

Israël, Jean-Marc; Damia, Violetta; Bonci, Riccardo; Gibran, Watfe,

Research Report

The analytical credit dataset: A magnifying glass for analysing credit in the euro area

ECB Occasional Paper, No. 187

Provided in Cooperation with:
European Central Bank (ECB)

Suggested Citation: Israël, Jean-Marc; Damia, Violetta; Bonci, Riccardo; Gibran, Watfe, (2017) : The analytical credit dataset: A magnifying glass for analysing credit in the euro area, ECB Occasional Paper, No. 187, ISBN 978-92-899-2849-6, European Central Bank (ECB), Frankfurt a. M., <http://dx.doi.org/10.2866/813758>

This Version is available at:
<http://hdl.handle.net/10419/162204>

Standard-Nutzungsbedingungen:

Die Dokumente auf EconStor dürfen zu eigenen wissenschaftlichen Zwecken und zum Privatgebrauch gespeichert und kopiert werden.

Sie dürfen die Dokumente nicht für öffentliche oder kommerzielle Zwecke vervielfältigen, öffentlich ausstellen, öffentlich zugänglich machen, vertreiben oder anderweitig nutzen.

Sofern die Verfasser die Dokumente unter Open-Content-Lizenzen (insbesondere CC-Lizenzen) zur Verfügung gestellt haben sollten, gelten abweichend von diesen Nutzungsbedingungen die in der dort genannten Lizenz gewährten Nutzungsrechte.

Terms of use:

Documents in EconStor may be saved and copied for your personal and scholarly purposes.

You are not to copy documents for public or commercial purposes, to exhibit the documents publicly, to make them publicly available on the internet, or to distribute or otherwise use the documents in public.

If the documents have been made available under an Open Content Licence (especially Creative Commons Licences), you may exercise further usage rights as specified in the indicated licence.



EUROPEAN CENTRAL BANK

EUROSYSTEM

Occasional Paper Series

Jean-Marc Israël, Violetta Damia,
Riccardo Bonci, Gibran Watfe

The Analytical Credit Dataset

A magnifying glass for analysing credit
in the euro area

No 187 / April 2017

Contents

Abstract	2
Executive summary	3
1 Introduction	
The concept of AnaCredit	6
2 The development of AnaCredit	10
Box 1 How data from Central Credit Registers can support monetary and macroprudential policy	14
2.1 Centralised v decentralised reporting	16
2.2 Loan-by-loan v borrower-by-borrower	19
2.3 Coverage of instruments, lenders and borrowers	22
2.4 Reporting threshold	26
2.5 Individual v consolidated reporting	28
2.6 Derogations	30
2.7 Counterparty reference data	32
2.8 Interoperability with existing granular datasets	34
3 The way forward	37
References	41
Acknowledgements	42

Abstract

In May 2016 the Governing Council adopted the AnaCredit Regulation (ECB/2016/13) providing the legal basis for the European System of Central Banks (ESCB) to collect granular information on loans from banks to corporates and other legal persons based on a core set of harmonised concepts and definitions. Starting with reference data from September 2018, credit institutions in the euro area, and possibly elsewhere in the EU, will report to the ECB via the national central banks (NCBs) individual credit exposures falling within the reporting scope. The reporting framework is the outcome of in-depth discussions within the ESCB involving several rounds of consultations with users, the industry and other stakeholders. As set out in the Regulation, AnaCredit will, already in Stage 1, significantly enhance the value for analysis on credit and credit risk in the euro area by providing detailed, timely and harmonised information on individual exposures to legal entities as counterparts. The new data will be useful for several key tasks of the ESCB for a better analysis of credit distribution to the economy, e.g. for monetary policy analysis and operation (risk and collateral management), financial stability, economic research and statistics. The scope of the project might be further expanded in future stages to cover additional lenders, borrowers and instruments. The purpose of this paper is to reflect and illustrate the methodological work and process leading to the definition of the AnaCredit requirements that were eventually included in the Regulation.

Keywords: analytical credit dataset, central credit registers, loan-by-loan data, credit risk, central bank statistics.

JEL codes: E58, G21, E51, C81, E44.

Executive summary

AnaCredit is a shared multipurpose database containing loan-by-loan information on credit extended by credit institutions to companies and other legal entities. On 18 May 2016 the Governing Council of the ECB adopted Regulation ECB/2016/13 on the collection of granular credit and credit risk data (AnaCredit) establishing Stage 1 of a shared database for the European System of Central Banks (ESCB) as of September 2018. The database will contain 88 attributes, updated mostly on a monthly basis, based on harmonised concepts and definitions common to all participating countries.

The AnaCredit Regulation is the outcome of in-depth discussions within the ESCB, involving several rounds of consultations with users, the industry and other stakeholders. The purpose of this paper is to reflect and illustrate the seminal methodological work and foreseen steps as elaborated by the Joint Statistics Committee (STC)/Financial Stability Committee (FSC) Task Force on Analytical Credit Datasets (JTF). This work eventually led to the definition of the AnaCredit reporting requirements as laid down in the Regulation.

The intensive collaboration among the relevant stakeholders showed first and foremost the usefulness of AnaCredit data for central banking purposes. The potential of a granular dataset like AnaCredit appeared very high from the very first contacts with users. Detailed information on credit and credit risk was deemed useful for several key tasks of the ESCB, including monetary policy analysis and operation (risk and collateral management), financial stability, economic research and statistics. In fact, user consultations showed that there were more than one hundred business cases for AnaCredit data in the realm of the ESCB's tasks. In a nutshell, AnaCredit is considered to offer a magnifying glass to analyse credit and credit risk in the euro area.

AnaCredit will also benefit reporting agents by enhancing their ability to assess borrowers' creditworthiness and through simplified reporting. The benefits of AnaCredit will not be limited to users within the ESCB. In the age of financial globalisation, credit flows and risk taking do not stop at national frontiers. It is therefore ever more important for credit institutions to assess the total exposures of their borrowers across borders. Used in conjunction with the Register of Institutions and Affiliates Database (RIAD), AnaCredit will allow for the unique identification of all counterparties (i.e. lenders and borrowers) and will offer a high degree of harmonisation of concepts and definitions, therefore allowing for a meaningful calculation of the total indebtedness of a borrower (company) vis-à-vis all its lenders (credit institutions). Reporting agents, as users of the granular credit information e.g. via feedback loops or simply via a better internal flow of information triggered by AnaCredit, will be able to perform a finer and more robust analysis of their own exposures than is currently feasible. Moreover, and as importantly, over time the reporting of such granular information is expected to mitigate the reporting

burden via more stable requirements and less ad hoc requests, thanks to the high flexibility of the new dataset.

The final reporting framework represents a fair balance between the numerous and relevant needs expressed by users in different business areas of central banking and the associated costs for setting up, collecting and processing the necessary information. The most relevant issues discussed during the definition of the AnaCredit requirements were (i) a centralised versus a decentralised (i.e. via NCBs) reporting; (ii) a loan-by-loan versus a borrower-by-borrower reporting; (iii) the appropriate coverage of instruments, lenders and borrowers; (iv) the reporting threshold (absolute versus relative, common versus country-specific); (v) individual versus consolidated reporting; (vi) scope for possible derogations to small reporting agents; (vii) definition of the relevant counterparty reference data; (viii) interoperability with other granular datasets (e.g. on securities holdings and issuance). Each feature of the final AnaCredit design as laid down in the Regulation is the result of a detailed merits and costs procedure, which involved the users, the NCBs and the financial industry. Moreover, the JTF based its work on best practices from national central credit registers, benefitting from the long-lasting experience available in some central banks that operate such systems. In conjunction with a continuous involvement of all stakeholders, a fair cost-benefit balance was thus achieved. In order to ensure the proportionality of the reporting obligations and mitigate the burden on smaller institutions, NCBs were also eventually left the discretion to grant derogations to small reporting agents, within certain common limits and taking into account the country-specific structure and concentration of the banking industry.

In order to gradually improve the informative value of the dataset over time, AnaCredit is envisaged to further extend its scope with respect to types of lenders, borrowers and instruments in the future. Extensions could eventually include the enlargement of the lender and the borrower population, as well as the scope of the instruments covered. Further enhancements could be introduced with a view to increase the usefulness of AnaCredit for (macro- and micro-) supervisory purposes. Moreover, ongoing efforts to integrate AnaCredit with existing granular databases will continue, possibly also resulting in adjustments to the requirements in the future. In any case, any such extension would be subject to a dedicated merits and costs procedure and would need to be announced at least two years prior to its implementation, to ensure sufficient lead time.

AnaCredit is an important building block in the 'data strategy' of the ECB. The ECB statistical function is working, also with the involvement of the financial industry, on designing and implementing a coordinated data management comprising information collected under different statistical and legal frameworks, with a view to simplify data reporting and reduce the burden for reporting agents. The European Reporting Framework (ERF) aims at collecting all data required for different statistical purposes and (in a second step) for banking supervision under an integrated and harmonised approach in all countries. The Banks' Integrated Reporting Dictionary (BIRD), developed with the voluntary participation of a high number of commercial banks, describes a possible model for the input information

banks need to report to authorities, as well as the transformations necessary to generate these reports. Both these initiatives are aimed at providing financial institutions with an integrated perspective that would reduce their reporting burden and increase the data quality over time. AnaCredit provides a unique opportunity to put these initiatives in practice, and in fact the first release of the BIRD is fully compliant with AnaCredit.

1 Introduction

The concept of AnaCredit

“ESCB statistics will continue to provide the ‘big picture’ of economic developments. But we should also offer a magnifying glass. Looking at the details beyond the aggregates enriches our understanding of economic phenomena and at the same time increases our flexibility to respond to unexpected policy needs, contributing to even better statistics.” (M. Draghi, 2016)

On 18 May 2016 the Governing Council of the ECB adopted Regulation ECB/2016/13 on the collection of granular credit and credit risk data (AnaCredit) establishing Stage 1 of a shared database for the European System of Central Banks (ESCB) as of September 2018. The creation of AnaCredit marks a new era for central banking statistics and is a genuine paradigm shift triggered by the need to “move beyond the aggregates”, to quote the title of the 8th Statistics Conference organised by the ECB in July 2016.

Content

AnaCredit is a shared multipurpose database that will contain loan-by-loan information on credit to companies and other legal entities extended by credit institutions and their foreign branches on a monthly basis. Based on compelling requests from users in a large number of central banks’ business areas, AnaCredit data collection has been designed with a view to obtaining a complete picture of a) the total credit exposure of the reporting population and b) the total indebtedness of borrowers across all lenders. The information collected consists of 88 different attributes based on harmonised concepts and definitions and covers various aspects of the credit exposure. The dataset is organised in several tables based on three distinctive elements: instruments, counterparties and protection received. The attributes were selected based on extensive user consultations and a merits and costs procedure, whereby the most costly attributes were identified and eliminated. Section 2 describes the merits and costs procedure and looks at the discussions that led to the final design of AnaCredit.

Reporting requirements

The AnaCredit Regulation applies to reporting agents resident in the euro area. Credit institutions in other EU Member States may participate in AnaCredit provided that the reporting requirements are adequately transposed into their national law. Where applicable, reporting agents have to report credit and credit risk data, including relevant counterparty reference information. Credit refers to any type of

transaction which gives rise to a credit risk exposure to the observed agent.¹ Instruments to be collected in Stage 1 of AnaCredit include outstanding financing under any of the following types of credit:

- deposits other than reverse repurchase agreements
- overdrafts
- credit card debt
- revolving credit other than overdrafts and credit card debt
- credit lines other than revolving credit
- reverse repurchase agreements
- trade receivables
- financial leases
- other loans.

Generally, credit data related to the above-mentioned instruments are to be reported if the credit risk arising under the instruments lies with the observed agent. The scope of data collection might be extended to other types of instruments in future stages of AnaCredit. Consequently, only credit extended to legal entities, as defined in Article 1(5) of the Regulation, falls under the scope of the data collection. The threshold for an instrument to be subject to reporting is a total commitment amount of €25 000 at borrower level at any point within the reference period. This includes both the drawn and the undrawn (i.e. off-balance- sheet) amounts of the instrument.

In order to ensure the proportionality of the reporting obligations, the relevant NCB may grant derogations to small reporting agents while respecting certain limits at national level.² Derogations may be granted to small reporting agents provided that the combined contribution of all reporting agents that are granted derogations to the total outstanding amount of loans reported by all reporting agents resident in the reporting Member State does not exceed 2%. Moreover, derogations allowing for a temporarily reduced reporting frequency (from monthly to quarterly) may be granted to small reporting agents if their combined share in the national total outstanding amount of loans does not exceed 4%. Derogations can be granted in full or in part (at attribute level).

¹ Reporting agents in AnaCredit are credit institutions and their foreign branches resident in a reporting Member State. An observed agent is an institutional unit whose activity as creditor or servicer is reported by the reporting agent. All observed agents are jointly referred to as the reference population. An observed agent is always related to a reporting agent. Depending on the reporting agent itself, there may be just one or several observed agent(s) affiliated with the reporting agent. The rationale for distinguishing between reporting agents and observed agents is that AnaCredit is designed to collect credit data with a view to obtaining a complete picture of the credit exposures of reporting agents in the reporting Member States, taking into account the area of economic activity of the reporting agent on the one hand, and its credit exposure in a specific country on the other. For more information, please refer to the AnaCredit Manual available on the ECB website.

² The relevant NCB is the NCB to which the reporting agent has to report.

According to the Regulation, data collected for AnaCredit can be part of a broader national reporting framework in which more extensive data is collected.

Stepwise implementation

The go-live of Stage 1 of AnaCredit is scheduled for 1 September 2018 whereas the first data transmission will be in mid-November 2018 with 30 September 2018 as reference date.

While the Regulation focuses on Stage 1 of AnaCredit, future stages are foreseen. The Regulation stipulates that the Governing Council will take its decision on any future widening of AnaCredit at least two years prior to its implementation to ensure sufficient lead time for reporting agents and NCBs. Future stages could include inter alia the following extensions:

- a reassessment of national discretion with respect to granting derogations for small reporting agents (Recital 11)
- an extension of the reporting population to non-deposit-taking institutions and other financial corporations in later stages to deposit-taking corporations other than credit institutions, asset management vehicles and other financial corporations (Recital 12)
- an extension of the instruments to be reported to derivatives, other accounts receivable, off-balance-sheet items (such as financial guarantees) and credit extended to persons other than legal persons, including to sole proprietors (Recital 12)
- the requirement to report on a consolidated basis (Recital 12).

Future stages will be prepared by conducting additional merits and costs exercise(s) and taking into account market developments and experience gained in Stage 1. The procedure is expected to apply best practices of transparency.

Guideline and Manual

Whereas the AnaCredit Regulation is addressed to reporting agents and therefore contains binding rules for primary reporting, an ECB Guideline is necessary to lay out the provisions governing how NCBs report the required credit data to the ECB in the phase of secondary reporting. In particular, the Guideline will include additional provisions regarding the secondary reporting, such as confidentiality requirements and data quality management elements as well as a framework governing the submission of counterparty reference data. In the near future, it will also include provisions regarding the feedback loops to reporting agents.

The legal acts are complemented by a detailed Manual mainly addressed to reporting agents and NCBs describing in detail the AnaCredit model and its reporting

features. With a view to creating a common database, the Manual promotes the alignment of concepts, definitions and reporting practices across countries to ensure a sound methodological background and accurate and comparable information. In this context, it explains the methodology underpinning the data collection and the data model, and it also includes guidance on the preferred approach in case the respective ECB legal act gives room to different interpretations. The Manual contains no additional requirements and has no binding legal status, as it merely aims to clarify and provide examples related to the requirements and definitions previously laid down in the AnaCredit Regulation.

2 The development of AnaCredit

The rationale for developing an analytical credit dataset

Credit registers alleviate the information asymmetry typically characterising credit markets. As such, they play an essential role for an efficient allocation of credit and for monitoring credit risk. Eventually, they are crucial for a sound functioning of the financial system. By providing a common platform for lenders to share information about actual and potential borrowers, either voluntarily or due to regulatory obligations, credit registers can contribute to a lower interest rate premium on loans by alleviating moral hazard. From an economic perspective, credit registers are recognised to have several advantages. First, as fewer borrowers tend to be priced out of the market, the presence of credit registers expands the credit available in the economy (Pagano and Jappelli, 1993). Second, information sharing provides an incentive for borrowers to prevent default, thus mitigating the risk of moral hazard (Padilla and Pagano, 2000). Third, credit registers counteract information monopolies as they provide all participating lenders with credit histories of borrowers (Padilla and Pagano, 1997). Finally, credit registers potentially reduce over-indebtedness, by revealing aggregate indebtedness per borrower across all lenders (Bennardo et al., 2014).

The nature of credit registers

Credit registers are usually categorised according to ownership and the purpose of the data collection. Central Credit Registers (CCRs) are usually operated by national central banks. They commonly serve the purpose of banking supervision and, occasionally, provide data for research. Private credit bureaux are usually run by private companies and serve as a platform for participating lenders to share information on borrowers' creditworthiness.

The landscape of credit registers in the EU

In the EU, the landscape of credit registers is rather heterogeneous. In 2009, the Expert Group on Credit Histories, mandated by the European Commission, collected information on public and private credit registers in EU Member States. They found that 15 Member States run CCRs whereas 24 Member States have one or more private registers.³ There is only one Member State which has neither a public nor a private credit register.

The cross-border dimension

Financial globalisation in general and a growing internationalisation of lending to companies, in particular, are trends that also have implications for the sharing of credit information. In the EU, the free flow of financial services strengthened by monetary union triggered an initiative for information sharing between various CCRs. In 2003, the CCRs of Belgium, Germany, Spain, France, Italy, Austria and Portugal signed a Memorandum of Understanding (MoU) containing principles for exchanging information. The aim was (and still is) to provide users with a way to obtain information on cross-border exposures of lenders and aggregate indebtedness of borrowers when lending in different Member States. The ESCB Working Group on

³ Croatia joined the EU after the report was published. It has been included in the numbers given here.

Credit Registers provided a forum for fine-tuning the MoU. The MoU was thereafter amended when the CCRs of the Czech Republic and Romania joined the endeavour.

The collapse of Lehman Brothers in September 2008 triggered a global financial crisis and revealed an urgent need for more granular, interconnected and flexible statistics. In fact, data gaps made it almost impossible for financial regulators and market participants to assess the scale of exposures to Lehman and affiliated firms. The collapse of financial markets was basically the result of uncertainty and mistrust due to insufficient knowledge about the real state of counterparties. This was the motivation for the G20 Data Gaps Initiative, launched in 2009 to improve the availability and comparability of economic and financial data. In this context, the Issing Committee⁴ identified the creation of "a global credit register" as one of the necessary elements for a more resilient international financial architecture. In particular, referring to the data from a credit register, the Issing Committee (2009) observed that "while the value of such information is appreciated almost universally on a national level, there is nothing commensurate on an international level", hence suggesting that "given the current high level of international lending and exposures, a global credit register will greatly enhance risk management, both at the firm level (improving due diligence of cross-border exposures), and at the systemic level (adding a cross-border dimension to financial stability stress testing, and to an evaluation of real effects on the economy)." In concrete terms, they proposed that "a harmonised approach should be adopted, where harmonisation refers to the standardisation across countries."

In the case of credit data, the information coming from CCRs is neither complete nor comparable, as a workshop organised by the ESCB in 2009 showed. It was acknowledged that while data from CCRs are already very useful to analyse risks at individual country level, at the same time it has severe shortcomings when it comes to analysing cross-country or supranational developments. First, there is no internationally used common identifier for borrowers. Second, the reporting thresholds across credit registers in the EU vary considerably.⁵ Finally, different concepts and definitions are used in each country. Against this background, the notion of AnaCredit took shape, aiming at developing a fully-fledged granular dataset on credit and credit risk based on a common set of definitions and covering all participating Member States.

The process leading to the definition of AnaCredit

The AnaCredit Regulation is the result of a long process. In fact, already in 2006, the Statistics Committee (STC) discussed the possible creation of a credit register or loan-by-loan database. It decided to perform an inventory of the situation regarding the existence, use, and content of existing national credit registers. In 2007, the Statistics Task Force of the ECB Governing Council identified the possibility of

⁴ Besides Otmar Issing, the Committee involved Jörg Asmussen, Jan Pieter Krahenen, Klaus Regling, Jens Weidmann and William White.

⁵ Cf. Table 2

enhanced reuse and sharing of micro data, in particular those available in existing CCRs. A workshop organised jointly by the Working Group on Monetary and Financial Statistics (WG MFS) and the Working Group on Credit Registers (WG CR) and hosted by Banco de Portugal in the summer of 2009 showed that many business areas could benefit from granular credit and credit risk data. Identifying increased demand for more granular data since the financial crisis, the participants also recognised the potential of granular data for reducing the reporting burden for reporting agents.

Overall, data collected by credit registers vary considerably across countries. With a view to achieving a higher degree of convergence of credit register data, the World Bank set up a Credit Reporting Standards Setting Task Force together with the Bank for International Settlements (BIS). In 2011, the Task Force's final report encouraged central banks to set up credit registers. Moreover, it defined general principles for setting up and running credit registers. The principles cover the areas of data, data processing, governance and risk management, the legal and regulatory environment as well as cross-border data flows. AnaCredit was developed in compliance with these principles and, more specifically, includes a provision on feedback loops, which was a main point raised in the Task Force's report.

The methodological work and the process leading to the definition of the AnaCredit requirements that were eventually laid down in the Regulation were mainly shaped by the Joint Statistics Committee (STC)/Financial Stability Committee (FSC) Task Force on Analytical Credit Datasets (JTF) – see the composition of the Committee in the “acknowledgements”. The JTF was preceded by the Task Force on Credit Registers (TF CR), which worked on the foundations of sharing credit register data at ESCB. The TF CR was mandated in November 2011 by the Working Group on Monetary and Financial Statistics (WG MFS) assisting the STC and the Working Group on Credit Registers (WG CR) assisting the FSC to explore three main issues. First, to identify the necessary data attributes and required level of harmonisation of definitions and methodologies to meet (central banking) user needs. Second, to explore governance, legal and confidentiality issues. Finally, to work on the required identification of entities and loans.

The outcome of the TF CR's pilot exercise

In addition, the TF CR undertook a pilot exercise with the aim of studying the feasibility of ensuring the provision of a minimum set of credit data based on ECB user needs. The exercise showed that, despite the level of heterogeneity across the different central credit registers, loan-by-loan or borrower-by-borrower data may extensively support several central banking purposes such as monitoring the transmission to the economy of monetary policy measures, analysing credit risk and its development and monitoring the degree of financial integration within the euro area. Besides, these data may become a primary source of information for the micro-supervisory tasks of the ECB. The final report of the TF CR listed numerous recommendations for the creation of AnaCredit based on the exercise and the discussions in the TF.

The JTF's mandate

On the basis of the work of the TF CR and its conclusions, the ESCB STC and FSC agreed to directly commission a new group in a broader composition to further

prepare and eventually establish the AnaCredit endeavour. Hence, in 2013, the JTF was mandated to pursue work on the following key issues:

1. Further elaboration on the scope of the AnaCredit dataset, based on user needs
2. Analysis and consideration of harmonised concepts and definitions and methodological enhancements to core data, metadata and attributes
3. Estimation of the impact of all the potential user needs on the (set-up and maintenance) costs to be incurred by the ESCB and the reporting agents for the collection and reporting of granular credit datasets
4. Estimation of the costs to be incurred by the ESCB (at margin and in full) in adapting their systems or creating new platforms to handle granular credit datasets
5. Consultation of the Legal Committee as regards the appropriate legal instrument as well as its potential content for running the longer-term approach.

The organisation of the JTF

The JTF pursued its work in a broader composition than the TF CR as it also included countries where no central credit register was previously available. It thus allowed for all European countries to participate in the development of AnaCredit. The JTF's work took the form of several physical meetings, teleconferences and numerous written procedures. One outcome of the JTF was the so-called "vision" for the development and future running of analytical credit data sets (the "vision paper"), which formed the basis for subsequent discussions.⁶

The rationale for developing AnaCredit based on user needs

To identify user needs, the JTF launched an extensive users' consultation across the ESCB, ESRB and other EU institutions and organised several follow-up workshops with users; this process led to the identification of more than one hundred business cases from various business areas of the ECB. The feedback received confirmed the very high importance given to granular credit and credit risk data for a number of ESCB tasks. For the purpose of monetary policy analysis, AnaCredit would make it possible to better address relevant issues relating to the provision of credit with a variety of counterparty breakdowns (size of firms, economic activity, new and undrawn credit lines, etc.) and the functioning of the transmission mechanism, especially in fragmented markets. It would also 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.

In terms of risk management, AnaCredit would increase the ability of the ECB to 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. AnaCredit would also support financial stability surveillance and macro-

⁶ The JTF was co-chaired – as the TF CR had been – by Ms Ana Margarida de Almeida (Banco de Portugal) and Mr Ramón Santillán (Banco de España) while a team of the ECB's Directorate General Statistics provided secretarial and methodological support.

prudential analysis as well as quantitative risk assessment, notably in the context of macro-stress testing. From a statistical point of view, AnaCredit would help to meet ever stronger and multiform statistical and analytical needs and breakdowns which require agility through granular datasets.

The database would furthermore serve research purposes for supporting credit risk analysis across euro area countries, monitoring changes in bank lending standards in particular for non-financial corporations. Research could use AnaCredit data also to extract observed failure rates across sectors by means of credit rating, by investigating interactions between monetary policy conduct and financial stability policy and assessing their impact on the non-financial economy.⁷ Moreover, assessing the evolution and trend of non-performing loans across the euro area could be scrutinised much more thoroughly with granular credit and credit risk data. Finally, AnaCredit has a multitude of usage options in the supervisory process (off- and on-site, including the use in risk assessment systems) and permits analysis options otherwise not covered by regular reporting as well as complementing other reporting systems' information.

Box 1

How data from Central Credit Registers can support monetary and macroprudential policy

In their paper titled “Hazardous Times for Monetary Policy: What Do Twenty-Three Million Bank Loans Say About the Effects of Monetary Policy on Credit Risk-Taking”, Jiménez et al. (2014) analyse whether low short-term interest rates induce higher (credit) risk taking by banks. They use a comprehensive dataset from the Spanish Credit Register (CIR) containing information on 23 million loans covering a period of 25 years. The dataset includes information on borrowers, lenders and the characteristics of each loan including default information. This allows the authors to identify borrowers' creditworthiness and to control for bank-specific characteristics.

The added value of loan-by-loan information

Previous research had been limited to conclusions derived from aggregate credit data. This allowed for analyses of total credit volumes and average interest rates potentially hiding many facets of credit risk taking by banks. In contrast, the loan-by-loan information contained in the CIR enabled the authors to use a magnifying glass and to exploit differences in loan characteristics.

The paper finds evidence that a prolonged period of low short-term interest rates induces banks to take on more credit risk. Moreover, the longer short-term interest rates stay low, the higher the probability of default of borrowers and the higher the credit risk banks are taking. Using a granular credit data set therefore enabled the authors to empirically examine a key question at the nexus of financial stability and monetary policy. The analysis was based on Spanish data as loan-by-loan information from credit registers is only available in certain countries and is hardly comparable. AnaCredit will enable researchers to make cross-country comparisons and enrich our understanding of key issues related to various ESCB tasks.

⁷ Aggregate data from AnaCredit is planned to be accessible to the public once data quality is ensured.

Merits are also expected for the reporting entities. The set-up of harmonised granular credit databases would further enhance the credit institutions' assessment on the creditworthiness of cross-border potential borrowers in particular, beyond what the current MoU offers, due to the use of a unique borrower identifier and a much higher degree of harmonisation of definitions and data attributes across granular credit datasets in many countries in Europe. In turn, reporting agents as users of the granular credit information will be able to perform a finer and more complex analysis than is currently feasible. The reporting of such granular information may, in turn, lead to the reduction of the reporting burden over time if the revisited dataset is more stable (i.e. less prone to new regular or ad hoc data requests, as already experienced for security-by-security reporting) and the envisaged level of detail significantly minimises the aggregation which thus far had to be done by reporting agents prior to transmitting the data.

When developing a new data collection framework with a potential impact on the reporting burden of respondents, the ESCB systematically assesses the merits for users, but also the costs for the ESCB and reporting agents via a merits and costs procedure. For AnaCredit, the procedure was run between April and October 2014, as prescribed by the legal basis for the AnaCredit Regulation, which is Council Regulation (EC) No 2533/98. The procedure included a wide-ranging assessment of costs for NCBs and reporting agents and a consultation of user committees and other stakeholders (including expected benefits for reporting agents). The merits and costs were then matched to fine-tune the requirements for the dataset and for the purpose of priority setting. In this process, many alternatives to the final outcome were considered and carefully assessed. In addition, a draft version of the Regulation was posted on the ECB web site in December 2015, and comments were received from several stakeholders. Some further amendments were made, culminating in the version eventually approved by the ECB Governing Council in May 2016.

After the TF CR had prepared the ground with recommendations on the way forward for developing an analytical credit dataset between 2011 and 2013, the JTF worked on solutions for the cornerstones of AnaCredit. It drafted the Regulation which was then finalised by its successor, the Working Group on AnaCredit (WG AnaCredit). The major achievement of the JTF was the harmonisation of concepts and definitions of up to 88 data attributes for the loan-level database across the euro area. In this context, the efforts of the JTF were inspired by the work of the Technical Working Groups that led to the creation of the European Data Warehouse, a comprehensive, centralised repository of data on European asset-backed securities. In particular, the template on small and medium-sized enterprises (SMEs) asset-backed securities, developed by the respective Technical Working Group, containing 81 mandatory and 96 optional attributes for loan-level data, proved to be a useful precedent.

The following subsections provide an account of the discussions in the JTF regarding the most pertinent issues in the development of AnaCredit.

2.1 Centralised v decentralised reporting

As for the actual collection of the AnaCredit data, the fundamental decision to be taken was whether reporting agents should report directly to the ECB (centralised approach) or rather to the respective NCBs (decentralised approach) for the latter to eventually pass the information onto the shared ESCB system. The Regulation stipulates that AnaCredit is a shared dataset collecting granular credit data based on harmonised ECB statistical reporting requirements. The dataset is shared between Eurosystem members comprising input data from all Member States whose currency is the euro. Reporting agents report to the respective NCB (primary reporting), which, in turn, report to the shared ESCB systems (secondary reporting).

Pros and cons

The centralised approach would be conceptually and technically simple and efficient insofar as only one central system would have to be developed instead of several systems at the level of the NCBs. At the same time, all relevant stakeholders (ESCB users and producers, and reporting agents) would directly access its information content. However, in practice, entirely centralised data collection would be a complex project with many intricate challenges. First of all, both the implementation and running of the fully integrated database would be very resource intensive. This applies to both comprehensive amendments of existing granular CCRs as well as to newly established data collection activities by NCBs. Secondly, a centralised system would not provide for NCBs to address national needs as well – their further integration could be assisted only by a process of convergence over time. Finally, it would be hard to establish a solid and sound legal basis for a centralised database that replaces national CCRs with all their different functions (e.g. feedback loops to reporting agents). In addition, other practical challenges such as data quality management (e.g. verifying the information reported by reporting agents) and, in particular, ensuring the unique identification of counterparties would be more significant under this option compared to the decentralised approach.

The second option, a shared ESCB database to be fed via participating NCBs, has benefits that alleviate most of the problems a centralised approach would pose. A shared system would be efficient insofar as it would allow NCBs to continue running their existing CCRs and leverage on their experience and well-established communication channels with respective reporting agents. The primary reporting would be the responsibility of the NCBs and would thus lessen the burden of the ECB also with respect to data quality management and possible provision of feedback loops. At the same time, the decentralised system requires continuous interaction between the central database at the ECB and the national datasets and an alignment of primary and secondary reporting. This would require detailed rules for secondary reporting from the NCBs to the ECB. Moreover, it would be a challenge for countries without a CCR to establish reporting at granular level to meet the ECB's requirements under the shared system.

Towards a shared database

Early on in the process defining what AnaCredit would look like, the JTF deemed a shared system (decentralised approach) the preferred option. In April 2013, it was emphasised that the intention was not to establish a pan-European credit register replacing the current functions of the existing national CCRs but instead to build a

distributional system whereby NCBs would choose their operational model – upgrading their CCRs where applicable, or defining other granular credit datasets; the shared dataset would contain a defined set of core harmonised data for satisfying different user needs. In this context, any direct contact and feedback loop with reporting agents would remain under the responsibility of NCBs, based on national arrangements. The AnaCredit dataset would support the further enrichment of such feedback loops with more extensive and harmonised information on cross-border loans and borrowers. In the course of the discussions, the benefits of guiding NCBs' efforts to provide, harmonise and extend feedback loops became clearer. Hence, the final Regulation contains an article on feedback loops that allows NCBs to use AnaCredit data for this purpose.

The STC, at its thematic meeting in September 2013, supported the view of the JTF by strongly supporting the second option, i.e. a decentralised approach. Option 1 was considered more complex and only suitable, if at all, for the very long term. Hence, a unanimous preference was expressed for the decentralised approach as a more realistic medium-term solution. However, it was thought that some elements of the centralised approach should be considered by the JTF in order to share best practices. This should include, for example, further harmonisation of data and information feedback to the industry. The former should be a particular focus of the JTF.

A third way: different layers of centralisation

Based on this mandate, the JTF agreed that a wide coverage should be achieved and that the concepts and definitions for data collection should be harmonised to a large extent. As this way forward would be similar to a centralised approach, the JTF sought other ways to limit, at least in the first stages, the perimeter of the database while allowing users to perform most of their analyses. Seeking a middle way between the practical merits of setting up a shared system and the desire of users to obtain harmonised data, the ECB team proposed a stratification of the dataset that would allow for different levels of centralisation according to the size of borrowers and instruments.⁸ In Layer 1, data of large borrowers would be collected and stored at individual level for both the shared database and the NCBs. Layer 2 would contain aggregated data with distribution measures in the shared dataset. While NCBs would collect data on a loan-by-loan or borrower-by-borrower basis from reporting agents, NCBs would provide aggregated data in this layer. The last layer would concern smaller borrowers, giving the NCBs discretion to decide whether they needed to collect the information in individual or aggregated format.

During the following months, a sub-group of the JTF worked on the details of the stratification proposal. During the JTF meeting in November 2013, views were exchanged on the criteria to be used for such stratification, for example, the size of a borrower and/or the size of a loan. Bearing in mind the cost implications, the JTF agreed that the scheme needed to be simple. The JTF also discussed the difficulties arising from the shift of borrowers among the different layers of information. Nevertheless, whatever the criterion (or criteria) used to define the stratification, two important elements were raised considering the future, shared microcredit database:

⁸ Cf. Table 1 for an illustration.

(1) the accessibility dimension (Layers 1 and 2 under consideration do not differ for NCBs, allowing national flexibility, while any changes to Layer 3 may prove very costly for NCBs) and (2) the stepwise implementation and prioritisation of the different strata to serve multidimensional purposes. It was finally agreed that a more detailed user consultation needed to follow.

Table 1

A proposal for stratifying the AnaCredit database according to the sector and the size of the borrowers.

	Borrowing sectors	Non-financial corporations (by size of the borrower)	General Government (by inst. sector & threshold)	Financial corporations (by inst. sector and threshold)	Households (by collateral type)	Other borrowers (by threshold)
Proposed layers						
Layer 1	<p>Individual in the collecting NCB</p> <p>Individual in the shared Analytical Credit Dataset</p>	Large enterprises	Central Gvt and other large GGvt borrowers	Large financial corporations	NA	for consideration
Layer 2	<p>Individual in the collecting NCB</p> <p>Aggregated in the shared Analytical Credit Dataset</p>	Medium enterprises	GGvt (above a given threshold)	MFIs and other financial corporations (above a given threshold)	NA	Above a given threshold
Layer 3	Aggregated in NCBs and in the shared Analytical Credit Dataset	Small and micro enterprises	Below given threshold	Below given threshold	Mortgage loans and loans for other purposes	Below a given threshold

Source: ECB.

Based on the outcome of a user workshop in February 2014, the sub-group recommended that Layer 2 be merged into Layer 1. The JTF agreed with this recommendation for the following reasons: (1) more granular information will be directly available to ESCB/SSM users, (2) a tool for the users to retrieve granular information from the NCBs will not be necessary, and (3) the existence of Layer 2 would be an additional burden for NCBs for the aggregation of granular credit data collected at national level. In turn, the integration of the two layers may also enhance transparency. However, stratification of the borrower population would more generally pose a challenge to users. If they needed to obtain granular data from Layer 2 they might have to liaise individually with many or all NCBs. Data from Layer 3 could severely lack comparability as NCBs would be free to decide on data collection.

The final outcome

Due to these difficulties with the stratification, the JTF, and subsequently the WG AnaCredit, prepared the ground to make AnaCredit a shared database with harmonised concepts and definitions without different strata. The resulting draft of the AnaCredit Regulation stipulated that AnaCredit would be decentralised insofar as reporting agents would have to report to the NCBs while the NCBs reported to the ECB in a second step. Having found a balance between harmonisation and decentralisation, this aspect of the AnaCredit database was not altered.

2.2 Loan-by-loan v borrower-by-borrower

AnaCredit is a dataset of high granularity. In 2013, the vision paper pointed to the need to decide on the level of granularity as one of the key features. Two ideas were considered: collecting borrower-by-borrower or loan-by-loan data.

Definitions

In its final report, the Task Force on Credit Registers (TF CR) defined a borrower-by-borrower dataset as a dataset where either the collection or the disclosure of credit data is based on the borrower itself and allows for the identification of individual borrowers or a reporting institution, but not on their individual loans with it. The model allows for the identification of instruments recorded only at a certain level of aggregation, which depends on the number and detail of the loan characteristics being collected. From the perspective of the reporting agent, aggregated data of the loans taken out by each borrower are to be reported, rather than all its individual exposures. In contrast to that, collecting data on a loan-by-loan basis means that the disclosure of credit data is based on the loan itself and allows for the individual identification of the lender, the borrower and their respective loans. In other words, each reporting agent reports the details of each single credit transaction of each borrower.

Existing credit registers

As regards the status quo, as indicated by the WG CR's early fact-finding exercise in May 2007, most existing national CCRs used to collect data at the borrower level. Romania and the Czech Republic were the exceptions. Since then, Latvia has introduced a loan-by-loan credit register in 2008 and Spain has switched from a borrower-level CCR to a loan-level CCR in 2013. Private credit bureaux collect information usually on a loan-by-loan basis as the purpose of their databases is different from CCRs and requires a higher level of detail. In 2012, the European Data Warehouse, a privately owned operator collecting loan-level data for asset-backed securities, was launched becoming the first European loan-level database.

Table 2
Reporting systems and reporting thresholds in national CCRs

	AT	BE	CZ	DE	ES	FR	IE	IT	LV	LT	MT	PT	RO	SI	SK
System	BbB BbB LbL	BbB LbL	LbL	BbB LbL	LbL	BbB LbL	LbL	BbB LbL	LbL	LbL	LbL	BbB LbL	LbL	LbL	LbL
Threshold	35,000	0	0	1,000,000	6,000	25,000	500	30,000	0	290	5,000	50	4,440	0	0

Sources: National central banks, ECB

Merits as defined by users

As usual, when defining new ECB reporting requirements, the choice to be made between a borrower-by-borrower (b-b-b) and a loan-by-loan (l-b-l) approach was to be based on an in-depth merits and costs analysis. Before running a formal merits and costs procedure, users within the ESCB and beyond were consulted and asked for their requirements. During its first meeting in 2012, the TF CR invited users in the areas of monetary policy, financial stability, research and statistics to express their data needs. In 2013, users including market operations, risk management and banking supervision were consulted as well. The JTF held another user workshop in February 2014. With respect to the choice between borrower-level and loan-level data collection, users expressed a clear preference for data on a l-b-l level.

During a workshop in 2014, users clarified the reasons for this choice. For the purpose of monetary operations, I-b-I data allows a cross-check to verify that the collateral features reported are correct, given that counterparties can put forward single loans as collateral for underlying Eurosystem credit operations. For research purposes, I-b-I data is essential to disentangle different types of loans and to allow a more precise analysis of the transmission channels of monetary policy. On the other hand, the European Stability Mechanism (ESM) indicated that information on a borrower level is useful for credit assessment, since it provides a complete picture of the risk exposure and the concentration towards a single borrower, as well as an overview of the commercial and pricing policy in the sector.

Based on the strong consensus expressed by users, already in February 2013, in its final report, the TF CR said that loan-level information would be strongly encouraged, allowing the highest usability and flexibility. At the same time, it acknowledged that b-b-b loan data can also provide comparable information, if adequately structured and complemented by additional details. However, the latter approach was considered to hamper a fully-fledged analysis of the loan portfolio, given the difficulty of distinguishing between distinct loans with similar characteristics.

A third way

As it was clear a priori that I-b-I information would be more costly to collect, store and disseminate than b-b-b data, a third option was envisaged. Under this option, the dataset would, in principle, be at the borrower level but loan-specific attributes would be added for each borrower, thus allowing for analysis that would go beyond borrower granularity and tend towards a loan-level dataset. It turned out, however, that users did not see added value in the proposed compromise, as the "mixed approach" was broadly deemed as inferior to pure b-b-b data collection.

Other possible options were also considered. Initially, the JTF investigated the possibility for the level of granularity to be different by sector, industry activity and borrower size as well as by the nature and type of loans. At a later stage, a phase-in I-b-I approach was considered, whereby loan-level reporting would become mandatory only after a certain phase-in period. Finally, it was proposed that the choice of the level of granularity could be at the discretion of each national central bank.

Pros and cons of the two data models

Being a very fundamental decision, the choice between b-b-b and I-b-I was one of the key issues discussed from the beginning. During the JTF's first meeting in June 2013, the transition from borrower-level to loan-level data collection of the Spanish CCR was presented to the participants. Another JTF sub-group would from then on deal with the users' view of the I-b-I v the b-b-b approach.

The work of the AnaCredit team at the ECB as well as the discussions in the JTF led to a refined list of pros and cons of the I-b-I approach. First of all, loan-level data would allow for a higher degree of usability, meaning that more detailed analyses could be performed. Examples include analysing the credit risk of a loan and the deterioration of credit assets for microprudential supervision. Moreover, loan-level data would allow for the monitoring of restructured or refinanced loans as well as tracking the mobility of loans and interest rates. I-b-I data collection would also

increase the flexibility of the AnaCredit dataset, allowing for further ex-post aggregation as needed, without defining ex-ante all combinations of variables. As aggregation can be done at ESCB level, reporting agents would be relieved of this burden. Finally, a loan-level dataset would allow for a link to other micro datasets within the ESCB as well as the European Data Warehouse and data from private credit bureaux.

There were three basic factors mitigating against the introduction of a loan-by-loan data model. Firstly, loan-level data collection would be costly as existing CCRs would have to switch from borrower-level data collection. Secondly, loan-level data collection would be very demanding for the collecting authority in terms of data storage and data quality management. Finally, b-b-b data could be a good proxy for l-b-l data if further loan-specific variables were included.

The AnaCredit team of the ECB assessed the pros and cons of an l-b-l system by comparing the two options along ten dimensions: degree of detail, system flexibility, complexity of the reporting scheme, complexity of the analytical tool, reporting externalities, IT demand for reporting agents, IT demand for the ECB and the NCBs, amount of data, data quality management, and user requirements. The ECB concluded that six out of these aspects weighed in favour of l-b-l while only three aspects were supported in a b-b-b approach.

Discovering the merits of loan-by-loan reporting

A draft of the vision paper from November 2013 contained an important point in favour of an l-b-l data model. Under l-b-l reporting, as users of the information, reporting agents would be able to perform a finer and more complex analysis than is currently feasible on the basis of the data exchanged under the MoU. Crucially, the reporting of loan-level information may lead to the reduction of the reporting burden since the level of detail would minimise the need for reporting agents to aggregate data based on criteria that may be alien to their internal reporting schemes. This could, in turn, make reporting requirements more stable over time as already witnessed for security-by-security reporting for other ESCB databases.

In February 2014, the JTF conducted a dedicated workshop on the content of existing CCRs. An interesting proposal was raised, i.e. to define an l-b-l data collection model whereby, under l-b-l reporting, reporting agents would provide a sort of copy of their internal records. Based on this raw data, the collecting authority would then compile and calculate whatever output is needed by users. Seen from this perspective, l-b-l would be a more natural approach for data collection, as it is closer to banks' business models. On the downside, l-b-l would increase the workload of the collecting authority in terms of data aggregation and data quality issues. Eventually, the outcome of the workshop was clear. While six ECB business areas favoured AnaCredit to become a loan-level dataset, only one business area preferred a borrower-by-borrower data model.

Taking the clear preference of the users into account, in its April 2014 report, the JTF acknowledged that l-b-l is the optimal reporting model for granular information in the long term. However, it was still envisaged that both reporting models would be accepted and that reporting models should be aligned over time towards loan-level reporting.

Costs

While the JTF favoured a loan-by-loan approach by early 2014, the final agreement on the matter was subject to the outcome of a detailed cost assessment. The cost assessment was conducted as part of the larger merits and costs procedure. The outcome indicated that reporting on a I-b-I basis would be slightly more costly than the b-b-b option. However, the overall picture depended largely on whether the cost estimate took into account the model of the CCR in each country. For countries with a b-b-b CCR, the average costs would rise considerably, although they were largely driven by the costs of one country. On the other hand, for countries with a I-b-I CCR or with no CCR, the cost increase for reporting on a I-b-I basis would be very low. It was therefore concluded that the bulk of the cost associated with the reporting on a I-b-I basis is concentrated in very few countries and that, overall, the I-b-I reporting would imply minor additional costs.

The JTF's alternative proposal and the STC's decision

Given the overall benefits of a loan-by-loan dataset as well as the limited costs, in autumn 2014 the JTF proposed progressively establishing AnaCredit as a granular dataset on a I-b-I basis. As a result, loan-level reporting would become mandatory after a certain time period while during an initial stage borrower-level reporting would also be accepted. However, by the end of 2014, reporting agents had expressed their clear preference for a one-off solution, i.e. a reporting framework that would be stable from its inception. Moreover, the timetable had proved to be increasingly ambitious. The ECB team therefore proposed a delayed, but more definitive, implementation of AnaCredit, including the introduction of I-b-I reporting from the start of the data collection. The STC accepted the proposal due to the more realistic timetable and taking account of the views of both the users and the reporting agents.

2.3 Coverage of instruments, lenders and borrowers

The scope of the AnaCredit database can be measured according to four dimensions, relating to the coverage of (i) instruments, (ii) lenders, (iii) borrowers, and (iv) the reporting threshold⁹. The AnaCredit Regulation that was eventually adopted stipulates that the instruments to be reported are (mainly) loans, while credit derivatives and debt securities are not included, at least not in Stage 1. As for the threshold, €25 000 at borrower level was considered the most viable compromise between a good coverage of loans to small and medium sized enterprises (SMEs) and a reasonable reporting burden. In terms of lenders, AnaCredit covers credit institutions, a subset of Monetary Financial Institutions (MFIs). Other Financial Institutions (OFIs) are excluded for now. On the borrower side, legal entities are covered, which will mainly mean non-financial corporations (NFCs) (and especially SMEs, given the low reporting threshold), but General Government is also included. Households might be included at a later stage.

Defining a wide coverage

The widest possible coverage conceivable for AnaCredit would include all instruments that give rise to credit risk. This would include both credit derivatives and debt securities. Given that all of these instruments would be reported on, a picture of

⁹ The discussions on the reporting threshold are covered in section 2.4.

a borrower's overall indebtedness could be reliably obtained by users. Moreover, all lenders involved in issuing loans and debt securities as well as loan-like instruments could be included. The widest possible coverage of borrowers would also include households, alongside companies and General Government.

Pros and cons

From a users' perspective, AnaCredit should indeed have the widest coverage as described above. It would enable users to analyse the overall indebtedness of a borrower and allow for analysis relevant for various ESCB tasks, as requested by user committees. One other example is the analysis of monetary policy transmission to various types of borrowers, including households. However, the costs for reporting agents as well as for the NCBs and the ECB could be substantial. The cost of reporting and quality check data varies depending on the instrument, type of lender and type of borrower. Other limitations for developing AnaCredit with maximum coverage are, for instance, legal and confidentiality issues with respect to reporting households as counterparties.

Considerations regarding instrument coverage

In 2013, the TF CR proposed prioritising loans of all types, including loans derecognised from the balance sheet. However, it was recommended to also cover credit derivatives and debt securities. The latter would not need to be collected directly through AnaCredit, but through existing databases such as the Securities Holdings Statistics Database (SHSDB) and the Centralised Securities Database (CSDB), while ensuring a good interoperability between the systems involved allowing for an easy combined use of information.

The draft vision paper reiterated the usefulness of maximum instrument coverage, at least from the perspective of banking supervision and financial stability analysis. From the supervisory perspective, all the financial assets of a reporting institution would ideally be within the scope of a granular credit database available to supervisors and financial stability experts since otherwise an assessment of a bank's solvency position conditional on a scenario is at best partial (with only a fraction of the loan portfolio being under scrutiny). A prioritisation should be introduced by means of thresholds rather than by asset class, thus not introducing an undue prejudice with regard to the borrower or the purpose of a loan which would be lacking a material justification from a risk perspective.

The JTF, in its April 2014 preliminary report, recognised that the definition of credit is a key element that may heavily influence the information content of AnaCredit. Taking a general approach, the JTF defined credit as an asset or off-balance-sheet item giving a contractual right to receive payment from a counterparty, i.e. loans, off-balance-sheet items, derivatives and (debt) securities. According to the JTF, this necessitates the widest coverage possible. Against this background, the JTF agreed to focus on information on loans, off-balance-sheet positions and derivatives. However, in a forward-looking manner, further harmonisation on a set of core attributes would be necessary, also covering data on securities regardless of where these are stored so as to enable an integrated view of all types of exposures and risks to be made available to the users. Hence, while avoiding the duplication of the CSDB and SHSDB, these datasets could be reviewed, also in terms of frequency and timeliness, to assess the extent to which they would cater for such an integrated view and identify which amendments may be necessary.

The merits and costs exercise confirmed the view of many users that all exposures would be absolutely necessary as otherwise the overall exposure and risk could not be calculated accurately. Regarding debt securities, several users indicated that such information could be retrieved from other systems provided that similar granularity for these exposures is ensured. The STC pointed out that information on securities, although not yet with the requested level of granularity, can be provided by the SHSDB in combination with the CSDB. It should be noted that, to ensure full coverage of the exposure in the form of securities, information on both instruments with and without an ISIN code would need to be collected. While an extension of the details currently covered by the SHSDB could be envisaged for instruments with an ISIN code (and indeed the Regulation of the SHS extension has recently been approved by the Governing Council), for instruments without an ISIN code, which are currently outside the scope of the CSDB, the inclusion in AnaCredit seemed the most appropriate way forward. This criterion would ensure that all instruments are covered and that there is no double reporting of the same information by reporting agents. Derivatives, on the other hand, were considered less important.

On the cost side, derivatives and debt securities without an ISIN code were considered very costly by reporting agents and NCBs. However, there was a large dispersion in cost estimates depending on whether these instruments were already reported or not in the respective jurisdiction. Debt securities with an ISIN code as well as other off-balance-sheet exposures were considered less costly.

Based on the outcome of the merits and costs exercise, the JTF proposed to include debt securities without an ISIN code in Stage 1 of AnaCredit, while debt securities with an ISIN code could be covered through appropriate extensions of the SHSDB and the CSDB. Moreover, off-balance-sheet exposures should be collected from the start, given the strong demand from users along these lines. On the other hand, given the high costs of derivatives reporting, these instruments should be reported from Stage 2 onwards.

After further discussions in the STC, the draft Regulation of December 2014 contained reference to five categories of instruments to be reported reiterating the importance of maximum instrument coverage: (1) loans granted or serviced (2) deposits held in other institutions (3) credit derivatives (4) off-balance-sheet exposures and (5) any credit risk exposure not reported through SHS. The outcome of the further discussions in 2015, however, emphasised the high costs of collecting additional financial instruments, including both derivatives and non-ISIN securities. Therefore, the final Regulation left out these instruments for the time being.

With regard to the definition of the reporting population, the STC initially favoured including all types of financial institutions. Beyond credit institutions, the shadow banking sector was also proposed to be considered by the JTF provided that users would see a clear added value. While noting that credit institutions are the only class of lenders currently reporting to all national CCRs, the vision paper pointed to the increased credit intermediation of non-credit institutions. Collecting microcredit data from these entities would therefore be even more important. Finally, in countries where loan securitisation is used widely, it could also be useful to include financial vehicle corporations in the reporting population.

Considerations regarding lender coverage, i.e. the reporting population

Users, however, proved to be mainly interested in granular credit data from monetary financial institutions (MFIs) while other lenders were less of a priority. Based on this prioritisation and against the background of the envisaged stepwise implementation of AnaCredit, the JTF proposed initially including only credit institutions in the reporting population. In future stages, financial corporations involved in lending would be included as users expressed their need for this data although according to a lower priority than data coming from credit institutions. It was proposed to include loan securitisation through the reporting of the credit institutions servicing the loans.

As with all the other decisions regarding the design of AnaCredit, the final proposal was subject to the outcome of the merits and costs procedure. In the framework of the cost assessment exercise conducted in June 2014, NCBs reported sizeable costs of including non-credit institutions. Hence, in the course of 2015, the decision was taken by the STC that AnaCredit would cover only credit institutions in Stage 1, possibly widening the scope of lenders in future stages.

Considerations regarding the coverage of borrowers

With regards to the scope of borrowers to be included in AnaCredit, the STC initially prioritised non-financial corporations with a particular focus on SMEs and possibly the General Government sector. At its meeting in November 2013, the JTF concluded that, with respect to the inclusion of households, benefits and costs need to be assessed with a special emphasis on the level of granularity, the types of household loans and the degree of anonymisation required to satisfy user needs.

In February 2014, the users echoed the suggestion of the STC. While non-financial corporations were deemed an overall priority, the view was that data on government and household borrowing should also be collected. On this basis, the JTF recommended proceeding with a stepwise implementation of borrowers in AnaCredit. In Stage 1, non-financial corporations and General Government would be included in the borrower population. In Stage 2, households would be introduced so that loans for housing purposes were covered in AnaCredit. All remaining exposures would be reported from Stage 3 onwards.

After further discussions within the JTF, this proposal was changed. Accordingly, in Stage 1, all exposures of legal entities would be included, so that it comprised non-financial corporations, General Government, supranational institutions and financial institutions. The concept of legal entity was defined as any entity which, under the national law to which it is subject, can acquire legal rights and obligations. Furthermore, households were split more explicitly according to the purpose of the loans. Stage 2 would include exposures to households for housing purposes while Stage 3 would include other exposures to households and non-profit institutions serving households.

By the end of 2014, the initial timetable had become increasingly ambitious. Therefore, the ECB team developed an alternative proposal to the outcome of the JTF discussions. As reporting agents asked for a more thorough one-off implementation, the ECB team proposed postponing the implementation date further and including all exposures from the start. While the STC initially welcomed the proposal, it was finally decided in March 2015 that exposures to households and sole proprietors would not be included until Stage 3. However, the STC noted that NCBs

would have the flexibility to collect this information already from Stage 1. The AnaCredit Regulation therefore contains all legal entities as borrowers.

2.4 Reporting threshold

As reporting thresholds are one of the key differences across CCRs in Europe, the JTF engaged in discussions about the introduction of a harmonised reporting threshold in AnaCredit. The reporting threshold is the outstanding amount of an instrument (in the case of a l-b-l reporting) or of instruments (in the case of a b-b-b reporting) above which the instrument or instruments have to be reported at a given reference date. AnaCredit will have a reporting threshold of €25 000 at the level of the debtor referring to the total outstanding amount (i.e. including off-balance-sheet amounts). The threshold is the same across all borrowers and all countries.

Pros and cons of a low and unique reporting threshold

A low and unique reporting threshold has several clear advantages from the perspective of users. Generally speaking, the lower the threshold, the higher the coverage of instruments and borrowers. This is, in practice, particularly relevant when there is a need to capture small to very small companies as well as households as borrowers. Analysing monetary policy transmission to SMEs, for instance, requires a low threshold in order to capture a sufficient share of the SME population. Another example for which a low threshold would be crucial is a financial stability analysis of risks originating from the household mortgage sector. Moreover, a unique threshold will allow for extracting harmonised and therefore comparable statistics from the database. Different thresholds across countries, for instance, would impede the usability of the information, as is currently the case with existing national CCRs.

On the other hand, the lower the threshold, the higher the costs for both the authority collecting the data and the reporting agents, as the resulting volume of information to be reported grows exponentially. The merits and costs exercise has shown that the marginal costs of lowering the threshold vary considerably across countries depending, for instance, on the degree of fragmentation of the financial system and on the average loan size. Crucially however, for countries not running a CCR or similar credit reporting, overall the costs were higher and not significantly affected by (the increase of) the threshold. A low threshold involves processing large amounts of data and, in particular, the workload for the unique identification of (more numerous) borrowers, including very small ones with hardly accessible reference information, and data quality management. Conversely, the automation process is greatly facilitated with standard processing of all data pertaining to a certain category of contracts. Besides, the downside of a unique threshold for all countries is that it creates imbalances between large and small countries, for instance, depending on the degree of concentration of the banking industry and on the average firm size.

Defining a threshold in AnaCredit

A fact-finding exercise in May 2007 found thresholds varying from €50 in Portugal to €1.5 million in Germany.¹⁰ Generally, the reporting thresholds of CCRs were found to

¹⁰ Cf. Table 2. The threshold in the German CCR was subsequently lowered to €1 million.

be higher than those of private credit bureaux. This stems from the different purposes of the two types of register. Based on this finding, in its final report in 2013, the TF CR recommended that (1) thresholds should be reasonably low and harmonised and that (2) they consider each relevant instrument. This implied that the threshold should not vary across countries but should instead be different for each type of instrument. The Task Force therefore highlighted the need for the different thresholds to converge over time.

Following the suggestion made in the draft vision paper in September 2013 to stratify the borrower population (as described above), the JTF also considered applying different thresholds to different borrowers, e.g. depending on the sector or size. Each of the strata could be a representative sample while leaving the relevant NCB in charge of the design of the system and the particular threshold needed to fulfil the condition of representativeness. Being in charge of assessing the viability of various solutions with regard to the reporting threshold, the JTF sub-group was of the opinion that a unique threshold across borrowers would simplify the collection of data as well as allow for an adequate comparability of information. The sub-group proposed a unique threshold of €25 000 with an exception for non-performing loans which should be collected irrespective of the outstanding amount due to the high value of information for these exposures.

While users gave the feedback that the threshold should be as low as possible – ideally €10 000, it became clear that the distribution of credit exposures across countries differs. The same euro amount threshold applied to all countries would imply large differences in the coverage of both outstanding exposures and number of borrowers. Consequently, the JTF sub-group recommended implementing different reporting thresholds across countries based on a common percentage of total outstanding amounts, e.g. 95%. This would imply a different euro amount threshold for each country. Finally, as the application of a percentage to calculate a euro amount threshold would have the undesired outcome of many multiple different thresholds (possibly one per country), in order to harmonise the samples of information reported, only a few thresholds may be used to delimit the sample of information to be reported to AnaCredit on a granular basis: €500, €10 000, €25 000 and €100 000. In its April 2014 report, the JTF reiterated this proposal and added that the threshold should be computed using credit exposures to NFCs in the form of all instruments, i.e. loans, derivatives and securities.

Assessing merits and costs

All proposals up until then were explicitly made conditional upon the forthcoming merits and costs exercise. The exercise asked three questions related to the threshold. The first was asking for the merits and costs of a specific unique threshold whereas the other two asked for the merits and costs of country-specific and instrument-specific reporting thresholds v unique thresholds. While the change in the cost of lowering the threshold was estimated to be rather small, the highest marginal benefit seemed to be with a threshold of between €10 000 and €50 000 depending on the user asked. Revealing further the advantages of a single reporting threshold across countries and across instruments, the merits and costs exercise showed that the benefits of a single threshold outweighs the cost. From a cost perspective, the exercise showed that there is a) a preference for a common threshold across

countries b) a preference for sufficient coverage of NFCs and c) a preference for defining the threshold on the basis of loans.

Therefore, at the end of 2014, the JTF proposed to the STC to proceed with the implementation of a low, single threshold of €25 000 calculated at borrower level and on the basis of loans and off-balance-sheet amounts (e.g. undrawn credit lines). This proposal was based on the feedback received from users, the majority of whom indicated that such a calculation would ensure homogeneous treatment across exposures, thus enhancing the consistency of the system, especially taking into account that such exposures were to be the main focus of AnaCredit. This view was also shared by most countries. However, since some users and some NCBs strongly favoured taking into account all relevant exposures in defining the threshold, an early draft of the Regulation contained a threshold along these lines.

In the end, however, a calculation on the basis of loans and off-balance-sheet amounts only – as favoured by a large majority of users and JTF participants – was deemed a superior solution. This was not only due to its simplicity, but also because in the end some types of exposures other than loans were not included in the AnaCredit Regulation.

2.5 Individual v consolidated reporting

Another important topic discussed by the JTF was whether reporting should take place on a solo or consolidated basis. Several mixed options were discussed as well. The discussion was conducted against the background of the new microprudential supervision task of the ECB, which is largely based on a consolidated view of banking groups' exposures. The question was whether this clear user requirement would justify the (high) cost of requiring reporting agents to report consolidated data. For the purposes of AnaCredit, data is reported on a solo basis for the time being. However, a recital in the Regulation mentions that data on a consolidated basis might be collected in future stages of AnaCredit.

Outcome and definition

Article 3(2) of the Regulation lays down that reporting agents shall report data on an individual basis. Article 1(26) clarifies that "on an individual basis" means with reference to a single institutional unit, including institutional units that are part of a legal entity. This is in contrast to reporting on a consolidated basis under which groups of legal entities report instead of their individual units. The consolidated approach is particularly relevant for banking supervision as this is conducted at group level, i.e. stocks and flows within the group are consolidated. In principle, data on a solo basis can be used to generate data on a group basis by appropriate consolidation. In practice, however, this is hampered by different accounting standards in different countries, which limits the possibility to aggregate and/or consolidate data.

Pros and cons

The central reason for including data on a consolidated basis in the future is the new user requirement of banking supervision to have readily comparable data on credit exposures. However, a consolidated approach means higher costs for either the

reporting agents or for the ESCB, depending on the option chosen for generating the consolidated data. It could, furthermore, involve operational risks for the ECB and, to a lesser extent, the NCBs.

Discussion

In 2013 the TF CR recommended defining the reporting population on the basis of solo institutions while a consolidated approach for the purposes of banking supervision requirements could be taken into consideration. In the months that followed, the JTF engaged in detailed discussions. There were two general options on the table. First, AnaCredit could require reporting agents to report directly on a consolidated basis. As a second option, compiling information on a group basis from information reported on a solo basis was discussed.

The first option would make reporting agents directly responsible for the consolidation. This would avoid placing a higher burden on the ESCB in terms of data consolidation, but would come at a considerable cost for reporting agents. It could furthermore generate overlaps with information reported on a solo basis. This procedure, nevertheless, could be followed in some cases, especially to support prudential supervision. This may be the case for flagging large exposures.

Alternatively, information on a group basis could be compiled from information reported on a solo basis. In this approach, the set of granular information reported on a solo basis may be used to build information on a group basis for different groups and can thus ensure full consistency and allow users to delve further into any questions to get the full picture. This solution is possibly less costly for reporting agents compared to additional reporting on a group basis. NCBs and the ECB would need to integrate the granular exposures based on the information about the group structures, placing a higher burden on the ESCB as a whole.

More precisely, due to difficulties in handling different accounting frameworks (e.g. different methods of valuation, provisions, loan recognitions, netting of exposures, etc.) in good time, the resulting reporting obligations may be as costly as the direct reporting on a group basis for the reporting agents; they would additionally involve operational risks for the final compilers - the NCBs and the ECB.

On this basis, the consolidation of group information based on solo data came to be considered as a longer-term objective that might not be included in AnaCredit from the start. However, an intermediate step could be that information is reported on a solo basis and is enriched in such a way that it supports analyses made on a group basis.

Assessing merits and costs

While the JTF considered this issue important, it took until the merits and costs exercise to form a nuanced view on the matter. First and foremost, the merits assessment confirmed the importance of consolidated data. It also showed the importance for many users of information on foreign branches outside the euro area. The SSM gave this information the highest rating and indicated that, to be consistent with the supervisory approach of supervising the banking group as a whole, information on foreign branches and subsidiaries would be necessary. Ideally, this information could be reported on a consolidated basis. However, due to the high

complexity of preparing this information, it suggested that the focus should be on solo data for the time being.

The cost assessment, on the other hand, indicated very high cost increases for reporting agents when including foreign subsidiaries and the reporting of exposures on a consolidated basis. In fact, collecting data on a consolidated basis was deemed one of the most costly features of AnaCredit. In contrast, the inclusion of foreign branches in the reporting population as well as the identification of entities within a banking group would cause significant, but smaller, cost increases. Moreover, there were large discrepancies in cost estimates across countries.

In an endeavour to find a better balance between the merits and the costs of consolidated reporting, the JTF put forward a new proposal. To counterbalance the high costs for the provision of consolidated information therefore, only a subset of information assessing the credit risk of the banking group should be collected on a consolidated basis. Nevertheless, as the parent company reporting on a consolidated basis and the financial institutions which were part of the banking group, would use the same identifiers for all the information reported, the analysis of the complete dataset would be possible both at consolidated and individual level, and would avoid double reporting. This proposal initially found support in the STC, which agreed on an obligation for the 120 most significant banking groups to report consolidated data to AnaCredit.

However, in 2015, the cost implications of including consolidated information in AnaCredit came under discussion on whether to cater for it already at that stage and, if so, to establish the criteria for consolidation. The JTF proposed not to include consolidated reporting in AnaCredit in Stage 1 but instead to collect data on a solo basis for foreign branches and use group structures via the RIAD register to get a proxy of the consolidation on a "home solo basis" (i.e. to combine information between legal entities and their foreign branches in the EU, and the possibility for some countries to extend to all foreign branches; conversely, foreign subsidiaries would not be covered in Stage 1). Hence, the AnaCredit Regulation stipulates that the data have to be reported on an individual basis and foresees the introduction of a consolidated approach for a later stage.

2.6 Derogations

The ECB has always attached great importance to the need to keep the burden on reporting agents to a minimum. The possibility to grant derogations to small reporting institutions is an important tool to achieve this goal. For AnaCredit, derogations for some reporting agents began to be considered in 2013. Discussions mainly revolved around two points: first, the distinction between absolute (in terms of euro amounts) and relative (in terms of % of the overall potential information) thresholds; second, the issue of national discretion v a unified criterion for all banks across the euro area.

The AnaCredit Regulation allows NCBs to grant derogations to avoid imposing an undue reporting burden, particularly on small reporting agents with limited total credit

Outcome

exposure. For each reporting Member State, there is scope within the AnaCredit Regulation to grant such derogations to a subset of reporting agents provided that the total amount of loans reported pursuant to the BSI Regulation¹¹ by such reporting agents does not exceed 2% of the overall amount of loans reported pursuant to the BSI Regulation in a given reporting Member State. The granting of derogations is at the discretion of the relevant NCB.

Moreover, the AnaCredit Regulation provides that certain reporting agents may be temporarily exempted from the obligation to report on a monthly basis, at the NCBs' discretion. Specifically, in line with Article 16(2), small reporting agents in a given reporting Member State may report on a quarterly basis until the end of 2020 provided that the total outstanding amount of loans of such institutions (as reported in accordance with the BSI Regulation) does not exceed 4% at national level. The AnaCredit Regulation leaves it up to the relevant NCB to determine whether a certain reporting agent may report on a quarterly instead of a monthly basis.

Pros and cons

Derogations for small reporting agents were considered for two reasons. Firstly, the possibility for NCBs to grant derogations ensures proportionality with respect to reporting requirements. Data from small banks might be insignificant in terms of volume to make a substantial difference and thus these banks should not be obliged to have the same requirements as larger banks. This argument is supported by the fact that small banks are more affected by reporting requirements compared to their larger peers which have more resources or opt to automate part of the reporting. Secondly, derogations are in line with the concept of subsidiarity. Having more detailed knowledge about their respective reporting agents, NCBs are in a better position to judge which small reporting agents should be granted a derogation.

On the other hand, there are two important reasons for not having derogations for AnaCredit reporting requirements. First of all, data coming from small reporting agents can be just as important as the data coming from larger agents, taking into account the role of contagion and confidence in the financial system. The information could be used to analyse smaller banks in particular, for instance when analysing monetary policy transmission to small banks relative to larger banks. From a financial stability perspective, the introduction of derogations would represent a risk to overlook a blind spot, i.e. the condition of small banks within the banking system. Secondly, NCBs deciding on derogations to their reporting agents runs counter to the idea of a level playing field across reporting agents in the euro area. Two banks competing across borders could be granted a derogation in one country but not in another, with potential consequences for competition.

Discussion

The first reference to derogations during the JTF discussions was made already in November 2013. It was contemplated that, depending on the current state of CCRs in the Member States, temporary derogations could be granted to some NCBs or groups of reporting agents. Consequently, only a subset of information would have to be reported for a certain period of time. Comments from the industry, however,

¹¹ Regulation ECB/2013/33 concerning the balance sheet of the monetary financial institutions sector.

emphasised that it was the size of reporting agents that should be the criterion for determining derogations. This approach was also put forward by the STC.

Since then, numerous proposals have been discussed in the JTF regarding derogations for small reporting agents. The proposals can roughly be categorised into three groups: derogations with relative thresholds, derogations with absolute thresholds and partial derogations. Proposed derogations with relative thresholds were, firstly, based on the combined contribution of derogated reporting agents to the national MFI balance sheet. Another proposal was based on a threshold relative to the combined contribution of the derogated reporting agents to the national outstanding amount of the instruments. Finally, one proposal made derogations dependent on the combined commitment amount of all derogated reporting agents. The proposed thresholds varied between 1% and 5%.

One proposal put forward in the JTF was a derogation based on an absolute threshold of total loans and undrawn credit lines below €1 bn for a small reporting agent to be granted a derogation from the relevant NCB. Combining an absolute threshold with a relative threshold was also considered and giving NCBs the discretion not to exempt a small reporting agent in case it belonged to a particular category of reporting agent.

In striving for a good balance between providing a complete and representative dataset for users on the one hand, and a proportionate reporting burden for small reporting agents on the other, the JTF reflected on partial derogations. A partial derogation could be granted by an NCB to a small reporting agent effectively exempting the latter from reporting a certain set of attributes to AnaCredit. The JTF also considered partial derogations applying only temporarily.

The final Regulation contains both the possibility of a derogation with a relative threshold and a partial derogation. Firstly, it stipulates that NCBs may grant derogations to small reporting agents, provided that the combined contribution of all reporting agents granted a derogation does not exceed 2% of the total outstanding amount of loans reported by all reporting agents in the reporting Member State. Secondly, these derogations can be either full or partial with respect to the amount of attributes to be reported. Finally, the Regulation also contains the possibility for NCBs to grant temporary derogations for the reporting frequency provided that their share does not exceed 4% of the total outstanding amount of loans.

2.7 Counterparty reference data

Another important topic relates to the counterparty reference data needed for a meaningful disaggregated analysis of credit and credit risk (e.g. by sector, size or location of the counterparty) and for the unique identification of counterparties (lenders, borrowers, guarantors, etc.) in AnaCredit. The latter aspect, in particular, is central for a shared supranational system like AnaCredit. However, that does not mean that such data, which are functional to the analysis and classification of the AnaCredit exposures, need to be stored and managed together with the credit data.

A different dedicated dataset could also be envisaged, especially considering the benefits of using such "reference dataset" for several purposes, i.e. to serve other granular datasets of the ESCB beyond AnaCredit.

Moreover, information collected from reporting agents via the AnaCredit Regulation could be complemented and cross-checked by using other sources of information available at national level, such as business registers. The latter option, though, raises confidentiality and accessibility issues. An important aspect was the definition of a clear governance of the reference data, allowing the construction of the "true" representation of the reference data of a counterparty based on possibly conflicting information coming from different reporting agents in different countries.

Outcome

The Regulation contains a recital which mentions that the framework for the collection of credit data should be set up with a view to ensuring interoperability with the Register of Institutions and Affiliates Database (RIAD) maintained by the ESCB. Although no further reference to RIAD is made elsewhere in the Regulation, the JTF proposed, as the WG AnaCredit also did later, using the RIAD system (already including information on financial institutions) as a repository to store and maintain the reference data on all AnaCredit counterparties thus includes non-financial corporations and other legal entities.

Pros and cons

Using RIAD for the purpose of AnaCredit has several advantages. In general terms, storing reference data in a separate dataset acknowledges the difference between credit and reference data. Reference data, for instance, has to be updated only in case changes take place while credit data has to be reported on a continuous basis to fulfil the purpose of AnaCredit. Moreover, reference data received in the framework of AnaCredit can be used for various other micro ESCB databases beyond AnaCredit itself. In fact, the unique identification of counterparties is a prerequisite for the integration of micro databases, which is strongly requested by users willing to assess the overall exposure of a lender to a borrower (e.g. both in the form of loans and debt securities). Moreover, the RIAD system and related data model already covers several functionalities which are also needed for AnaCredit. Finally, data from RIAD on ownership relations can be exploited to consolidate exposures at group level.

However, the use of a shared reference dataset requires the resolution of conflicts potentially arising between NCBs when they upload different data on common counterparties; this could easily happen in cross-border exposures where a borrower has taken out loans from more banks in different countries and therefore several NCBs besides the resident NCB have to report its reference data. More crucially, using RIAD in the framework of AnaCredit has significant implications for the functioning of RIAD, as around 15 million legal entities are expected to be included in the repository dataset as AnaCredit counterparties.

Discussion

One possible option considered at an early stage of the discussion was to exchange information (and resolve conflicts) on a bilateral basis by the two NCBs involved, in line with the practice followed under the Memorandum of Understanding (MoU) for the exchange of information among national central credit registers for the purpose of passing it on to reporting institutions. However, the JTF deemed this process too

cumbersome and resource intensive given the high level of expected data volumes; it was hence not suitable for AnaCredit, also in light of the strict reporting deadlines for NCBs.

Instead, in its final report, the TF CR noted that information on all counterparties is better recorded in RIAD than in AnaCredit itself. This point was taken up in the vision paper prepared by the AnaCredit team at the ECB. As for the need to identify all entities uniquely in the system and avoid duplications, while the use of the Legal Entity Identifier (LEI) was recognised as a long-term goal, the benefits of using RIAD were acknowledged, as this would provide access to a broader set of identifiers.

Accordingly, in April 2014, the JTF proposed combining the RIAD and the AnaCredit datasets, with the former providing the repository for the counterparty reference data included in the Regulation. The JTF was of the opinion that the most efficient solution was the use of RIAD as a shared platform, accessible by all NCBs, to store and maintain reference data on all borrowers. Under such an approach, both reference data on purely domestic borrowers and on entities with cross-border exposures is stored and maintained in RIAD by the "competent NCB" (the resident NCB), or, outside the AnaCredit perimeter, directly by the ECB. In turn, such an approach could serve other shared datasets showing data on legal entities such as the CSDB and the SHSDB beyond the financial sector in the future.

For the process to work, the JTF identified a workflow whereby mutual cooperation between all NCBs involved in the cross-border exposure is requested to ensure the unique identification of counterparties in RIAD, under the assumption that, because the large majority of cross-border borrowers are in the first place also domestic borrowers, a bilateral exchange will be needed only on the margins, i.e. to resolve issues of no identification or multiple identification of the borrower. The WG AnaCredit pursued this idea further, including the possibility of a central matching mechanism allowing the automatic detection of potential duplicates among all registered entities, for the competent NCB to take a final decision on the need to proceed with an actual deduplication.

Outcome

As a result, the RIAD system is currently undergoing a significant capacity expansion (RIAD 4.0 project) to make it fit for use to manage the high volumes expected and the additional tools (i.e. matching tool) requested to support the AnaCredit business needs.

2.8 Interoperability with existing granular datasets

Another issue that was discussed, already at the start of the AnaCredit initiative, was its interoperability with other micro datasets. While for the purpose of reference data, AnaCredit was envisaged to be connected to RIAD, a particular focus was laid on its interoperability with the ESCB's databases on securities holdings and issuance.

The Securities Holdings Statistics Database (SHSDB), collected on a security-by-security basis, provides information on securities held by euro area resident sectors, broken down by instrument type and issuer countries. The Centralised Securities

Database (CSDB) is a securities reference database that holds complete, accurate, consistent and up-to-date information on the issuance of all individual securities relevant for the statistical purposes of the ESCB.

Outcome

The AnaCredit Regulation contains two references to the ESCB's securities database in the recitals. Accordingly, the framework for the collection of credit data should be set up with a view to ensuring interoperability with the databases on securities statistics. Moreover, the Regulation stipulates that credit data will improve the usability of the securities databases. While not being mentioned more explicitly in the legal text, work is currently underway at ESCB level in order to integrate AnaCredit with the aforementioned databases.

Discussion

The TF CR already recommended that AnaCredit be interoperable with the SHSDB, the CSDB and other data sources for credit exposures. Noting the interest of users in connecting data that allows them to get an overall image of total borrowers' indebtedness and lenders' credit exposures, the JTF was mandated to further investigate the role of securities within AnaCredit. Three options were considered: (i) to use RIAD to combine the information of the three databases without further modifications; (ii) to harmonise the information collected in AnaCredit and the SHSDB by collecting securities without an ISIN code (not currently reported for the SHSDB) in the set of exposures reported for AnaCredit, e.g. private placements, and extending the detail collected in the SHSDB so that both databases provide consistent information allowing the assessment of all exposures of single credit institutions; and (iii) to integrate both frameworks in a single Regulation. For this purpose, the JTF set up a sub-group on securities to analyse the optimal integration of securities in AnaCredit and the connection with the CSDB and SHSDB.

The sub-group conducted a fact-finding exercise in June 2014 which was meant to examine the possible ways of combining the two datasets with AnaCredit and find out if and how the identifiers used in the SHSDB and CSDB can serve this. In addition, it also tried to identify the possible methodological mismatch between the CSDB, SHSDB and AnaCredit. The outcome of the exercise showed that the integration alone of AnaCredit with the existing statistical framework would not lead to a full coverage of exposures. The JTF was thus of the opinion that AnaCredit should either comprise all types of exposures, in particular vis-à-vis NFCs, including securities, or the existing (shared) databases on securities would need to be enhanced and made more granular; their appropriate connection to AnaCredit via RIAD would then provide the required overview.

While the former approach may lead to double reporting, the latter may be more appropriate, but could take longer to implement; this would consist of increasing the level of granularity and coverage of instruments by e.g. including securities without an ISIN code, as indicated by the outcome of the fact-finding exercise.

In January 2015, on the basis of a note confirming the need to collect additional information regarding issuance and holdings of securities by credit institutions and possibly other financial entities (as previously identified by the JTF), the STC agreed on an update of the SHS Regulation as well as on necessary changes in the existing securities databases (CSDB and SHSDB). Four areas of change were proposed.

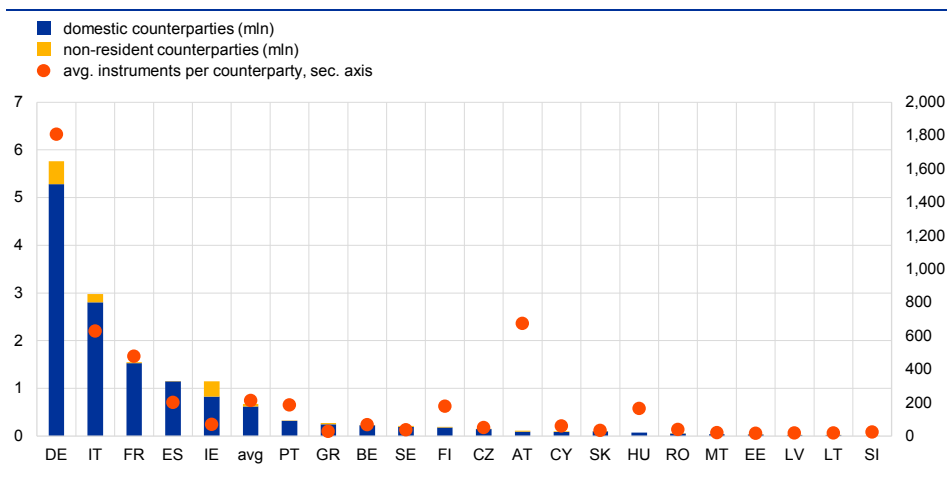
First, the level of granularity needs to be aligned as AnaCredit contains data on individual lenders whereas the SHSDB contains data aggregated by institutional sector or at the consolidated banking group level. Second, the data coverage should converge as reporting populations, thresholds, instruments and attributes differ. Third, integration requires the same identification method which could be facilitated by introducing RIAD as a common register. Finally, reporting frequency and timeliness would need to be aligned.

In the end, it was decided that debt securities are not to be covered in AnaCredit. The amended SHS Regulation, as approved by the Governing Council in April 2015, now incorporates the reporting of securities without an ISIN code. Work is underway within the ESCB to identify to what extent the systems allow for a combined use of the respective information and assess any possible future amendments to overcome bottlenecks hampering their full interoperability.

3 The way forward

The first data transmission to the shared AnaCredit database will take place in November 2018 with September 2018 as the data reference period. From then on, data collection will continue on a monthly and quarterly basis with an estimated volume of more than 100 million records per month. Until then, the ECB will continue to cooperate closely with NCBs, through the WG AnaCredit, and with reporting agents to ensure a smooth implementation of AnaCredit. For this purpose, an ECB Guideline, addressing secondary reporting by NCBs to the ECB, is currently being drafted. In addition, the Manual is being finalised to provide reporting agents with clear instructions on how to report the requested information, based on a number of specific examples.

Chart 1
Volume estimates for the first Stage of AnaCredit



Source: National central banks, ECB
Note: not based on the final AnaCredit Regulation

Implementing AnaCredit

Preparatory work has already started at the ECB and at participating NCBs, and progress will be monitored on a regular basis to ensure that any issue is detected at an early stage and suitable solutions are promptly identified and implemented; this will also include ongoing sharing of existing national experiences in the field of granular credit data collection.

The current situation with respect to credit data is quite different across countries. While some countries have been running a credit register similar to AnaCredit for several years now, in others, this is a completely new endeavour, also for reporting agents, and therefore the experience of the first group of NCBs is constantly shared to promote best practices. The banks are starting from quite different situations too. Big institutions are generally already well-equipped to fulfil the information requirements of AnaCredit, also because the kind of data requested is not different from that already used by banks in their daily activities (risk control, credit

assessment, etc.); and their systems are increasingly well-integrated and centralised to cater for the requirements under BCBS 239.¹²

Smaller banks may have to amend their systems significantly to comply with AnaCredit requirements; this is true at least of banks which are not part of a network and do not have shared IT services performing this task. Actually, most of the small banks that are not part of a network might be granted derogations.

Extensions of AnaCredit in the future

The go-live in September 2018 will mark the start of Stage 1 of AnaCredit. In future stages, extensions of the database are foreseen but not yet planned in the AnaCredit Regulation.¹³ In order to provide more useful data to its users, AnaCredit is envisaged to extend the scope of reporting with respect to types of lenders, borrowers and instruments.

The reporting population, i.e. the lenders obliged to report to AnaCredit, could be extended in scope. Stage 1 of AnaCredit only includes lending by credit institutions. As mentioned above, the discussions in the JTF pointed to the fact that lending by non-credit institutions has gained in importance in recent years. In order to capture the lending from shadow banks, future stages of AnaCredit could therefore incorporate deposit-taking corporations other than credit institutions as well as asset management vehicles and other financial corporations. Moreover, foreign subsidiaries of these entities as well as those of credit institutions could be included so as to enable users to obtain a more complete picture of lending activities by euro area financial corporations.

Secondly, the borrower population covered in AnaCredit could be enlarged. As repeatedly emphasised by ESCB users of statistical data, the lack of information on household exposures could significantly hamper the usefulness of AnaCredit in various contexts such as financial stability analysis, banking supervision and the analysis of monetary policy transmission. An extension to household exposures would likely be limited to loans for house purchases as these exposures are given the highest priority by the users. In any case, this implies an even stronger ECB focus on data confidentiality and an ability to handle ever larger volumes of data in the future.

Finally, the scope of instruments covered in AnaCredit could be widened. While covering mostly loans in the initial stage, future stages could extend the coverage to derivatives, other accounts receivable and off-balance-sheet items such as financial guarantees. In addition, the work on integrating data from AnaCredit with data from the CSDB and the SHSDB will continue so as to enable users to obtain a holistic picture of total borrower indebtedness and total lender exposure including securities.

Reporting on a consolidated basis

In view of increasing the usefulness of AnaCredit for users in banking supervision, subsequent stages might involve the inclusion of reporting on a consolidated basis. This would support the tasks of the ECB in microprudential supervision, in particular

¹² Standard published by the Basel Committee on Banking Supervision on "[Principles for effective risk data aggregation and risk reporting](#)"

¹³ Cf. Recital 12 of the AnaCredit Regulation.

with respect to conducting stress tests on credit risk, to identifying risks from significant cross-border concentrations and to monitoring banks' internal models and risk parameter estimates.

The process for future extensions

Any future extension of AnaCredit will be based on the experience gained in Stage 1. Furthermore, the expansion of AnaCredit will proceed in a stepwise approach so as to gradually extend the data collection without overburdening reporting agents. This will entail a second merits and costs procedure for all additional reporting requirements and will give NCBs and reporting agents at least two years to implement the additional requirements.

In order to further harmonise reporting requirements, another issue to be discussed for future stages will be national arrangements in the framework of AnaCredit. In fact, the Regulation foresees a reassessment of national discretion with respect to granting derogations for small reporting agents.¹⁴

AnaCredit in the context of other strategic ESCB projects

AnaCredit is an important building block in the strategy of the ECB with respect to statistics. The ECB statistical function is working with the involvement of the financial industry for designing and implementing a coordinated data management comprising information collected under different statistical and legal frameworks.

An important workstream in this field relates to the development of the Banks' Integrated Reporting Dictionary (BIRD)¹⁵, which describes a possible model for the input information banks need to report to authorities as well as the transformations necessary to generate these reports. Thereby, BIRD aims to foster cooperation in the field of regulatory reporting which alleviates the reporting burden for banks and improves the quality of data reported to the authorities.

The harmonised data model developed by the BIRD initiative describes precisely the data which should be extracted from the banks' internal IT systems to derive the reports required by the authorities. In addition to this, there are clearly defined transformation rules to be applied to the data extracted from the banks' internal IT systems in order to produce a specific final regulatory figure. The univocal categories of the data to be extracted from the banks' internal IT systems, the so-called "input layer", together with the transformation rules, make up the BIRD.

The purpose of the BIRD is to provide a service to the banks. The BIRD is available, as a public good, to banks and all interested parties. The adoption of the BIRD by banks is fully voluntary. It can be used as additional documentation (with respect to regulations and guidelines) or as an "active dictionary" for procedures developed by banks. The BIRD represents an input approach because it does not stop at the regulatory requirements, but goes back to the data in the banks' internal systems. However, banks remain responsible for the organisation of their internal reporting system. The first release of the BIRD is largely focused on AnaCredit as a new endeavour and constitutes a sizeable development for which stability of the underlying concepts and definitions is key. It fully defines the input layer for

¹⁴ Cf. Recital 11 of the AnaCredit Regulation.

¹⁵ More detailed information on the BIRD can be found on the website of the initiative.

AnaCredit and thus contributes to lowering compliance costs for banks and to increasing data quality.

Another significant example is the work, still in its early stage, towards the definition of an integrated European Reporting Framework (ERF). The idea behind the ERF is rather ambitious but potentially beneficial, i.e. to collect all data required for different statistical purposes and (in a second step) for banking supervision using an integrated and harmonised approach in all countries.

Both of these initiatives go in the direction of providing financial institutions with an integrated perspective, including integrated definitions of concepts and data attributes, with a twofold goal: (i) for the banks, to alleviate the statistical reporting burden by avoiding duplications and increasing clarity; (ii) on the side of the ECB, to ensure data quality and consistency and allow a combined use of all granular information (e.g. data on debt securities and, in future, credit exposures). In this context, given the scale and relevance of the project, AnaCredit provides a unique opportunity to put these initiatives in practice.

References

Bennardo, A., Pagano, M. and Piccolo, S. (2014), "Multiple bank lending, creditor rights, and information sharing", *Review of Finance*, Vol. 19, pp. 519-570.

Draghi, M. (2016), "Central Bank Statistics: moving beyond the aggregates", Speech given at the Eighth Statistics Conference, Frankfurt am Main, 6 July.

Issing Committee (2009), "New Financial Order Recommendations", CFS White Paper, No. 1, February.

Jiménez, G., Ongena, S., Peydró, J.L. and Saurina, J. (2014), "Hazardous Times for Monetary Policy: What Do Twenty-Three Million Bank Loans Say About the Effects of Monetary Policy on Credit Risk-Taking?", *Econometrica*, Vol. 82, No 2, pp. 463-505.

Pagano, M. and Jappelli, T. (1993), "Information sharing in credit markets", *Journal of Finance*, Vol. 48, No 5, pp. 1693-1718.

Padilla, A.J. and Pagano, M. (1997), "Endogenous communication among lenders and entrepreneurial incentives", *Review of Financial Studies*, Vol. 10, No 1, pp. 205-236.

Padilla, A.J. and Pagano, M. (2000), "Sharing default information as a borrower discipline device", *European Economic Review*, Vol. 44, No 10, pp. 1951-1980.

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

Acknowledgements

This paper is based on the work of the Joint STC/FSC Task Force on Analytical Credit Datasets in the following composition:

Ana Margarida de Almeida	(Banco de Portugal) – Chairperson of the Joint STC/FSC Task Force on Analytical Credit Datasets (JTF)
Ramón Santillán	(Banco de España) – Chairperson of the Joint STC/FSC Task Force on Analytical Credit Datasets (JTF)
Jean-Marc Israël	(European Central Bank) – Chairperson of the Working Group on Monetary and Financial Statistics (WG MFS)
Michael Ritter	(Deutsche Bundesbank) – Chairperson of the Working Group on Credit Registers (WG CR)
Violetta Damia	(European Central Bank) – Secretary of the Joint STC/FSC Task Force on Analytical Credit Datasets (JTF)
Orestes Collazo Brananova	(European Central Bank)
Ines Correia	(European Central Bank)
Vasileios Georgakopoulos	(European Central Bank)
Alberto Lorenzo Megia	(European Central Bank)
Gerhard Winkler	(Oesterreichische Nationalbank)
Uta Hammer	(Oesterreichische Nationalbank)
Bernhard Hirsch	(Oesterreichische Nationalbank)
Patrick Bissot	(Nationale Bank van België/Banque Nationale de Belgique)
Jiří Klučka	(Česká národní banka)
Irena Zykanová	(Česká národní banka)
Ingo Techet	(Deutsche Bundesbank)
Ingrid Seitz	(Deutsche Bundesbank)
Philip Sonne	(Danmarks Nationalbank)
Birgitte Vølund Buchholst	(Danmarks Nationalbank)
Gonzalo Sainz Caballero	(Banco de España)
Kristiina Kibin	(Eesti Pank)
Elisabeth Hintikka	(Suomen Pankki - Finlands Bank)
Jérémi Montornes	(Banque de France)
Benoit Fauvelet	(Banque de France)
Pashalina Karampasi	(Bank of Greece)
Konstantinos Zavandis	(Bank of Greece)
Nikolaos Kamberoglou	(Bank of Greece)
Aisling Menton	(Central Bank of Ireland)
Jean Cassidy	(Central Bank of Ireland)
Laura Mellone	(Banca d'Italia)
Maria Di Noia	(Banca d'Italia)
Darius Diečkus	(Lietuvos bankas)
Shirin Madani-Beyhurst	(Banque centrale du Luxembourg)
Irena Krumane	(Latvijas Banka)
Laura Ausekle	(Latvijas Banka)
David Josef Buhagiar	(Central Bank of Malta)
Wim Goes	(De Nederlandsche Bank)
Paula Casimiro	(Banco de Portugal)
Carmen Maria Lungu	(Banca Națională a României)
Daniela Radovici	(Banca Națională a României)
Maria Eriksson	(Sveriges Riksbank)
Pavol Kodaj	(Národná banka Slovenska)
Irena Klemencic	(Banka Slovenije)

Jean-Marc Israël

European Central Bank, Frankfurt am Main, Germany; email: jean-marc.israel@ecb.europa.eu

Violetta Damia

European Central Bank, Frankfurt am Main, Germany; email: violetta.damia@ecb.europa.eu

Riccardo Bonci

European Central Bank, Frankfurt am Main, Germany; email: riccardo.bonci@ecb.europa.eu

Gibran Watfe

European Central Bank, Frankfurt am Main, Germany; email: gibran.watfe@ecb.europa.eu

© European Central Bank, 2017

Postal address 60640 Frankfurt am Main, Germany
Telephone +49 69 1344 0
Website www.ecb.europa.eu

All rights reserved. Any reproduction, publication and reprint in the form of a different publication, whether printed or produced electronically, in whole or in part, is permitted only with the explicit written authorisation of the ECB or the authors.

This paper can be downloaded without charge from www.ecb.europa.eu, from the [Social Science Research Network electronic library](#) or from [RePEc: Research Papers in Economics](#). Information on all of the papers published in the ECB Occasional Paper Series can be found on the [ECB's website](#).

ISSN	1725-6534 (pdf)	DOI	10.2866/813758 (pdf)
ISBN	978-92-899-2849-6 (pdf)	EU catalogue No	QB-AQ-17-007-EN-N (pdf)