

Standardised granular credit and credit risk data

Violetta Damia and Jean-Marc Israël¹

Abstract

The recent financial crisis has further highlighted that, although a wide range of data on credit are already available, more granular, frequent and flexible credit and credit risk data are considered of high relevance within the European System of Central Banks (ESCB) for monetary policy, financial stability and research analyses, as well as for the development and production of ESCB statistics. Such granular credit and credit risk data are also critical for micro-prudential supervisory purposes.

In this context, central credit registers, which are operated by several National Central Banks (NCBs) in the EU, appear as a major data channel. Such databases have proven to be very valuable sources of information by the financial industry itself for assessing creditworthiness of potential borrowers and benchmarking credit risks, and by supervisory authorities in assessing credit risk borne by credit institutions and other lenders. In many countries, (complementary) granular credit data may be available from private credit bureaus or via surveys.

Still differences in concepts, definition and coverage across countries are being revisited by the ESCB with a view to increasing their value for credit and credit risk analysis to support policy making, and for assessment of their risks by lenders.

This paper presents the current ESCB work towards harmonisation in concepts and definitions and convergence in data coverage and content so as to share a significant subset of granular credit and credit risk data to better support and serve many policy and analysis needs at European and national levels.

Keywords: European System of Central Banks (ESCB), credit and credit risk data, credit exposures, creditworthiness, indebtedness, Central Credit Registers, granular data, harmonisation, convergence

¹ European Central Bank.

1. Background

The recent financial crisis has highlighted that, although a wide range of data on credit are already available, more granular, frequent and flexible credit and credit risk data are considered of high relevance within the European System of Central Banks (ESCB) for monetary policy, financial stability and research analyses, as well as for the development and production of ESCB statistics. Such granular credit and credit risk data are also critical for micro-prudential supervisory purposes.

From the input side, central credit registers (CCRs) or similar granular (loan-by-loan or borrower-by-borrower) credit reporting systems (herein called granular-credit datasets) are considered as major data channels. CCRs are 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.

The ESCB has explored the potentials of granular-credit datasets, in particular to understand to which extent their content may be enhanced and adapted to euro area and EU statistical, supervisory and analytical needs, i.e. to meet the above mentioned user requirements while at the same time alleviating respondents' reporting burden and increasing transparency. In this context, several related ESCB initiatives have not only proven the analytical usefulness of such granular datasets 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

2. ESCB Task Force on Analytical Credit Datasets

On the basis of earlier work, an ESCB Task Force on Analytical Credit Datasets (hereinto mentioned as the TF), comprising experts from the statistical as well as from credit registers' areas, was mandated to investigate the following main issues:

² For more than 5 years, CCRs data (or granular data on credit in general) have been investigated by the ESCB in order to verify the possibility to (re-)use them for statistical purposes. Although there was a broad agreement on the fact that granular data may alleviate the reporting burden, at the same time it was widely acknowledged that a lot of work is still required to reach the necessary level of harmonisation. In May 2009, the Expert Group on Credit Histories set up by the European Commission, recommended for Credit Registers to converge, in particular as regards the concepts and definitions used (e.g., bad debt, arrears, default, loan types, etc.). It, however, recognised several legal and other obstacles in doing so. Based on some work organised by the ESCB in 209–2010, a task force elaborated in 2012 some recommendations for the identification and common definition of core attributes on loans. It also, ran a pilot exercise and, based on the feedback received from a representative sample of ECB internal users, concluded that harmonised granular credit and credit risk information is key for monetary and financial stability analyses. Since then, the ECB Banking Supervision (in the making) also strongly supported such an endeavour.

- (1) identify a core set of information to meet main users' needs in the long term and elaborate on its scope
- (2) to further analyse and consider harmonised concepts and definitions and methodological enhancements to core data, metadata and attributes;
- (3) estimate the costs to be incurred by the ESCB and the reporting agents; and
- (4) consider the governance, legal and confidentiality issues prevailing at national and EU levels³ and prepare the appropriate legal instrument.

To launch its work, the TF organised in February 2014 a workshop related to ESCB users' requirements on credit and credit risk data, where ESCB users were invited to present concrete business cases of using credit data in their areas that are already implemented and related data needs. The workshop confirmed the very high importance given to granular credit and credit risk data for a number of ESCB and ESRB tasks; the availability of a granular credit dataset would:

- better address a number of **monetary policy analysis** relevant issues relating to the provision of credit with a variety of counterparty breakdowns (size of firms, economic activity, undrawn credit lines, etc.) and the functioning of the transmission mechanism, especially in fragmented markets;
- play an important role in supporting the direct use of credit claims in **monetary policy operations** and in calibrating potential credit support measures to monitor bank lending and liquidity in the euro area money market;
- adequately calibrate the different **risk control and collateral management measures** of the Eurosystem, including adequate pricing, credit risk assessment and haircuts, and to allow an in-depth analysis of credit claims pledged with the Eurosystem credit operations;
- support **financial stability surveillance and macro-prudential analysis** as well as quantitative risk assessment, notably in the context of macro-stress testing; a key expected benefit, also for micro-prudential supervision, will be the assessment of creditworthiness of borrowers (via probabilities of defaulting) by credit institutions using internal-ratings based approach;⁴
- meet ever stronger and multiform **statistical and analytical needs** and breakdowns which require agility through granular datasets;
- serve **research purposes** for supporting credit risk analysis across euro area countries and various other financial research work, also assessing their impact on the non-financial economy; and, last but not least,
- enable a multitude of usage options in the **supervisory process** (off- and on-site, including the use in risk assessment systems) and permit analysis options

³ 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 law, and that accessibility to CCR raw data is often restricted for supervisory purposes.

⁴ So far these are scarcely available as only few large credit institutions use internal-ratings based approach in each Member State. In the AnaCredit shared dataset there will be more data on probabilities of defaulting, allowing several analyses and comparisons, including possible benchmarking.

otherwise not covered by regular reporting as well as complementing other reporting systems' information.

Following users' confirmation on the analytical value of such datasets, the TF launched its work aiming at establishing **a core set of analytical granular credit data (the so-called analytical credit dataset or AnaCredit)** to be shared, on a need-to-know basis, across the ESCB and other relevant institutions (e.g. national supervisory authorities associated in the Single Supervisory Mechanism, SSM, the European Banking association) and the EU Commission.

To achieve the long-term objectives, the TF provided input to the preparation of **Decision ECB/2014/6**, adopted on 24 February 2014,⁵ concerning the organisation of preparatory measures for the collection of micro credit data by the ESCB (hereinto the "Decision"). The Decision, which was established to provide the necessary legal basis supporting the preparatory work, sets out a list of measures to be implemented by all euro area countries concerning

- (a) the definition of data attributes and data transmission arrangements;
- (b) the elimination of data gaps due to non-existent or insufficient granular databases in some Member States; and
- (c) the regular monitoring of the progress achieved.

The preparatory measures also include the arrangement of annual collections of semi-annual data during the preparatory phase, for the purposes of advancing with the establishment of a harmonised granular credit data framework in line with the identified user needs.

3. Preliminary recommendations

The TF work resulted in a number of preliminary recommendations, referring to the gaps on credit and credit risk data, and including proposals to coordinate efforts to improve granular data among NCBs on lenders, borrowers and credit attributes, as well as methodological issues.

The recommendations cover the following areas:

1. Data Coverage

a. Definition of the set of granular information: as credit and credit risk data may focus on small and medium-size enterprises, as large enterprises may more easily find other funding sources, e.g. issuing securities, any threshold to minimise the reporting burden needs to be rather low. In general priority for sharing data will focus on credit to legal persons, e.g. financial and non-financial corporations, and government agencies. Loans to households, in particular housing (often mortgage) loans would also be subject to thresholds and anonymisation. All non-performing loans may need to be reported.

⁵ For more information please ref. http://www.ecb.europa.eu/ecb/legal/pdf/oj_jol_2014_104_r_0008_en_txt.pdf.

b. Lenders: The data collected should provide a clear insight into the borrowers' credit positions. This could be done with data from credit institutions and other financial intermediaries.

A Memorandum of Understanding among national CCRs⁶ shows that credit institutions are reporting to all CCRs, and may remain a core reporting population, in particular for the forthcoming Banking Supervision purposes. However, as credit intermediation has increased among other financial institutions granular credit data may also be collected from these.

c. Borrowers: All users, in particular the Banking Supervision, have assessed that all types of borrowers would be of interest but that priority should initially be given to non-financial corporations and general government.

As regards the borrower identification, the use of a unique identification code (for individual undertakings, all the more as components of groups) would be ideal. The Legal Entity Identifier may, in the future, become the standard. In the mean time, the codification and identification used in national business registers and the ESCB "Register of Institutions and Affiliates Database" (RIAD),⁷ which already includes reference data for credit institutions and other lenders, will be used to share existing codes, and map them with the LEI, where available. It is envisaged that access to lists of institutions may be granted to facilitate the use of common identifiers with respondents.

Registers also provide also key information for the identification of lenders and borrowers, for their stratification (e.g. based on turnover or other variables) and for defining whether stand-alone entities or parts of identified groups. They need timely updates for which support of the industry may be relevant.

2. Types of credit

According to the users' assessment priority should be given to loans of all types, including securitised loans derecognised from the balance sheet, in particular when they are still serviced by the originating lender. Loan provisions and credit derivatives would also be relevant. The distinction is made according to statistical manuals where tradability determines the difference between loans and securities.

From the supervisory perspective ideally all financial assets of a reporting institution, including risk-based variables, would be within the scope of a granular credit database.

As regards issuance and holdings of securities, interoperability will be sought with other relevant data sources of debt or credit exposures, e.g. the ESCB securities

⁶ For more information please ref. <https://www.ecb.europa.eu/pub/pdf/other/memoxinccreditregisters201004en.pdf>.

⁷ For more information please ref. <http://www.ecb.europa.eu/stats/money/mfi/html/index.en.html>.

database on issues (the Centralised Securities Database)⁸ and holdings (the Securities Holdings Statistics Database).⁹

3. Level of granularity

Concerning the level of granularity, i.e. whether the reporting should be performed on a loan-by-loan (l-by-l) or borrower-by-borrower (b-by-b) basis, the level of detail can be collected using these systems may comprise (a) a collection on a l-by-l basis, (b) a multi-dimensional b-by-b system (containing e.g. "outstanding amount" of "trade receivables", "securitised", "collateralised", "derecognised" and with "maturity lower than 1 year"), or (c) less detail obtained in a simpler b-by-b model.

Mixed solutions also exist. It is recognised that the optimal reporting model for granular information, in the long term, is loan-by-loan, although the shared AnaCredit dataset will be built up to receive data on a l-by-l and a (mixed) multi-dimensional b-by-b basis to allow for a stepwise alignment over time.

4. Stepwise implementation

Last but not least, given the complexity and costs of this endeavour, the TF is of the opinion that AnaCredit should be developed in a stepwise manner, to accommodate the difficulties and different states of both the existing and non-existing granular reporting systems across Europe, taking at the same time due account of the users' priorities.

4. Challenges ahead

In view of identifying 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 are currently pursued by the ESCB:

On-going merits and costs procedure

As for all new or substantially enhanced statistics, the ESCB follows a "merits and costs procedure" so as to design a cost-effective approach to best fulfil user needs, while minimising the reporting burden. Such a procedure, rather similar to the "impact assessments" run by the European Commission is required to support the decision-making process towards establishing the long-term framework

In May 2014, the TF launched the cost assessment exercise covering (a) a number of attributes/features to be included in the reporting framework; (b) implementation issues regarding the provision of information on specific sets of

⁸ For more information please ref. <https://www.ecb.europa.eu/pub/pdf/other/centralisedsecuritiesdatabase201002en.pdf>.

⁹ For more information please ref. Regulation ECB/2012/24 http://www.ecb.europa.eu/ecb/legal/pdf/l_30520121101en00060024.pdf and Guideline ECB/2013/7 http://www.ecb.europa.eu/ecb/legal/pdf/l_12520130507en00170033.pdf.

exposures on a granular and aggregate basis; (c) the provision of information of group structures, and (d) the integration of information with existing datasets such as registers or securities databases.

Overall, the cost assessment showed high implementation costs (IT systems and work load) for both NCBs and reporting agents. However, the situation widely differs, e.g. between those countries where central credit registers (CCRs) are in place and those without any granular credit reporting, also translating into different cost assessments across countries. In addition, some features were reported as having a significant impact on the overall cost assessment, namely the additional reporting of credit exposures on a consolidated basis (further to solo reporting), the possible inclusion of foreign subsidiaries in the reporting, the timeliness of the reporting schemes, or the compilation of detailed multi-dimensional aggregates.

The questionnaire also included questions on the foreseen merits of AnaCredit to reporting agents. Overall, the reporting agents recognised that (1) the set-up of a harmonised granular euro area-wide (or even EU-wide) credit database may further enhance the credit institutions' assessment on the creditworthiness, in particular of cross-border borrowers, (2) such granular information may lead to reducing the reporting burden if the envisaged level of detail minimises significantly the aggregation which has to be done by them, and, if appropriately defined, (3) the reporting requirements are expected to be more stable over time (as already experienced for security-by-security reporting) which is an important factor minimising costs in highly automated systems.

Following the detailed cost assessment, the TF is about to launch a users' consultation, i.e. ESCB users, the EU Commission, the European Banking Authority and the European Systemic Risk Board, will be asked to review their business cases, in the light of costs potentially incurred, and, in particular, to clearly prioritise their needs for a stepwise implementation.

The requirements for the dataset will eventually be transposed into a legal act to be submitted (likely in 2015 Q2) to the ECB Governing Council for approval.

Credit exposures at group level

The reporting is foreseen to be established on a solo basis. However, as institutions may be part of more complex group structures, the TF investigated an appropriate approach to provide information on entities at group level, important from a financial stability and supervisory viewpoint.

Two main approaches were considered to allow for analysis of credit exposures at group level: the direct reporting on a group basis (e.g. reporting of exposures at the consolidated banking group level), or additional data on group structures. The former approach is considered more costly but provides more accurate information on exposures at group level. The latter method only uses the information reported on a solo basis and, therefore, is less costly for reporting agents. However, it requires retrieving the relevant information on group structures from other (supervisory) reports or from business registers. Although more demanding for compiling agencies, it provides flexibility to reflect changes in group structures as long as the relevant information on entities and the group is available and up-to-date.

As banking groups are not confined to the countries of parent institutions, it is important to obtain the exposures of foreign subsidiaries credit institutions. Exposures from subsidiaries resident outside the euro area would be collected through the resident parent companies –following the home approach.

Finally, the TF assumed that the RIAD system could serve as the database to store information on group structures as it already contains some criteria used to identify entities in the group. Still, should it be the case, the information in RIAD needs to be enhanced with additional data.

The TF will continue work in this area.

Identification of lenders and borrowers

The TF is working on the establishment of a procedure to (i) uniquely identify legal persons from different data sources (e.g. business registers including RIAD, AnaCredit or securities databases) and (ii) integrate the information from different sources into a unique record, under a shared responsibility within the ESCB.

The identification of lenders and borrowers is considered a key issue for the establishment of AnaCredit to ensure a consistent treatment, an accurate reporting and a correct exchange of information with other datasets. To facilitate the exchange of information without forcing a change of codes at national level (also as the Legal Entity Identifier may soon be available for lenders, but may take much longer for borrowers), RIAD is envisaged to serve as a hub where the different identifiers as aliases for a single entity or group. In RIAD, flags could indicate which identifier/alias is used in which dataset so as to allow the necessary reconciliation. For non-EU borrowers, unique identification, so far as possible via the Legal Entity Identifier, will be implemented for all lenders (the inclusion of foreign lenders in the reporting population is still under discussion, although likely at least for branches and subsidiaries headquartered in the EU) and large borrowers.

However, there are currently legal provisions in some countries hindering the connection at national level with the business register. The TF will further investigate the establishment of best practices/alternative solutions.

The RIAD system should be fed overall on a host-country principle, i.e. the country where the exposure was originated and where the (EU) borrower is resident will be the criterion for defining the NCB responsible for the data checking. However, the home-country approach would be used for foreign branches and subsidiaries outside the euro area/EU (e.g. the exposures of a Deutsche Bank in US would be reported by the German parent institution). As regards non-resident borrowers, the TF will further discuss how to ensure data quality management, e.g. some NCBs may express interest for some regions in the world, the ECB acting as a hub and ensuring coordination for borderline cases.

Overview of credit exposures and indebtedness

In order to provide users with an overall picture of the credit exposures of credit institutions or other lenders to borrowers (be it on a solo or on a consolidated basis), granular information on loans and derivatives (to be collected in AnaCredit) should be complemented by data on debt securities held by lenders. Information on

issuance and holdings of debt securities is already available in the CSDB and the SHSDB, respectively.

Acknowledging the need of users (including the SSM) for an assessment of the overall exposure of lenders, i.e. beyond loans and credit derivatives, an efficient approach is to combine the AnaCredit dataset with the two securities datasets mentioned above (CSDB and SHSDB).

In both databases there are ways to uniquely identify lenders and borrowers and possibly connecting them to AnaCredit. However, on the methodological side the CSDB and SHSDB deviate from the foreseen content (e.g. level of granularity and consolidation) of AnaCredit. A light fact-finding exercise conducted by the TF has shown that either AnaCredit should comprise all types of exposures (including securities), in particular for non-financial corporations as borrowers/issuers, or the ESCB-wide shared databases on securities would need to allow for the necessary level of granularity. In the latter case, the appropriate connection between information on debt securities and AnaCredit, to be mediated via RIAD (for mapping identifiers and aggregating for group structures where appropriate), would then provide the required overview of exposures/indebtedness. This second approach may be more accurate, although costly, and require a longer implementation phase, as the level of granularity in the SHSDB would need to increase (for the sectoral data, and possibly also for the disaggregation of reporting banking groups). A borderline case relates to those instruments issued as private placements, which in principle are securities, although they are often issued without ISIN codes. As indicated by the fact-finding exercise, they may be significant in several countries and may be significant when considering credit to the non-financial economy.

The TF will investigate more in depth possible options for combining the SHSDB and CSDB with AnaCredit and ways to overcome possible methodological and practical mismatches between these datasets.

A proper IT application

The work may be further supported by the establishment of an IT solution for receiving, storing and disseminating credit and credit risk information on a euro area (or an EU-wide) scale, which would be sourced from CCRs or other similar granular datasets and would include the most important attributes on loans, lenders (likely abridged from the 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). It needs to help protect confidentiality and handle different levels of access to the datasets (aggregated, granular anonymised, non-anonymised). The remainder of the granular information would be handled at national level, under the responsibility of NCBs who 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 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 to create interconnected statistical databases to be used according to the increasing demand for data.

Accronyms

AnaCredit:	Analytical Credit Datasets
CCR:	Central Credit Register
CSDB:	Centralised Securities Database
EBA:	European Banking Authority
ECB:	European Central Bank
ESCB:	European System of Central Banks
ESRB:	European Systemic Risk Board
NCB:	National Central Bank
RIAD:	Register of Institutions and Affiliates Database
SHSDB:	Securities Holdings Statistics Database
SSM:	Single Supervisory Mechanism

References

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

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

"Report of the Expert Group on Credit Histories.pdf": http://ec.europa.eu/internal_market/consultations/docs/2009/credit_histories/egch_report_en.pdf.

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

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

De Almeida, Ana M. and Damia, V. (2013), Challenges and prospects for setting-up an EU shared system on credit, Workshop on integrated management of micro-databases, Banco de Portugal: http://www.bportugal.pt/en-US/Estatisticas/Conferencias/Documents/01%20Almeida_D1_M_A.pdf.

Sola, P. and F. Strobbe (2010), Addressing data gaps revealed by the financial crisis: European Central Bank statistics on holdings of securities, Irving Fisher Committee on Central-Bank Statistics, IFC Bulletin No. 34: Proceedings of the IFC Conference on "Initiatives to address data gaps revealed by the financial crisis", Bank for International Settlements, Basel, August 2010.