input in both regulatory and risk models. The "dissemination" activity is centralized using a Data Hub and there are different means of dissemination, including access to data bases, reports, cubes, etc.

The model has several uses and users, including inside users in *Banco de México*, users in other financial authorities and the general public. Figure 2 presents some of the main uses of the information model.

Main Uses of the Financial System Information Model – Frequency, domain, processes and uses of the Financial System Information Model

Figure 2

Primary Information Processes			Description	Risks	Analysis	Costs	Regulation
Markets: Microdata, Transactional or huighly detailed information	Daily	Derivatives	Operation by operation OTC and Exchange Traded: Life Cycle and Snapshot Approches (= Derivatives Repository)	✓	✓		✓
		Debt Securities	Operation by operation, securities lending, repos and spot sales/purchases	✓	✓		✓
		Interbank Funding and Time Deposits	Detail of interbank funding, funding concentration and time deposits	✓	✓		✓
	Monthly	Commercial Credit	Commercial Credit Registry Loan by loan	✓	✓	✓	✓
	Bimonthly	Credit Cards	Card by card balance, interest and payments & Transaction by transaction from switches (include debit cards)	✓	✓	✓	✓
		Other Consumer Credits	Loan by loan : 1) payroll, 2) personal, 3) automovile, 4) urable goods, 5) gruop and 6) otros	✓	✓	✓	✓
Financial S	Monthly	Bank Balances	Financial balances with sectorial breakdown	✓	✓		✓
		Other Regulatory Reports (other authorities)	ortgages, operative reports, securities portfolio of investmet funds, pension funds and other institutions	✓	✓		✓
		FX Cash operations	FX cash operations	✓	✓		✓
	Qrt	Payment Systems	Checks, transfers, cards, ATMs, costs of payment systems		✓	✓	
Regulatory Compliance	Daily	FX regulatory Regimes	Liquidity and Exchage Risk	✓	✓		✓
		Fees and Commissions Registry	Costs of Deposits and Credit Financial Products (SMEs and Households)		✓	✓	✓
	Monthly	Capital Requirements (Basel III)	Templates with high level of detail	✓	✓		✓
		Liquidity (Basilea III)	High level of detail	✓	✓		✓

Information collected by Banco de México

Información collected by other authority

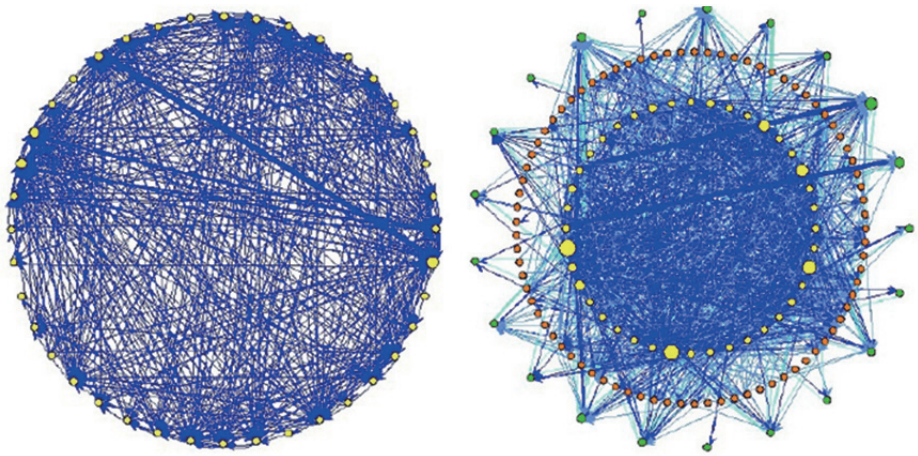
Some Uses of Information	
Risks Micro y Macroprudential	Contagion Models
	Market Risk Models
	Capital Requirements Model
	Liquidity Risk
	Financial Institutions
	FX positions
	Indebtness by Sector
Macrofinancial Analysis	International Banking Statistics (BIS)
	Monetary and Financial Aggregates
	Balance of Paymets
	Fincial Positions of Households and Firms
	Financial Programming
Costs of Fianancial Services	Interest Rates and Total Anual Costs
	Reports on Comparative Costs
Regulatory Compliaance	Fees and Commissions Registry
	Derivados, Valores, Tasas, Capitalización, Moneda Extranjera

One of the uses of the information model, the analysis of systemic risk (contagion) based on detailed network models is a good example of intensive use of micro-data and the identification of counterparties available.

The data model for counterparty risk analysis is fed with daily information reported by Banks, Brokerage Houses, Investment Funds and Pension Funds. Martinez *et al.* (2012 and 2013), for example, build-up network models to analyze stress scenarios and the contagion channels through the financial system. They conclude that previous research results showing that contagion mechanisms may not be relevant are probably biased by the lack of a complete set of information and assumptions like maximum entropy.

Figure 3 shows the network of exposures in Martinez *et al.* (2013).

Exposures networks – Interbank and interfinancial exposures networks — Figure 3



(a) Interbank exposures (b) Interfinancial exposures

Conclusion

In our experience, the use of micro-data (transactional or highly granular data) has been of great help in attending different information requests from the users with existing information requirements. It has also improved data quality as it allows comprehensive cross checks between data sets. More importantly, however, are the benefits to data users due to the increased analytic value of expanding data to any level of disaggregation and the possibility for supervisors to follow the impact of a particular operation through all data sets.

In times of financial turmoil the advantages of having the precise information surpasses any maintenance costs associated with such a model, nonetheless, there are also great benefits in steady times.

References

Martínez Jaramillo, S., Kabadjova, B. A., Bravo Benítez, B. & Solórzano Margain, J. P. (2012). "Systemic Risk Analysis by means of Network Theory: An Empirical Study of the Mexican Banking System". *Banco de México Working Papers*, 2012–7.

Martínez Jaramillo, S., Solórzano Margain, J. P. & López Gallo, F. (2013). "Financial Contagion: Extending the Exposures Network of the Mexican Financial System". *Computational Management Science*.

Summary of session on "Links between Micro and Macro Data I"

Chair: Aurel Schubert[1]

Paula Menezes (Banco de Portugal) "Pre-conditions for integrating micro-databases"

In order to cope with the insufficiencies of multiple isolated systems, in some cases legacy systems, built with the sole purpose of addressing specific needs the BdP decided to develop a business intelligence (BI) architecture, to be used as reference in all future IS/IT developments in the statistical area and capable of promoting efficient data analysis. The framework is built upon three pillars: (i) a reference data management centre, (ii) a data warehouse (DW) and (iii) a common IT platform. The centralised reference database, developed on the basis of the already existing Reference Information Sharing System, provides the connecting elements of statistical data from different sources and, ultimately, is the main guarantee of the possibility of integrating the information, enabling cross-linking information from different sources and systems. The DW guarantees a central access point to every statistical data, independently of the input source or the production process; a common technological infrastructure across multiple information systems makes it easier to integrate and reuse components and promotes data access efficiency and transparency to final users. According to the BdP experience the recommended strategy in developing a BI architecture includes an incremental development of the individual projects as a way to reduce risk and achieve rapid results, joint project monitoring by IT and Statistics Department, shared responsibilities with the IT Department (i.e. benefiting from its specific technical knowledge and, by the same token, securing the degree of autonomy necessary to meet the changing requirements and business rules) and adequate post-implementation support on the part of the IT Department.

Ivette Fernández (Banco Central de Chile) "Use of tax micro-data for the compilation of the Chilean financial accounts"

In National Accounts or other statistics tax records are frequently used in case of incomplete information on the measured macroeconomic variables. In Chile aggregated (i.e. non-confidential) fiscal data are used as provided by the Internal Revenue Service with a monthly or annual frequency. However, it is crucial to analyse the information through statistical analysis and by comparing it with the detailed information of other sources and surveys, before utilising it. More precisely tax data is used to measure different macroeconomic variables, for both National Accounts and Financial stability, such as monthly indicator of economic activity, production and sector accounts, distributions of the profitability of enterprises; indexes of housing prices, real estate statement and various (internal) economic and financial studies. Currently the Banco Central de Chile is working on a medium-term project, which attempts to integrate different statistical-financial micro data sets, including tax data, financial statements, credit databases, securities-by-securities

[1] Director General Statistics, European Central Bank.

databases and real estate databases with the objective to integrate these databases in a data warehouse. In addition, it is foreseen to use business intelligence tools to visualize, analyze, aggregate and link information of different micro-data sets.

Johannes Turner (OeNB) "Micro-data – the Austrian concept with more granularity as a lesson learned from the financial crises"

The financial crisis has shown that any tailor-made aggregated reporting forms are eventually not flexible enough to cover new data requirements. Instead only micro-data and a suitable analytical framework can answer today's questions. This motivated the OeNB to start a joint project with the reporting banks to build up a flexible and harmonised data collection system.

A key element of the new forthcoming data collection system is the so-called ***basic cube*** that provides an exact, standardised, unique and hence unambiguous definition of individual business transactions and their attributes. While not being enforced by any legal obligation this basic cube is supposed to describe the entirety of a banks information basis and has been modelled jointly by banks and the OeNB. Based on this basic cube, banks aggregate or calculate and report (most of them with a jointly developed software) the second key element of the new system, the – legally obligated – multi-dimensional ***smart cubes***. These cubes allow the compilers and analysts in the OeNB to use the collected data in a consistent and flexible way. Examples of these micro data cubes would be ISIN cube or loan cube.

In particular the SSM opens a window for introducing a single, granular micro data model with harmonised and consistent data. These micro data allow, for instance, analysing individual banks and risks in an efficient and flexible way.

Saban Murat Cakir (Türkiye Cumhuriyet Merkez Bankası) "From data to information and from information to policy-making – the story of the Integrated Company and Industry Analysis Platform"

The final case study was the example of the big project to completely redesign the Turkish Company Accounts and Financial Ratios which illustrated the challenges of the introduction of new compilation and analysis technologies, such as the change management, psychological issues of users and stakeholders, promotion and marketing of new solutions and other financial and institutional constraints.

Integrating micro-databases for statistical purposes

Paula Menezes[1] and Luís D'Aguiar[2]

Introduction

Data are critically important in making well-informed decisions. Poor quality data or, *a fortiori*, lack of data can lead to inefficient allocation of resources and imposes high costs on Society. The strategy defined by the *Banco de Portugal* (hereinafter referred as "the Bank") to deal with the challenge of maintaining its statistics relevant to the users in a shifting and more demanding environment, while attending to the need to keep the respondents' reporting burden at an acceptable level, was to enhance the overall efficiency of the statistical framework by further exploring the largely unused statistical potential of already existing data sources (including the available administrative micro-databases). This approach has been facilitated by the advancements in information systems and technologies (IS/IT), network and communication protocols, database systems and multidimensional analytical systems, which have removed most of the potential shortcomings of having to deal with the huge amounts of data usually associated with the handling of micro-databases.

However, expanding the sheer amount of data available, *per se*, does not necessarily correlate with its value – *"not everything that can be counted counts"* (citation accredited to Albert Einstein). Indeed, there is a need for tools that enable rapid data exploration, permitting multidimensional analysis and cross-reference of multiple sources of information with different granularity. We this in mind, the Bank decided, back in 2008, to lay the foundations of a reference framework for the planning and implementation of IS/IT projects in the Statistics Department, so that the solutions developed within the scope of the different statistical domains could contribute to the incremental construction of a single, coherent and integrated repository of information.

Aiming at micro-data integration

The situation prevailing at that juncture could be described by the existence of multiple isolated systems, in some cases legacy systems, built with the sole purpose of addressing specific needs identified within the restricted scope of each development project. The statistical information was distributed across multiple data "silos" of limited accessibility (see Figure 1. below), and the use of each information system was almost entirely vertical, that is, confined to a business area.

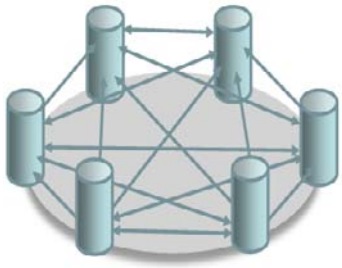
[1] Statistics Department, Banco de Portugal, pamenezes@bportugal.pt.

[2] Statistics Department, Banco de Portugal, laguiar@bportugal.pt.

From the standpoint of the access of the various applications to data from other systems, for the purpose of cross-checking or integrating information, the solutions that have been created consist in extracting and transmitting data from each system to the others, according to the different needs and in varied formats.

Direct integration of ISs

Figure 1



This type of integration model results in a plurality of interfaces, often with poor support; the processes of extraction, transformation and loading occur redundantly and with increased maintenance costs. Another problem associated with this form of integration lies in the difficulty in ensuring data coherence, and in guaranteeing that data updates are reflected equally in the various applications or that reference data are consistent in all databases.

Under this application architecture there were not sufficient resources to effectively respond to the growing need of integration and sharing of information. This problem manifested itself in different ways: from the point of view of those who use the information, by the difficulty of access and cross-check different databases, with diverse structures and formats, or simply by the lack of means to access data directly; from the statistical production side, by the effort involved in establishing and maintaining a complex web of interfaces for the exchange and conversion of data between applications.

To cope with those difficulties, the Bank decided to develop a business intelligence (BI) architecture, to be used as reference in all future IS/IT developments in the statistical area and capable of promoting efficient data analysis. The adopted BI framework aimed to contribute to the construction of a coherent statistical information system for the *Banco de Portugal* (*i.e.*, one that is not just the result of a simple juxtaposition of the ISs in each domain), by creating a framework for the development of the systems focused on the goal of integration. Such framework was built upon three pillars: (i) a reference data management centre, (ii) a data warehouse (DW) and (iii) a common IT platform. The centralised reference database, developed on the basis of the already existing *Reference Information Sharing System* (or SPAI, using the Portuguese acronym), provides the connecting elements of statistical data from different sources and, ultimately, is the main guarantee of the possibility of integrating the information, enabling cross-linking information from different sources and systems; the DW guarantees a central access point to every statistical data, independently of the input source or the production process; a common technological infrastructure across multiple information systems makes it easier to integrate and reuse components and promotes data access efficiency and transparency to final users.

A DW strategy typically falls in one of the following cases:

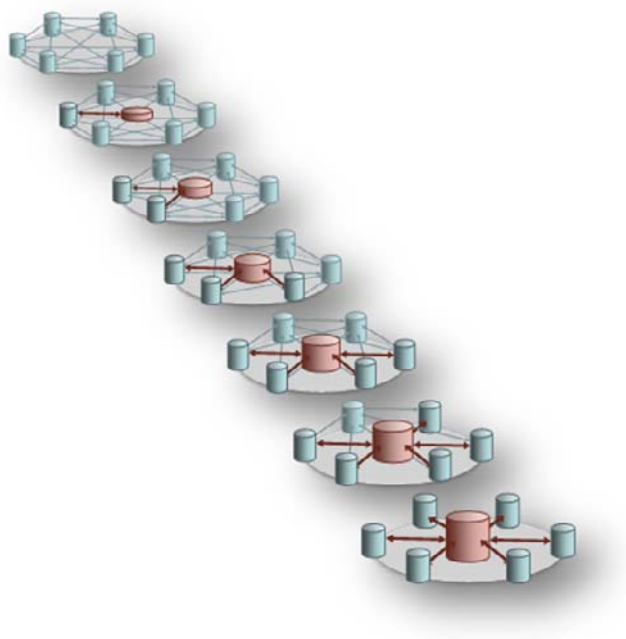
– *Top down* approach, whereby a central DW is first constructed, from which are then extracted specific data marts for each business process. It allows the integration of multiple information domains, achieved through a global development effort. However, its intrinsic difficulty, together with the time and cost required to obtain the first results, explain its high failure rate.

– *Bottom up* approach, whereby several data marts are first developed, on the basis of which the global DW is incrementally built. This allows for reducing the risk of the projects and speeds up the achievement of the first results. The problem with this approach is that it often results in new information silos, when successive isolated projects fail to promote the integration of data from different sources.

The Bank opted for the second approach, on the grounds that in our case there is an intermediate area – the so-called "working data store" –, where production processes take place, accessing data from different domains already stored in the DW – unlike the more common situations, in which data flows from source systems only in one direction and the construction of the DW has no upstream impacts. In fact, the additional complexity arising from this dual dependency, in which information providing processes are at the same time clients of the DW, suggests following an incremental approach.

Figure 2. illustrates a schematic representation of the evolution from the method of direct integration of various production systems to the final model, in which the integration will be done through a central repository of information coherent and consistent.

Phases of the incremental approach — Figure 2



As the number of embedded systems increases, the number of data streams becomes smaller. In this representation, the first attempts at implementing the architecture do not result immediately in a significant reduction of complexity, since all the other systems are still dependent on the current interfaces. Clearly, this process has the potential to last for too long.

Conversely, at later stages of the process each new integration project will allow a further reduction of the existing interfaces, which will be replaced by access to the central repository. This implies some intervention aiming to ensure compatibility with the systems already integrated (which is the cost of the incremental strategy).

One thing to take into account as regards the development of this paradigm is that the projects beyond its scope and objectives must be subordinated to a broader program and abide by a common reference, so that the DW emerges naturally from a set of shared but coordinated efforts.

Concluding remarks

The strategy followed by the Statistics Department to develop a BI architecture was based on the following principles:

1. Incremental development of the individual projects as a way to reduce risk and achieve rapid results.
2. Project monitoring by a joint IT Department/Statistics Department committee, to determine whether the objectives of the project are being successfully met and to keep the BI architecture updated.
3. Shared responsibilities with the IT Department, thus benefiting from its specific technical knowledge and, by the same token, securing the degree of autonomy necessary to meet the changing requirements and business rules, while guarantying reliability, performance and safety to the structural components of the solution.
4. Post-implementation support on the part of the IT Department, to accommodate new requests without compromising the performance of the system, and to promote proper use of tools.

On the basis of our experience we were able to identify and to deal with a number of key success factors, *inter alia*:

1. Rethinking the governance model adopted in 2008, with a view to rectifying a number of insufficiencies that were hampering its effectiveness. A new BI management model is being implemented, involving both the Statistics Department and the IT Department.
2. Taking stock of the information available in the different areas of the Statistics Department (creating data catalogues). In this context, the meta-information model is of paramount importance.
3. Assessing the capacity of the existing infrastructure and whether such infrastructure is appropriate for the BI project that was originally designed.
4. Ensuring a greater operational flexibility of the relational databases and cubes in the mix of solutions to adopt.

5. Overcoming the limitations of the current analytical systems, as regards its capacity to produce new metrics and/or indicators.
6. Training "power users", capable of providing support to other users.

In spite of the difficulties and complexity of these challenges the Statistics Department was able to create the minimum necessary conditions for data integration. Our micro-databases are now based on common reference information, allowing for new statistical products to be launched and new processes to be implemented, with clear advantages to the users. However, if it is safe to say that we are starting to reap the benefits arising from having a higher level of data integration, there are still many issues pending, which require new approaches and extensive research and development efforts.

References

D'Aguiar, L., de Almeida, A., & Casimiro, P. (2011) "Promoting enhanced responsiveness to users' data needs: the experience of the *Banco de Portugal* in exploring the statistical potential of micro-databases". *Proceedings of the 58th ISI Session*. Dublin, August 2011.

Lima, F. & D'Aguiar, L. (2009) "Credit risk transfer – Dealing with the information gap". *IFC Bulletin,* No. 33, August 2010.

Martins, C. (2008) *et al*. "Arquitectura de BI – Sistema de Informação Estatística", *Banco de Portugal*, June 2008 (not available in English)

Martins, C. & Aguiar, M. (2011). "Adding business intelligence to statistical systems – The experience of Banco de Portugal". *Eurostat Conference on "New Techniques and Technologies for Statistics"*. Brussels, February 2011.

Use of tax micro-data for the compilation of the Chilean National Accounts

Ivette Fernández[1]

Introduction

In the measurement of macroeconomic variables, it is common that the complete information is not always available. To get around this problem, tax records are frequently used -provided by the Internal Revenue Service- including variables for measuring various products calculated in the National Accounts.

In this context, this paper briefly explains the main sources and features of tax information that is used to measure macroeconomic variables, shows the information obtained from this data, and explains a project which is currently being developed regarding this matter.

Characteristics of the tax data

The fiscal data being used is provided by the Internal Revenue Service, the information presents a certain level of aggregation and has no reference of the actual informants (for confidentiality reasons) and the data is received both with a monthly and annual frequency, depending on the specific records. The main information used is:

Tax-data sources — Table 1

Information	Form	Selected Variables
Value Added Tax (VAT)	Form 29	Monthly information of 15 variables by industries and legally constituted body.
Income Tax	Form 22	Annual information of 21 proxy-variables by industries and legally constituted body.
Wage Statement (Annual Affidavit of Income of Article 42 N°1)	Form 1887	Annual information of wages paid by employers to salaried workers.
Real estate Statement (Statement of disposal and registration of Real Estate)	Form 2890	Quarterly information of purchase / sale of housing and non-residential properties.

Source: Author's elaboration based on Central Bank of Chile (2011).

The VAT returns include sales and purchases that companies declare monthly; this allows us to calculate the sales of companies for each industry.

[1] Statistics Division, Central Bank of Chile, ifernandez@bcentral.cl.

The annual income tax return contains the information used by the Internal Revenue Service for the calculation of the tax base. In the case of companies, the statements contain some records of the companies' accounts, including balance data sheets as well as data of income statements. In addition, companies must declare their employees and their salaries they have by using the Form 1887 (Annual affidavit of income for the salaried workers).

The real estate statements include the purchases and sales of housing and non-residential properties that people have made. Additionally, this source of information contains the location of these traded properties, along with the transaction prices involved.

However, the data occasionally contains outliers and missing values and therefore it is filtered using statistical methods.

Uses of tax data in the Chilean National Accounts

The tax data is used in order to measure different macroeconomic variables, for both National Accounts and Financial Stability purposes. Specifically in the National Accounts, the information is used to measure the IMACEC (Monthly indicator of economic activity), productions accounts and sector accounts.

In the first case, the IMACEC is calculated by using indicators for each of the different economic activities. Some of them are measured by using tax data, specifically using the monthly value added tax returns (Form 29). In general, this information is used to measure the following industries: some manufacturing industries, the wholesale and retail trade, the hospitality industry, some transport services, the business services, the private health sector and other services.

In the case of the production accounts, the quarterly reports are measured by using the value added tax returns (Form 29) for the case of some of the industries, while the annual accounts are calculated with the Income Tax Return (Form 22) for some specific activities. Income tax statements contain information of some variables that can be used as a "proxy" of the production account. For example, the data contains the information of Sales revenues that can be used as "Output", Cost of Sale as "Intermediate consumption", Wages and Depreciation as part of "Gross Value added", among others. Additionally, the Form 1887 is used to measure the compensation of employees.

In the sector accounts, the non-financial companies sector is measured by using both the information reported to the Superintendence of Securities and Insurance by its supervised companies (which have the requirement to declare their financial statements) and the information obtained from the Income Tax Return (Form 22) for the case of unsupervised companies. This data enables to determine variables such as assets, liabilities, equity, inventories, accounts receivable, interest expenses, depreciation, and others, and to calculate both the current and the accumulation accounts.

Moreover, this data contributes in the analysis of financial stability. For example, in the Financial Stability Report (Central Bank of Chile, 2012), the tax information is used to measure the distribution of the profitability of small and medium-sized enterprises, using a sample of companies that pay first-category

income tax. Additionally, some of the studies or economic papers utilize some variables of the income taxes.

Finally, the real estate statements (Form 2890) are used to estimate indexes of housing prices (not published yet) which collaborate for both the analyses of national accounts and the study of financial stability.

Micro-financial data project

Currently, we are working on a medium-term project, which attempts to integrate different statistical-financial micro data sets, including tax data, financial statements, credit databases, securities-by-securities databases and real estate databases.

The main objective of this project is to integrate these databases in a data warehouse (the chosen tool is Oracle). In addition, business intelligence tools will be used to visualize, analyze, aggregate and link information of different micro-data sets.

Conclusions

Tax information is very useful when it comes to the estimation of economic variables in the Chilean National Accounts. However, it is crucial to analyze the information through statistical analysis and by comparing it with the detailed information of other sources and surveys, before utilizing it.

In the future, we will use this information more thoroughly through the integration of the financial micro databases.

References

Central Bank of Chile (2012). "Financial Stability Report, Second half 2012", Santiago, Chile, December 2012.

Central Bank of Chile (2011). "*Cuentas Nacionales de Chile: Compilación de Referencia 2008*", Santiago, Chile, December 2011.

Guerrero, S., Luengo, R. Pozo, P. & Rébora, S. (2012). "*Nuevas series de Cuentas Nacionales encadenadas: Métodos y fuentes de estimación*". Studies in Economic Statistics, N°90, Central Bank of Chile, Pages 11–14, January 2012.

Pérez, J. (2010). "*Una caracterización de las empresas privadas no financieras de Chile*", Studies in Economic Statistics, N°83, December 2010.

Micro-data – the Austrian concept with more granularity as a lesson learned from the financial crises

Johannes Turner,[1] Erich Hille[2] and Günther Sedlacek[3]

1. Introduction

Experts from the Statistics Department of the Austrian central bank (Oesterreichische Nationalbank – OeNB) and experts from commercial banks are currently working on a joint project to improve the system of data collection. One of the key elements of this project is the collection of granular securities and loan data which can be used for different purposes in a harmonised way. This paper gives an overview about the background, the expected advantages and use as well as the state of play and next steps regarding the planned micro data reports.

2. Background

In earlier times different divisions and departments at the OeNB designed tailor-made reporting forms and processes to enable the collection of data for producing statistics for different purposes such as monetary policy, financial stability, research, supervision and establishing a credit register. These data collection procedures were not harmonised, leading to a lack of data consistency as well as inefficient data quality procedures. Furthermore, the financial crisis has shown that these tailor-made – mainly aggregated – reporting forms are not very flexible concerning new data requirements, so that each change of the collection system led to substantial costs and thus could not be implemented as quick as required by data analysts. All that motivated the OeNB to start a joint project with the reporting banks to build up a flexible and harmonised data collection system.

3. Expected advantages and use of micro data

A key element of the future data collection system is the so-called basic cube:

- The basic cube provides an exact, standardised, unique and hence unambiguous definition of individual business transactions and their attributes.

[1] Director Statistics Departement, Oesterriechische Nationalbank, johannes.turner@oenb.at.

[2] Senior Adviser Statistics Departement, Oesterriechische Nationalbank, erich.hille@oenb.at.

[3] Head of unit Statistics Departement, Oesterriechische Nationalbank, guenther.sedlacek@oenb.at.

- As the core banking systems of different institutions are in most cases not comparable, the basic cube establishes a harmonised database model at a very granular level within the banks.
- Consistency, the absence of redundancy and ease of expandability are key features of the basic cube.
- As suggested by its name, the basic cube comprises the primary data for all kinds of statistics, both supervisory and monetary.
- The basic cube could be used only as a means of data description without being physically realised.
- The basic cube has been developed jointly by banks and the OeNB, but OeNB staff will not be allowed to access the basic cube.
- The OeNB is strongly interested that banks will introduce the basic cube but it will not be a legal obligation.

Based on this basic cube, banks aggregate or calculate and report (most of them with a jointly developed software) the second key element of the new system, the – legally obligated – multi-dimensional smart cubes. These cubes allow the compilers and analysts in the OeNB to use the collected data in a consistent and flexible way. Both elements of the system should also guarantee a quicker implementation of new data requirements and reaction to ad hoc requests.

Two of these multi dimensional smart cubes are designed as "micro smart cubes": the "ISIN cube" and the "loan cube". In the first phase of the development of these micro data cubes the focus lies basically on unconsolidated data, whereas in a second phase it would make sense to investigate the feasibility of collecting such granular data on a consolidated basis.

ISIN cube

Within the ISIN cube assets and liabilities of credit institutions related to securities are collected on a security-by-security basis without any thresholds. The OeNB has already gained positive experience with it's security-by-security data collection system, which was originally developed in the nineties for external statistics purposes. Since the financial crisis this system has become more and more relevant for financial stability and micro-prudential supervision, although the focus of the current security-by-security system still lies on external statistics – partially justified by legal constraints for using these data for other purposes than external statistics.

The future ISIN cube is designed – as an enhancement of the current system – to meet additionally the user needs regarding monetary and financial, securities holdings and supervisory statistics. Thus the cube will substitute some existing reporting templates in the field of monetary and financial statistics as well as supervisory statistics. For this reason legal acts and the reporting templates have to be changed.

The main difference to the current security-by-security system is that some additional securities'/issuers' attributes which are mainly not available from commercial data providers will be collected from banks, e.g. indicators for

- Amount below 100% of repayment of nominal capital
- Trading book / banking book
- Intangible fixed asset / working capital

as well as some detailed information about credit linked notes (underlying information) and asset backed securities (e.g. type of securitisation, type of assets, attachment and detachment point). Furthermore, different categories of valuation are required, e.g. book value, market value, specific provisions, new issues and redemptions. Basic characteristics of issuers and securities, like issuer country and sector, nominal currency, interest rate and maturity information, will be derived from e.g. the Centralised Securities Database of the ECB.

Loan cube

The loan cube is principally based on the current central credit register (CCR) data collection where credit data are collected on a borrower-by-borrower basis with a threshold. At the moment CCR data are mainly used for financial market supervision purposes and for requests of lenders about the existing obligations of potential borrowers (determination of an obligor's degree of indebtedness). The planned loan cube has more dimensions than the CCR reporting template and could thus substitute – similar to the ISIN cube – some other existing reporting templates in the field of monetary and financial statistics as well as supervisory statistics. Additional loan dimensions currently not collected by borrower are for instance:

- Residual and original maturity
- Currency or interest related attributes
- Purpose of the loan
- Repayment attributes

Risk information, e.g. non performing loans, ratings and risk weighted assets, as well as collateral information will be collected on the assets side on a borrower-by-borrower basis similar to the current CCR templates.

4. Next steps

Whereas the design and future use of the ISIN cube is already in a stable phase, some topics regarding the loan cube and the related risk information have still to be clarified. However, there is a common understanding between banks and the OeNB that the final development stage should be a loan cube with integrated risk information on a loan-by-loan basis (where meaningful). But still some necessary questions regarding the timeliness, use, quality management and international environment regarding this loan-by-loan cube have to be answered before a final decision for such a cube can be taken. This will be a main task during the next phase of development.

References

Hille, E (2013). "Recent developments in restructuring the Austrian bank reporting system", *Journal of Banking Regulation* Vol. 14, 3/4, 269–284, July 2013.

Sedlacek, G (2009). "Practical examples of policy relevant uses of security-by-security data", *IFC Bulletin*, 29, 41–47, January 2009.

From data to information and from information to policy making – the story of the "Integrated Company and Industry Analysis Platform"

Çakir Şaban Murat[1]

With a huge experience of data management, after 20 years the Real Sector Data Division of CBRT has decided to overhaul its platform radically. It is and was a revolutionary step towards the challenges of data and information needs of the 21st century by the bank's very strict standards.

This presentation will try to give some flavour how such a complicated project is bound to succeed in an environment where it is doomed to face reactions from different stakeholders, how handling things had been hard, how some have been left temporarily and/or definitely unsolved, how some mildly and some toughly had been made accepted, how fast or slow the project had evolved since its very first inception, given the information needs of authorities and policymakers for them to make sound policy decisions in an ever more complicated global economy. The presentation summarizes the project within a data, technology and change management conceptualization and framework by not giving specific examples but by depicting what and how they have become issues to focus on.

The presentation will not give a standard way of looking at things but will try to exemplify how a classical problem is approached from the eyes of a practitioner.

Current situation – I

Company Accounts and Financial Ratios (The Quartiles) Reports are an aggregate of individual company financials and managerial information prepared and presented by sector and scale. It is bound to be published every year mid-November by law and in the Official Statistics Program's publication calendar. The reports involve the aggregation of the three consecutive year financials and other information by sector and by scale, and help

- Policymakers in assessing medium and long term policy decisions on different sectors and scale groups as well as shaping and taking their decisions about/on structural economic policies.
- Bankers in their credit granting decisions by providing them with sector and scale level financial statement and structure information and quartile ranges for financial ratios at these levels.
- Participating companies in their business level decision making by providing firm level feedback comparing them with their sector and scale group peers.

[1] Real Sector Data Division, Statistics Department, Murat.Cakir@tcmb.gov.tr.

- Researchers from different disciplines carrying out research on real sector by catering miscellaneous data and information.

Currently however, the importance and significance of the study is only valued by a focused group of users due to the current structure of preparation process with low frequency data and lags between the publication dates of comparable reports (though with smaller sample and hence lower representation). This (mis)perception of less usability and practicality due to the above mentioned lags and lower frequencies – though for it being the sole work with the highest representation rate of the whole economy and sectors in GDP sample – should be changed by persuading the users that this work in and of itself should be regarded as a valuable feedback for longer term decision making, by publishing it with a shorter lag, and by providing them with a higher frequency reports along with it using quarter term financials for short and medium term policy making.

How was the situation before the platform? – II

The situation before the platform could be characterised as/with

Sparseness of the system

1. Sparse and needing special effort, energy, knowledge, skills and a whole lot of experience for harmonization and integration of raw, processed and yielded data and information from different sources, platforms and systems,

Proneness to errors and risks

2. Prone/open to miscellaneous errors and mistakes leading to longer processing times, lags in final reports and hence resulting in lower efficiency,
3. With risks of all sorts of data and information losses on PC-based structure,

Deficiencies of personnel

4. Prepared by a considerably small group of field experts with considerable amount of individual endeavour,
5. Deficiencies of institutional solutions,
6. With deficiencies of systematic institutional solutions at some levels of data storage, processing and hardware and software,
7. Risks of authorisation,
8. Prone to risks due to flexibilities in data access described as levels of predefined and uneasy to alter authorization processes,
9. Obsoleteness of architecture,
10. With an although secure but outdated database architecture,
11. Extreme dependence on experts and specific software,
12. With expert dependence on the design and implementation phases of BO, VBA and excel applications and hence prone to operator risk,
13. With software dependencies for in-house applications on limited number of coding and design experts,

High inflexibility and complexity of software and applications

14. With outdated, expert dependent, less flexible and slower mainframe applications with poor visual reporting and processing properties,
15. With inflexible and expert dependent listing, data processing, controlling and monitoring static system structure,

Low automation

16. With less automated sub processes for even the slightest routines,
17. Low performance,
18. With performance inefficiencies due to dependencies in data processing and all sorts of data storage on PCs,
19. With performance and time losses due to the sparse structure of the system that used to consist of 5 different software and application media (MS office, Business Objects, Mainframe, in-house developed, Visual Basic and VBA applications),

Multi-layered and complex reporting

20. With so many layers in reporting,
21. Chaotic, uneasy to transfer knowledge and experience and hence prone to extensive loss of knowledge and experience due to HR mobility,
22. With only manual and slow – due to the bureaucracy – capabilities in the tiniest correction in software applications, and with difficult and expert dependent testing procedures.

What is targeted with the platform? – III

The targeted platform could be characterised as having features enumerated below

Optimal automation

1. With all the standard and non-standard ad-hoc automated interfaces and in routine procedures,
2. *Optimal levels of n-dimensional cube static and dynamic reporting,*
3. With data depots and warehouses that are enabled to store and provide as well as produce interim and final reports and (freeze snapshots) difference analyses,
4. With dynamic reporting facilities,

Batch and automated data collection, processing, analysis and reporting

5. With batch and automated data processing, controlling, transferring operations and with managed layers/levels of confirmation and approval,
6. With batch and automated data collection capabilities through commercial accounting software,
7. With capabilities of automated data reporting and publishing,

Flexible design, coding, visualisation and reporting

8. With facilities of automated and manual as well as flexible visualisation,

9. With flexible design and coding properties,
10. With all sorts of flexible reporting and processing media,

High performance IT capabilities

11. With powerful processing, storage, security and performance capabilities (remote server based),

Timely and fast reporting

12. With timely and fast reporting capabilities,

High flexibility on desktop and server applications

13. With high visualisation and high flexibility properties and powerful ad-hoc reporting on desktop systems and server applications,

Ability to do brand new analyses (Distress and rating)

14. Where rating and financial and operational distress analyses can be carried out,

Experience and knowledge transfer

15. With all types and levels of training and tutorial facilities on the platform so that experience and knowledge can be easily transferred.

Who and/or what are the stakeholders? – IV

The platform involves various stakeholders from different groups concerning the current and potential future costs and benefits. Without any rank of importance these can be enumerated as

1. Policy-makers,
2. Academic and independent researchers,
3. Economic agents composed of real and financial sectors and households,
4. Statistics, research and information technology units/departments of the institution (CBRT), and
5. Other private and public economic institutions and those producing statistics and the academia.

   Stakeholders can also be classified as inter and intra organisational.

What are the costs and benefits?

Costs and benefits related to the new platform regarding the both groups that is inter and intra organisational stakeholders:

Benefits:

1. Producing and presenting better quality and timely reports will
   a. Help policymakers make healthier and better targeted economic policies,

   b. Lessen the information asymmetry from the viewpoint of economic agents,
   c. Help research units making better quality analyses thus offering more meaningful and exact policy suggestions, and
   d. Help economic agents taking better economic and financial decisions.
2. Gathering, processing, analysing and presenting more detailed data about real sector will
   a. Help financial and real sector decision-makers shape their short, medium and long term decisions and investment behaviours,
   b. Help statistics units cater better quality processed data and reports, and
   c. Help policymakers to make finer adjustments in their short, medium and long term policies by offering them more exact, higher quality and detailed reports.
3. With a macro point of view
   a. IT's better quality and larger project experience, and with higher technology its capabilities to improve the product quality and produce better projects will increase internally, and
   b. Capability to scrutinize in depth the complexities and the interdependences of the economic relationships will be enhanced externally.
4. With high tech hardware and more efficient software
   a. Time costs will be lowered,
   b. Production and training unit costs will decrease,
   c. Data and information quality will be improved, and
   d. With faster feedback while the reaction times of all the stakeholders are diminishing, costs related to decision mistakes will fall.

Costs:

1. In terms of legal constraints
   a. Due to the necessity of definition of data customer relationships, time and legal costs may increase.
   b. The project budget, hardware and software procurement procedures being heavily constrained by laws in CBRT, related costs will inevitably augment during and after the project.
2. In terms of resources devoted to the project
   a. Time unit costs related to project requirement assessment, case analysis involving all procedures and processes, design and coding will increase.
   b. Hardware and software costs allocated to the project will increase.

Final Analysis:

In the medium and long run, cost reductions resulting from efficiency improvements and potential benefits from the quality enhancements of the final product will

surpass the project costs and will make them acceptable and the project be a viable one.

What lessons are learnt? – V

A new project is – though mostly a betterment to and an improvement of what it was before – usually perceived a challenge and thought of as a detriment to how things were/are handled before and now. Therefore, it is never easy to put new ideas into practice unless the one who comes up with them is willing to deal with all the issues throughout the whole process.

There are a couple of points to clarify in a new project to take care of and/or to be careful about so that the project will be successful. These points include the a) change and change management, b) the management as a whole, c) the psychological issues, d) marketing and persuasion management, constraints and e) other independent issues. This way of categorisation is way-out away from a typical classical project checklist, but as I promised in the beginning, this very heading is not a theoretical but a practical one. One should be aware of all the features of the above issues. These include:

1. ***Change:*** Brand new radical projects require change management rather than data and tech management. Change management is a painful long term ardour that demands relentless effort. Change demands leadership skills rather than management skills most of the time.

   Human side of the change is important. All levels of HR from the lowest to highest should be convinced about the change and the necessity of the change.

2. ***Management:*** There a more quite many management issues than one can imagine. Management demands continuity.

   Time management is one of the most important components of the project management. The longer the project takes – at all levels of hierarchy – the lower the faith in the process becomes; the whole process may enter into a vicious circle due to boredom, prostration and indifference.

3. ***Marketing and Persuasion Management:*** The target audience and the importance of the produced data and reports should be well communicated to the stakeholders so that the importance and the value of the project is well appreciated and enough support is provided by the stakeholders especially the upper management.

4. ***Constraints:*** Budget constraints and cost factor usually play the sole most important role in the projects although it should be considered together with the potential benefits from the project. Project costs should not be solely thought of as IT costs; potential gains and benefits should be considered together with them. The project customer should have its own say on the budget and be given the ownership of its own IT budget with respect to the project(s).

   Legal constraints on inter-organisational and governmental institutions wide data exchange, HR and budget constraints are the other most important factors/constraints to look after in project management.

5. ***Other Concerns:*** If one still is willing to proceed, one should bear in mind that:

   a. Dealing with data is inherently a dull, disrespected endeavour with low benefit, and hence depends heavily on the character, willingness to learn and to produce, and patience and love of the person who deals with it.

   b. IT units (though theoretically not a necessary condition) should take IT sophistication of the customer seriously and assess whether the customer requests are properly directed. (For instance, in this particular project, flow charts, conceptual designs of data entry forms and reports, and all pseudo-codes –thousands of lines- are all prepared by the customer unit's team leader.) Projects are not/should not be taken as granted as one man shows.

   c. Research units and authorities should take the theoretical sophistication of the unit seriously. Most of the time the one who produces the data knows it best and can analyse the outcome the best. Where there seems to be any flaws and/or shortcomings it may be appropriate to question the data and the outcome as there should always be a cross checking between and among the different units/departments of the organisation.

Final suggestions – VII

Intrinsically, project management most suffers from human factors rather than others. Project's own properties – though not as much as the human factors – play the second most important role as well. Hence, for a successful project management the properties of these factors should be well known. These properties can be summarized as follows:

Human Profile: All stakeholders

Project leader from the customer unit should be a generalist, a practitioner, an open-minded, a self-motivated, curious, patient, hard-working, detail-oriented/ obsessed field expert for persuasion, design, controlling and monitoring the process and budgeting phases. A generalist is a person who knows the IT procedures, the statistical and the economic theory behind the raw data production processes.

Project team members of the customer unit should be well aware of the whole process also being aware of what is being demanded from the IT department. They should also have been well persuaded about the change and the necessity of the project. The other properties of the team members overlap with those of the team leader.

All outsiders should have a reasonable say on what to demand from the providers and should be able to use the final product efficiently and effectively. Therefore, though indirectly they have the right to impact the whole process.

The upper management including the ultimate policy-makers should have been well persuaded about the necessity of what is planned as a final product and they should also have been convinced about the value of the data and information as well as the potential reports and derived information from the data.

Project Profile:

The project framework itself should possess some features so that the whole process can be dealt with smoothly.

The project structure should be less bureaucratic, it should have its own flexible budget and decisions on budgeting should be left to the project teams from both departments along with the responsibility to report on the decisions. It should be less hierarchical. The less the layers the faster the project will be finished. A matrix type organisation seems to be most appropriate for such structures. Finally, team work should be motivated and sought for a radical project to be successful.

Integration of different micro-databases – a significant value added for statisticians

Francesca Monacelli[1]

Some preliminary concepts

Micro-data are defined as the information describing the characteristics of the units of a population. With regard to the quantitative data collected by the Bank of Italy, such population can correspond to the reporting entities when the reporter supplies information about itself (for example in supervisory reports or in non financial enterprises direct reporting). Alternatively, it can refer to third parties when the reporter is a commercial data provider/other Authority or when data are derived from operational platforms (*e.g.* TARGET). There is also a third instance in which the reporter produces in the same data flow information about itself and about its related parties as a kind of personal attribute. This is the case, for example, of borrowers' data supplied with Central credit register reports.

Another family of micro-data we deal with are the Registers, which hold the qualitative information of the units of a population. The most relevant are the Subjects Register, which contains the identification variables of individuals, public entities and enterprises, and the Securities Register, which contains the descriptive variables of the securities issued in Italy and of foreign securities held in the banks' portfolios. Both Registers are fed by multiple sources.

Although these categories of micro-data are logically different to one another, our system considers them simply as different specific occurrences of a reported data. In fact our methodological approach and inquiry tools are generic so that we are able to treat them in the same uniform and unique way. For this reason, we indistinguishably call them Reported data.

The same approach and tools are also applied to aggregated/compiled data (*i.e.* macro-data) and, in our experience, uniform treatment of information represents a significant value added for statisticians.

The pursuit in the integration of micro data-bases

So far we have managed to build a single company-wide statistical Data Warehouse for multidimensional data. We are currently in the process of merging into the Data Warehouse also the time series data-base with macroeconomic statistics used in the Economic research Area so that, in the end, all users will be able to compute complex statistics and aggregations, by putting together multidimensional quantitative and qualitative data and time series, using the same tools and Data Dictionary.

[1] Statistics Collection and Processing Department, Banca d'Italia, francesca.monacelli@bancaditalia.it.

The Data Warehouse has been designed to be used for multiple purposes across the Bank (*i.e.* research, supervision, payment systems surveillance, statistical publications, statistical flows for external entities) and for any category of information such as quantitative, qualitative reported data as well as data quality management indicators.

Its distinguishing factors are that it is governed by a unique Data Dictionary with harmonised concepts, it is hosted on the same technological platform and it can be inquired with the same tools regardless of the category data. More in detail, with the same metadata managing software and by means of the transformation rules present in the Data Dictionary we are able to compute any complex statistical multidimensional or time series output on the basis of the different Reported data. The output can be additional Data Warehouse tables with ready to use statistics, a statistical flow for external parties or a publication. The common element of these 3 types of output is that they are generated from the same elementary data.

With regard to ready-made statistical tables which enrich the Data Warehouse, it is worth specifying that they represent an intermediate step between the Reported data and the final statistical production in order to facilitate the user computational tasks. In particular, ready to use statistics enhance the efficiency of the statistical analysis as they save computation time and space. On the other, the existence of data compiled following agreed and fixed definitions ensures that the same concept is homogeneously adopted by all of the Bank's users and in the external dissemination, thus reducing ambiguity.

Furthermore, when different groups of users require that one concept is shown with different facets (*e.g.* with regard to the level of detail/aggregation or because they are authorised to view the data only after a certain time lag) we create additional tables in the same section in order to accommodate these needs. Therefore it can happen that the same concept is shown across the tables of the same section according to different criteria.

It is important to specify that access rights to the Data Warehouse are organised according to strict internal rules based on the "need to know principle" whereby users are grouped according to the institutional area they operate in and are consequently allowed to view all or only some of the Data Warehouse tables.

The Data Warehouse is organised in different logical thematic sections (data-bases): Supervision, Financial instruments, Central credit register, Monetary policy operations, Money market, Payment systems, Registers etc. Each section comprises tables with the relevant Reported data which can be used in a flexible and personalised way; this represents the foundation layer of the Data Warehouse. In what follows I will describe the solution that we have put in place to create the Central Credit Register [CCR] section which, for its complexity, is a good example of integration of the different elementary data as a way to offer a service tailored to the specific and diversified needs of researchers and supervisors. I will also outline the solution found to integrate these data with other individual information stored in non harmonised Registers.

The internal dissemination of Central Credit Register's data

The CCR's section has been built with the aim of managing the following issues:

- efficient elaboration of large volumes of Reported data
- sensitivity of borrowers personal data in terms of privacy protection law
- ensuring that statistics on credit default rates and other commonly used concepts are compiled in the same way by all Banks' users
- identification variables of the borrowers are separately stored in the Subjects Register
- exploiting as much as possible the information concerning the borrowers characteristics and the relationship lender-borrower.

All of the above has determined the implementation of a rather articulated section based on the CCR reported data which are processed in different ways with a view of satisfying different user need and constraints. It is composed by the following sub-sections, the first 2 focusing on the borrower and the 3rd focusing on the lender:

1. CCR-N. The basic layer contains the data reported to the Central Credit Register; the borrowers are identified by the Subjects Register code. The individual data are also enriched with some ready made indicators. By means of the borrower's code, this information can be freely integrated by users with the Subjects Register or it can be cross related to other quantitative individual information collected with supervisory reports. Since it is possible to easily identify the borrowers, we maintain a detailed log of every single access to this data. Authorised users are supervisors, researchers and also applications (such as the one producing the Credit register personalised return flows of information, banking supervision models, statistical publications).
2. CCR-A It. is an exact copy of CCR-N, however the borrowers code is an alias and can be associated to a reduced version of the Subjects Register with more limited information which does not allow the identification of any subject specifically. As a consequence no access log is needed.. Since these data are anonymous, in addition to supervisors and researchers, authorised users are also restricted groups operating in other areas which have an institutional interest to analyse these data.
3. CCR-S. It contains several ready to use statistics based on a joint use of the CCR-N section and the Subjects Register focusing on the single lenders or on the credit system as a whole. The browsing tools and the authorisation level of this section are the same as for CCR-A.

Frequently users need to put together information related to the same group of subjects coming from the data-bases present in the Data Warehouse and from other data-bases with individual data acquired from commercial providers and therefore not harmonised. An example of the latter is the Central balance sheet data-base. The problem they face is to identify the same subject in the 2 data-bases when the 2 sources have different metadata and can describe the subjects with slightly different values. This activity requires specific skills and may be very time consuming if applied to large number of entities.

To solve this problem one can decide either to harmonise the metadata of external sources, so as to allow an integrated use with the rest of the Data Warehouse, or to leave the integration to the moment when the data are actually used by supplying the users with an appropriate way to reconcile different sets of metadata. We opted for the second solution by developing a general application that is able to compare the identification variables of any pair of Registers returning the pairs of codes that correspond to the same subject. The comparison is based on the same algorithm used in the Subjects register to assign a code to a new entity. The reconciliation of the unmatched subjects is performed manually by specialised staff.

The integration of the time series data-base

We are currently undergoing a major technical transformation with the aim to achieve a wider integration of data-bases by merging into a unique company-wide platform the 2 Data Warehouses currently available to internal users: one being the one described so far and the other containing macroeconomic information in the form of time series. In particular, this data-base is composed by the time series built on the basis of the Supervisory reports, the ones acquired by national and international institutions (National Statistical Office, OECD, IMF, ECB, etc.) and others created by the researchers as a result of statistical and econometrical analysis.

The project entails two fundamental preconditions. The first is the complete harmonisation of the Data Dictionary although we have decided not to harmonise the metadata derived from external sources, instead, for this data we have created specific tables to link the internal and the external codes. The second is the definition of a comprehensive Information model that not only it is capable to manage different categories of data, their rules and logical dependencies but also can cover the statistical production process as a set of integrated activities (collection, compilation, dissemination). Once the merge will be finalised the production of time series from Reported data will be more efficient, thus reducing the process overheads, and they will be updated at the same time as the elementary data. In addition it will be easier for researchers to analyse time series and other cross section data together.

Final remarks

The unification of different data-bases is not an effortless task but we are convinced that it will give a greater value added to our internal and external users. By means of the same Data Dictionary, same Inquiry tools and same data representation model we will be able to exploit all the micro and macro information held in the future integrated statistical platform to produce reports/analysis in a more efficient manner and, for the benefit of our external users, we will also be able use the same metadata in the statistical outputs, regardless of the category of data they are based upon.

Do macroprudential tools require micro-data?

Perttu Korhonen[1]

Overview

As they stand, macroprudential tools for the Financial Policy Committee (FPC) imply significant data requirements. The Draft Policy Paper published in January highlights ways in which data need to marry an economic and financial system level perspective with a level of detail not currently collected under any single reporting regime.

Particularly demanding tool is the Sectoral Capital Requirement (SCR). The policy makers must be able to identify the sector of financial activity where systemic threats are emerging. A sector can be a broad category of banking products and/or counterparties, and sectoral data are to some extent available through traditional micro-prudential or statistical collections. However, when required, the FPC seeks to take a more targeted approach, adjusting capital requirements to target only on exposures to those subsectors that contribute to the systemic threat.

It appears that unprecedented flexibility from the available data dimensions will be required. It may be proven that only micro-data can provide sufficient basis for the analysis and impact assessment by the FPC.

Macroprudential tools

The macroprudential tools build on the existing microprudential regime, under which capital requirements depend on an estimate of the riskiness of each exposure. Both the Countercyclical Capital Buffer (CCB) and SCR provide the FPC with a means to change the capital requirement when threats to financial stability are emerging.

The CCB allows the FPC to change capital requirements above normal microprudential standards for all UK loans and exposures. A potential issue is that an apparently profitable boom in one part of the economy or financial system can be left relatively untouched while lending in another may be significantly reduced.

The SCR tool is an important complement or alternative to the CCB. The SCR allows the FPC to change the capital requirements above microprudential standards on exposures to specific sectors that pose a risk to the system as a whole.

The SCR is therefore more targeted than the CCB while both tackle the same issue. In terms of data requirements, the CCB is generally highly aggregated while the SCR may require the same measures to be broken down by various dimensions.

[1] Statistics and Regulatory Data Division, Bank of England, perttu.korhonen@bankofengland.co.uk.

Various slices and dices may be needed which are not in the current data. The FPC will be able to adjust the SCRs for banks' exposures to three broad sectors:

- residential property;
- commercial property;
- other parts of the financial sector.

Sometimes the systemic threat can stem from subsectoral level. For instance in residential property exposures it may become necessary for the tool to target only high LTV or high LTI loans.

For intra-financial exposures, the FPC needs the capacity to differentiate at least between secured and unsecured loans, and derivative and bond exposures to financial sector entities. Such entities can be banks, building societies, investment firms, insurers, various types of funds and a range of other regulated and unregulated financial institutions.

When risks are particularly concentrated, it may be desirable to apply SCRs to certain financial sector exposures only. This could be done in two main ways:

- By adjusting the SCRs for exposures to specific types of financial institution. For example, prior to the current crisis that could have been monoline insurers, special purpose vehicles, or non-bank lenders.
- By adjusting SCRs for specific types of intra-financial system activity, or by instrument. Prior to the current crisis these could have been repurchase agreements using low-quality collateral.

In addition to the dimensions needed for identification of the appropriate sectors, there are also further requirements for the data provided.

Firstly, and consistently with the microprudential framework, the tools can be applied at both solo and consolidated level. Whilst regulatory returns generally provide the split, current statistical reports do not provide anything on the latter.

The SCRs can be applied whether exposures are held in banks' trading or banking book. That requirement may limit the usefulness of micro-prudential data which do not always treat banking and trading book exposures symmetrically.

A "look-through" approach will be taken in classifying the sectors. For instance, an increase in the SCR for banks' commercial property exposures will also apply to exposures to securitisations backed by commercial real estate loans. This level of detail generally is not collected on the existing forms.

Consistent with the microprudential framework, the tool will apply to undrawn credit lines and other such contingent obligations as well as loaned amounts. This requirement limits the usefulness of existing statistical collections.

In contrast to the microprudential framework, capital requirements may be applied to only one of the stock of exposures or the flow of new exposures. Again, microprudential reporting is generally not concerned with flows, which limits the use of current data in this respect.

The SCRs can be applied irrespective of the ultimate borrower's domicile; or their exposures to various countries. Currently, there is only limited and high-level data available about banks' international exposures.

Why might we need micro-data for the SCR?

Firstly, micro-data provides the flexibility required to determine the appropriate subsectors. In uncertain times, it is a general theme for users to require additional data to help answer a wider range of questions. A traditional form is generally a 2-dimensional cross-tabulation. Certain COREP forms are 3 dimensional as the same form must be reported several times for example in different currencies. And sometimes the forms have an extra dimension nested within the two or three main ones. For the SCR, it appears, even this amount of dimensionality in the input data is not enough. Singling out a relevant subsector may easily require five or more dimensions and it is very difficult to foresee which combinations will be needed.

Secondly, micro-data would enable pre-assessment of the impact. Before adjusting a capital requirement, the FPC must attain whether the chosen subsector is the appropriate one – *i.e.* whether it really is the one booming, or if in fact there is another subsector within that sector which should be the actual target. In addition to these first round effects, micro-data are likely to be required to model any second round effects.

Thirdly, the principle of 'look-through' requires very detailed data of the composition of different exposures. The FPC should be provided with means to monitor banks' exposures to a certain sector or subsector, regardless of the type of the instrument they are holding. An exposure to over 90% LTV residential property should be treated similarly, whether it is a direct loan or a part of securitisation.

Moreover, micro data allow for collecting and constructing the solo level and the different consolidation scopes (*e.g.* CRD, IFRS) from the same underlying data, and discussing differences between them. It also helps in shedding some light on the shadow banking sector as all firms' direct counterparties are recognised.

Individual counterparty-level reporting would also enable some important structural macroprudential analysis, such as financial network analysis. While the cyclical tools take the firms' RWAs as given, challenging the RWAs has been an important topic for the interim FPC. Collecting this in micro-data could allow for informed and therefore more effective challenge.

References

Bank of England (2013). "The Financial Policy Committee's powers to supplement capital requirements. A Draft Policy Statement."

The use of micro-data to develop remaining maturity data for debt securities and loans

Kuniko Moriya[1]

Introduction

The Bank of Japan (BOJ) is currently developing sectoral debt securities and loans with remaining maturity of short term and of long term to capture macro-financial risks by different sectors in the economy, making the use of micro-data of individual financial institutions. At the moment, BOJ publishes sectoral debt securities and loans in Japan's Flow of Funds Accounts (J-FFA) which comprehensively records financial assets and liabilities in matrices of 43 sectors and 51 financial transactions. However, as developing remaining maturity data entails more detailed data than J-FFA currently uses, BOJ has to rely on micro-data from individual financial institutions. This paper illustrates an example of linking micro-data to macro-data, and discusses some issues with the use of micro-data.

Some background

The current project is related to the recent international initiatives aimed to fill data gaps to improve the identification and monitoring of risk build-up in financial systems: the G-20 Data Gaps Initiative (DGI), and the IMF's SDDS Plus. DGI's recommendation #15 calls for developing data in internationally comparable sectoral accounts and balance sheets in common templates (matrices of sectors and financial instruments). The templates, similar to a typical flow of funds accounts' format, include debt securities and loans with remaining maturity of one year and less, and of more than a year, albeit as encouraged sets. Consistent with DGI, countries adhering to the IMF's SDDS Plus, a new top tier data dissemination standard, would also be expected to disseminate data for government debt with the same remaining maturities as DGI, albeit as memorandum items.

Given the international initiatives, BOJ has acknowledged the importance of developing remaining maturity data, respecting and fulfilling international data dissemination standards although reporting is on a voluntary basis. Independently, BOJ considers those data as useful for capturing Japan's financial and economic structures in its own merit. In fact, BOJ proceeds to additionally develop data of bank deposits with remaining maturities and loans by different types of interest rates (fixed and floating) to help identify and monitor the build-up of financial risks.

[1] Research and Statistics Department, Bank of Japan, kuniko.moriya@boj.or.jp.

Data availability and data collection

Macro-data of debt securities and loans by remaining maturities currently available in Japan are quite limited. A readily and publicly available macro-data source is "Financial Statements Statistics of Corporations by Industry (FSSC)," in which the Ministry of Finance (MOF) compiles loans to non-financial corporations. Other macro-data sources include Japan Securities Depository Center (JASDEC), which stores book entry transfer data for corporate securities and "Debt Management Report (DMR)," in which MOF publishes data on government bonds. With those data sources, we will be able to fill in most cells in the template on the liabilities side.

Data Sources for Compiling Statistics — Table 1

	Assets (holding) side	Liabilities (issuance) side
Debt securities	Micro-data of financial institutions	JASDEC, MOF (DMR)
Loans	Micro-data of financial institutions	MOF (FSSC)

In contrast, to fill in sectors on the assets side, we have to rely on micro-data, taking into account the reporting burden incurred by financial institutions. Although not systematically collected by BOJ, certain types of financial institutions are, by regulations, required to disclose data of debt securities and loans by remaining maturities: for example, the Banking Act applies to deposit-taking financial corporations and the Insurance Business Act to insurance corporations. For other types of financial institutions, credit unions, albeit not obliged by regulations, internally collect data of debt securities and loans for their own purpose of risk management. Although not required, most financial institutions have agreed to report their data to BOJ on a voluntary basis – otherwise, we would have to collect data from some 500 individual financial institutions' financial statements and reports by incurring a tremendous amount of administrative costs.

From micro- to macro-data

With the micro-data of financial institutions as well as the existing macro-data described above, we will be able to estimate the following items in the template: on the assets side, for debt securities and loans, deposit-taking financial corporations, insurance corporations, and public financial corporations; on the liabilities side, all domestic sectors for debt securities, and non-financial corporations for loans.

On estimation, after compiling micro-data, we calculate ratios between debt securities/loans with remaining maturity of one year and less, and of more than a year for subsectors. We then calculate macro-data in subsectors by multiplying those ratios with amounts outstanding in corresponding sectors in J-FFA (Table 2).

Derivation from J-FFA Data (Rec. #15 Template, partial) Table 2

	Non-Financial Corporations	Financial Corporations	General Government	Households	Rest of the World
F1 Monetary gold and SDRs	Data avairable in J-FFA				
F2 Currency and deposits	Data avairable in J-FFA				
F3 Debt securities (D)	Data avairable in J-FFA				
One year and less	a/(a+b), calculated from micro data, times D				
More than a year	b/(a+b), calculated from micro data, times D				
F4 Loans (L)	Data avairable in J-FFA				
One year and less	c/(c+d), calculated from micro data, times L				
More than a year	d/(c+d), calculated from micro data, times L				
⋮				⋮	

Issues with micro-data

One important issue surrounding the use of micro-data is administrative costs to collect and process micro-data. Accordingly, in cases where micro-data are not readily available or costs of data collection outweigh benefits for corresponding sectors in J-FFA, ratios for certain components are applied to impute their subcomponents. For example, if deposit-taking financial corporations consist of domestic and foreign banks, ratios for the former can be applied to impute the latter rather than collecting individual foreign banks' data. In fact, amounts outstanding for foreign banks in Japan are small enough to reasonably use such an imputation method without significant impacts on outcomes.

Another issue is that calculating methods of "remaining maturity" data vary across individual institutions, as the Banking Act does not stipulate them. While debt can be repaid all at once at the end term ("redemption at maturity") or by a series of instalments with predetermined schedule over the course of borrowing period ("scheduled repayment"), when *recording* debt on their books with the latter case, financial institutions could choose "redemption at maturity" as well as "scheduled repayment." Accordingly, as illustrated in Figure 1 below, outcomes reported by individual financial institutions could vary, depending on which methods to use. Instead of painstakingly investigating all individual financial institutions, after examining various sized samples, we have concluded that it is statistically acceptable to proceed with existing differences in reporting methods as differences would have negligible impacts on estimated ratios between remaining maturities.

Illustrated Example of Different Calculations of Remaining Maturities

(Face value: 100; Maturity: 5 years) Figure 1

Different outcomes calculated for installment debt

	One year and less	More than a year
Redemption at maturity	0	100
Scheduled repayment	20	80

Concluding remarks

BOJ has been working on a project to develop sectoral data of debt securities and loans with different remaining maturities by collecting micro-data from individual financial institutions as well as by making the use of the existing macro-data. BOJ is currently in the midst of collecting data from financial institutions with publication projected early 2014.

Using micro-data to develop macro-data has unveiled some specific issues such as projecting micro-data into macro-data, and standardizing data definitions across individual reporting parties. Furthermore, as we recognize the burden incurred from reporting micro-data can be significant, it is only after exhausting all the possibilities to use publicly available data, should we then consider choices of micro-data. While needs to fill in data gaps are ever increasing as evident from the recent international initiatives such as G-20's DGI and IMF's SDDS Plus, BOJ continues to improve and develop financial statistics it publishes by weighing benefits for users and costs for reporting parties.

Using trade repository data for systemic risk monitoring

Erik Heitfield[1]

The 2008 financial crisis revealed significant problems with the functioning of over-the-counter (OTC) derivatives markets. In the years and months leading up to the crisis market participants were able to build up large counterparty risk exposures to AIG Financial Products and other highly interconnected institutions. When these institutions ran into trouble in the fall of 2008, lack of transparency contributed to a breakdown in the functioning of some segments of the OTC swaps market. While most firms presumably had a clear understanding of their own direct exposures to particular troubled institutions, they could not clearly identify the exposures of other potential trading partners to those institutions. Faced with an environment in which they did not know who was at risk and who wasn't, many market participant simply chose to withdraw from trading for a time, harming market liquidity and thereby making it more difficult for all firms to manage their risks.

A year after the crisis, in September 2009, the Group of Twenty (G20) leaders committed to a long list of reforms aimed at strengthening the foundations of the world's financial system. Among these were commitments that by year-end 2012 all sufficiently standardized OTC derivatives contracts should be traded on exchanges or electronic trading platforms and cleared through central counterparties and that all OTC derivatives contracts should be reported to trade repositories. In its October 2010 report, *Implementing OTC Derivatives Reforms*, the Financial Stability Board enumerated the benefits of having swaps data reported to trade repositories.

By providing information to authorities, market participants and the public, trade repositories will be a vital source of increased transparency in the market, and support authorities in carrying out their responsibilities, including (i) assessing systemic risk and financial stability; (ii) conducting market surveillance and enforcement; (iii) supervising market participants; and (iv) conducting resolution activities.

In the United States, the G20 trade reporting commitment is being implemented through Title VII of the Dodd-Frank Act, which requires that data on swaps transactions be reported to Swap Data Repositories regulated by the Securities and Exchange Commission or the Commodity Futures Trading Commission. The European Market Infrastructure Regulation (EMIR) includes similar requirements for swaps involving European entities.

The Federal Reserve Board is currently using data on credit default swap transactions from the Depository Trust and Clearing Corporation's (DTCC's) Global Trade Repository to monitor the CDS market with an eye towards identifying developments that may constitute sources of systemic risk. We approach this problem in two ways: by monitoring the risk exposures of individual institutions that may be systemically important, and by monitoring the network structure of the CDS market as a whole.

[1] Research and Statistics Division, Federal Reserve Board, erik.heitfield@frb.gov.

At the micro level, we use trade repository data to monitor individual institutions' exposures to counterparties and market risk factors. We examine institutions' exposures to their trading counterparties by measuring the notional value of trading positions with other market participants with respect to single-name reference entities and CDS indexes, broader industry sectors, and the CDS market as a whole. We monitor the exposure of individual institutions to important market risk factors by examining the net notional value of their positions with all counterparties at the same levels of aggregation (reference entity, industry and total market). In addition, we aggregate information on the maturity dates of institutions' CDS positions for relevant sub-markets (corporates, sovereigns, etc.) to understand how those institutions would be affected by shifts in the term structure of CDS spreads. Similarly, we track institutions' net positions in the CDS tranche market to understand how they would be affected by changes in market implied correlation of default rates across reference entities.

We apply network analysis tools to identify sources of risk concentration and frailty in the CDS market. We use measures of degree centrality to identify particular buyers or sellers within the network of bilateral CDS positions that are particularly highly interconnected to other market participants. Network-wide centrality statistics tell us how buy-side and sell-side network concentration has changed over time. Other network statistics, such as clustering coefficients and measures of connectedness, provide information on the extent to which trading is intermediated through third parties and on the robustness of the network structure to the removal of key nodes. While the economic interpretation of these types of reduced-form network statistics is not well developed, tracking them over time helps us to understand how the structure of the CDS market is evolving.

Our use of trade repository data for on-going systemic risk monitoring is at an early stage, but our experience thus far suggests that detailed data on bilateral swaps positions can be an important source of information on the risk exposures of systemically important institutions and the health of vital financial markets. We hope to expand our analysis of trade repository data beyond the CDS market into other swaps market segments including interest rate and commodity swaps as these infrastructures develop.

New micro-data on household wealth in Denmark: perspectives for national accounts

Martin Oksbjerg[1]

Introduction

Danmarks Nationalbank is, in cooperation with Statistics Denmark, developing a new annual statistical micro data set on the full population of households' assets and liabilities. A preliminary version of the data set has been used in two analyses by *Danmarks Nationalbank*. A study of the Danish households' comparatively large gross debt used the data to match liabilities with household assets at the household level (Andersen *et al.* (2012 I)). It found that the largest share of liabilities, assets and income is concentrated in those households who generally have large net wealth. Also, Andersen *et al.* (2012 II) used detailed knowledge of the risk profile of individual household mortgage loans to show that the economic situation of most families is robust to shocks. Around the beginning of 2014, Danish research institutions will be given access to the new micro data set.

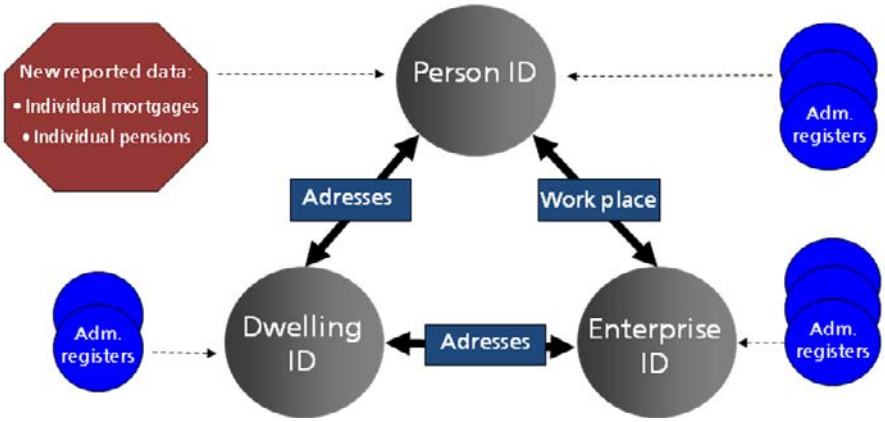
While developing the new data set, it is important to assess its consistency with other existing statistics. This paper compares the aggregated preliminary data with national accounts for the household sector. Several categories of household net wealth seem consistent between the two, but more analyses are planned in this regard.

The statistical information setup in Denmark

Danmarks Nationalbank is not the only central bank seeking detailed information on household balance sheets. Several others already have surveys and credit registers to collect this information. For instance, the Eurosystem Household Finance and Consumption Network recently completed the first harmonized survey on household balance sheets in the Euro zone (ECB, 2013). *Danmarks Nationalbank*'s goal is similar to that of the other central banks, but due to institutional specificities in Danish public administration, the approach is based on administrative registers instead of surveys.

[1] Statistics Department, Danmarks Nationalbank, mok@nationalbanken.dk.

The statistical information system in Denmark including new micro reports Figure 1



Source: Own chart with inspiration from Statistics Denmark.

In Denmark, public sector administrators collect and use centralized administrative registers containing information on all citizens concerning many parts of life. Figure 1 shows how these registers are based on three different units: individuals, dwellings and enterprises. Since all the units in the registers have an ID, additional information on addresses and workplaces can link together information in all registers.

The statistical micro-data set will combine several administrative registers on demography, education, income, socioeconomic situation and balance sheet information. The information on household assets and liabilities is collected mainly from administrative registers from tax authorities. Denmark has a pay-as-you-earn tax system where tax is deducted at the source. For instance, employers report and collect income tax to the authorities on behalf of their employees and banks do the same concerning tax redemption on interest payments. This systematic collection of information from household counterparties makes it possible for the Danish statistical authorities to collect information on most parts of the balance sheet from all 2.8 million Danish households with no additional reporting burden.

However, information on some balance sheet items is insufficient. Firstly, the registers contain no information on the comparatively large stock of Danish pension wealth. Secondly, data on mortgage debt do not include information on the risk profile (for instance, information on fixed interest, variable interest and/or deferred amortization). For that reason additional reporting forms are being set up in cooperation with the reporters. Figure 1 shows how the additional reports from life insurance corporations, pension funds and specialized mortgage banks will link this information to the existing administrative registers via personal IDs.

Danmarks Nationalbank and Statistics Denmark are currently defining the new reports on pension wealth and mortgage loans. The legal authority to collect these detailed data is given by the act governing *Danmarks Nationalbank*. However, the reports will be collected by Statistics Denmark, after which *Danmarks Nationalbank* will assess quality by comparing with data from the MFI and ICPF statistics.

Comparing data from households at micro level to national accounts

The main goal of the micro data set is to improve knowledge about the balance sheets of Danish households for analytical purposes. However, as *Danmarks Nationalbank* already publishes several macro statistics containing information on household assets and liabilities, it is natural to assess quality by comparison the different data sources. Hence, harmonization of micro aggregates and macro statistics is of high importance.

Figure 2 compares household assets (A) and liabilities (B) in an aggregation of the preliminary micro-data with national accounts. The comparison is a work-in-progress, as control for definitional differences between the two data sets are currently being done. In general, Figure 2 shows that several items are reasonably consistent between the two, while others diverge significantly.

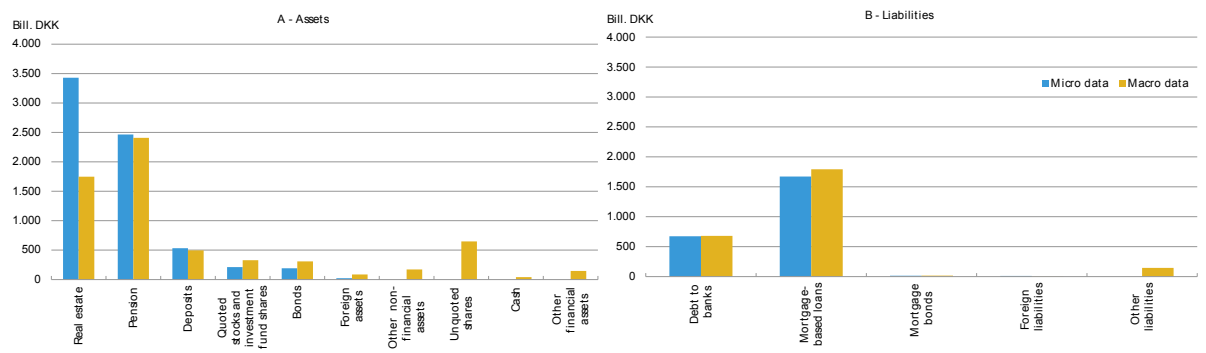
Three general issues emerge when comparing the data, and they are discussed below.

Valuation. The largest item on household balance sheets in the micro data is real estate wealth. However, real estate wealth in national accounts is only about half the size of micro data. This is because non-financial accounts use the Perpetual Inventory Method (PIM) to calculate real estate assets. PIM generally disregards price movements in the current real estate stock and the value of land. In the micro-data, estimated market values are calculated from the public valuation multiplied by the ratio of transaction price and public valuation of all real estate sold in the geographical area. Hence, the real estate assets from micro data currently better accounts for current market value than national accounts.

Comparison of preliminary micro-data and national accounts 2010

Part A compares the aggregated household assets in the first version of micro data with financial and non-financial accounts. Part B shows the same for liabilities.

Figure 2



Source: Quarterly financial accounts, annual non-financial accounts and own calculations.

Note: Pension wealth in micro data is estimated, since the individual information from future reports is not yet available. In this chart, macro data have been adjusted for known differences in classification between micro and macro data. This implies that the 380 billion DKK in pension funds administered by banks are moved from deposits to pensions. Also the 170 billion DKK in shares in bond funds are moved from quoted shares and investment fund shares to bonds.

Population. The micro data set on households includes all people living in Denmark aggregated at the family level while the household sector in national accounts also includes Non-profit institutions serving households (NPISH). The bigger population in national accounts is the most likely explanation for deposits in and debt to banks and mortgage debt being a little larger here compared to micro data. However, when national accounts implement the System of National Accounts 2008 in 2014, NPISH is excluded from the household sector. This will increase consistency between macro and micro data.

Coverage. Both on the asset and liability side of household balance sheets are items that are only covered by national accounts. Other non-financial assets (which mostly consist of estimated value of cars), unquoted shares, cash and other financial assets are not part of the micro data even though they add up to 15 per cent of assets in the national accounts. Most of these items would be difficult to include at the individual level, as data currently are not available. Other financial assets mainly consist of insurance reserves which by definition are difficult to disaggregate due to the basic risk sharing attribute of insurance.

The way ahead

Danmarks Nationalbank and Statistics Denmark are currently developing the reports from life insurance corporations, pension funds and specialised mortgage banks, but also other improvements to the data are in the works. Both market value of non-financial assets other than real estate (land, cars and boats) and current value of public servants' pay-as-you-go-based pension schemes are being added to the annual micro data set. It is expected that the new macro data can also help improving *Danmarks Nationalbank*'s existing statistics by comparison as done in this paper. Annual distributional information may in time make it possible to estimate national accounts disaggregated in terms of for instance income deciles or geographical location.

It is expected that the first version of the micro data will be published in 2014.

References

Andersen, A., et. al. (2012 I). "The Debt and Wealth of Danish Families", Monetary Review – Part 2, 2nd quarter 2012, pp. 1–40, Danmarks Nationalbank, 8 June 2012.

—— Andersen, A., et. al. (2012 II). "Danish Families' Financial Robustness, Variable Rates and Deferred Amortisation", Monetary Review – Part 2, 4th quarter 2012, pp. 1–53, Danmarks Nationalbank, December 7 2012.

ECB (2013), "The Eurosystem Household Finance and Consumption Survey – Results of the first wave", Statistics Paper Series, No. 2 April 2013, ECB.