

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

Frank Mayerlen²

Introduction

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

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

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

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

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

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

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

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

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

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

Iterative approach – increasing required completeness and accuracy over time

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

Wide data usage – with appropriate health warnings

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

Section 2: operating and further enhancing a securities database platform

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

Operation

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

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

On-line access and joint data quality management

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

The DQM framework is fully harmonised and supported by dedicated CSDB system functionality, which provides six DQM metrics to prioritise and steer the DQM work as well as the necessary system functionality to correct data, if needed. DQM metrics are fully automated and compare massive amounts of data. As an example, the update of the metric identifying relevant updates in the CSDB data compares 23 attributes for two points in time for over 6.3 million instruments, i.e. about 300 million data points.³ The DQM framework has been formalised in an ECB Guideline, which will apply as of 1 July 2013.⁴

Data distribution and usage

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

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

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

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

⁴ See Guideline ECB/2012/21 http://www.ecb.int/ecb/legal/pdf/L_30720121107en00890103.pdf.

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

System maintenance and further development

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

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

Required staff profile

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

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