Integrated statistical systems: evolution or revolution?

Maria do Carmo Aguiar [1]

Evolution: a gradual process of change and development
Revolution: a very important change in the way that people do things[2]

1.   Introduction

Information technology developments have led to a new trend in the way statistics are produced: traditional aggregated reporting is gradually being replaced by item-by-item reporting. The advantages of this approach are enormous, ranging from lower reporting costs to greater flexibility in compilation. Although item-by-item reporting requires dealing with large volumes of data, this has become easier with the evolution of network and communication protocols, database systems and multidimensional analytical systems. Consequently, old multiple heterogeneous collection and compilation systems are being replaced by integrated systems.

The statistical data on securities periodically released by the Banco de Portugal are compiled from the Integrated Securities Statistics System (Sistema Integrado de Estatísticas de Títulos, or SIET), which was developed by the Banco de Portugal's Statistics Department to be a single repository housing all of the information needed for compliance with securities reporting requirements. SIET makes it possible to meet user needs at both the national and the international level. Quite ambitious in its aims, the system has been a challenge for data quality managers, and a source of opportunities for data "explorers".

With developments in financial markets worldwide, securities statistics are increasingly important, and the coverage, quality and harmonisation of securities statistics produced in the various countries are of growing concern at the international level, and within the European System of Central Banks (ESCB) in particular. In this context, integrated statistical systems facilitate more efficient and harmonised production of statistical data. SIET, as an integrated system with data on issues and portfolios, and with coverage of all of the economy's institutional sectors, makes it possible to cope efficiently with most information requirements in the field of securities statistics.

This paper provides an overview of SIET's architecture – inputs, processing and enrichment modules, and outputs – examining whether this integrated system represents an evolutionary development from previous disaggregated systems, or whether the new possibilities it creates constitute a revolution. A glance into the near future foresees the integration of information from the Centralised Securities Database (CSDB), which is an ongoing project of the European Central Bank (ECB).

---

[1]   Banco de Portugal, Statistics Department, Av. Almirante Reis, 71, 1150-012 Lisboa, Portugal. E-mail: mcaguiar@bportugal.pt.

[2]   In "Cambridge Dictionaries Online" (dictionary.cambridge.org).

2. SIET's features and main components

SIET is an information system that stores data on securities issues and portfolios on a "security-by-security" (s-b-s) and an "investor-by-investor" basis, except for investors in the household sector, for which data are aggregated according to the investor's country. This means that data considered relevant for statistical analysis are collected, validated and stored for each security, each issuer and each investor. The existence of a reference database with individual information on securities and issuers allows statistical information to be collected from reporting entities on an s-b-s basis. This reduces reporting costs, since there is no need for reporters to aggregate background information according to multiple criteria. Furthermore, it facilitates better information monitoring and greater flexibility in exploring data and performing statistical analysis.

SIET includes two types of information: information on securities issues, and information on holders of securities. The issues side of the system collects information on securities issued by entities resident in Portugal, whether the securities are issued in the Portuguese market or in foreign markets. The portfolios side of the system collects information on the securities portfolios of the country's resident investors, as well as on Portuguese securities in the portfolios of non-resident investors. In addition, information on foreign securities held by resident investors is obtained from commercial databases.

SIET replaced a relatively extensive set of procedures for the collection of securities data, while expanding reporting to all investing sectors and enhancing quality control. While, on the issues side, procedures for researching and collecting information underwent some evolution, reporting on holders of securities was totally changed. Whereas previous reporting on securities portfolios covered only external transactions and operations conducted by monetary financial institutions (MFIs), today all sectors are covered.

Data on issues are collected from several sources. Under the provisions of Article 13 of the Banco de Portugal charter, and of Article 6 of the European Central Bank Guideline of 6 February 2003 (ECB/2003/2) on monetary and financial statistics, the Banco de Portugal is required to ensure the production of securities statistics covering issues by entities resident in Portugal, and the Bank may require any public or private entity to directly provide information deemed necessary for the compilation of these statistics.

On the portfolios side, information is reported according to a Banco de Portugal Directive on "Securities Statistics – Transactions and Positions". Based on this legal provision, detailed information is collected on investments by residents in domestic and foreign securities, and on investments by non-residents in domestic securities.

Figure 1 illustrates the architecture of the SIET system, which relies on two relational databases and one analytical database. Data are stored and validated (first level of quality control) in the "transactional database", which also includes estimates of missing data. Validated and "enriched" data are copied daily to the "exploration database". A second level of quality control is performed on aggregate data, by means of consistency tests and comparisons with other information sources. Statistical outputs are generated from the "exploration database" as well as from the "multidimensional database". This recently developed analytical database is a powerful tool, since it enables user-friendly multidimensional analysis of the information.

Figure 1

SIET architecture
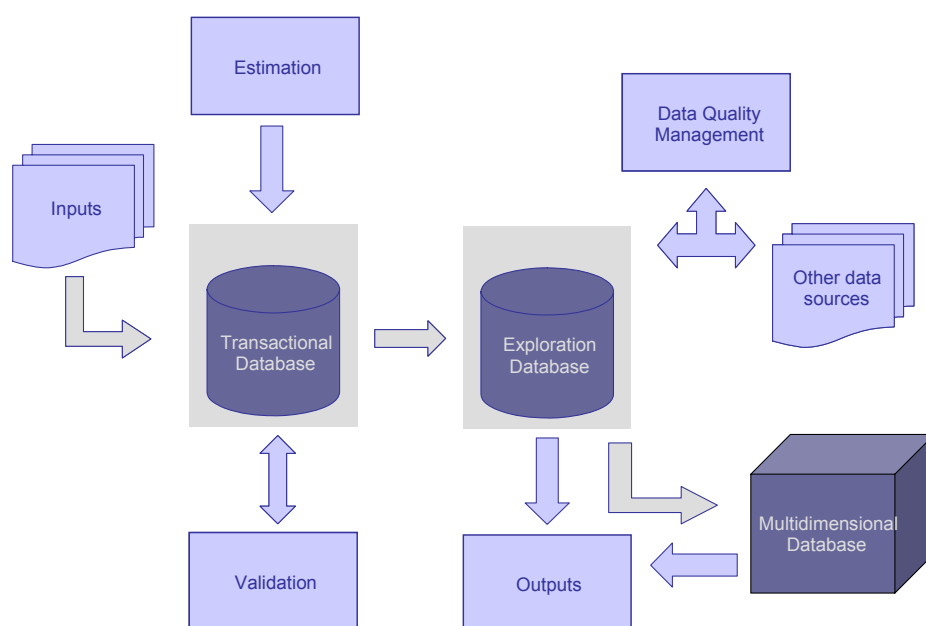
SIET stores information on type of instrument, institutional sector and residency of issuer/investor, prices (quotations), transactions and positions associated with securities issues (issues, redemptions and outstanding amounts), and transactions and positions associated with securities portfolios (purchases, sales, stocks). Classification of securities and entities follows the European System of National and Regional Accounts (ESA/95), complemented by Annex XIX of Guideline ECB/2003/2. Ideally, securities are identified by ISIN (International Securities Identification Number) code, and resident issuers/investors by NPC (the Portuguese acronym for Legal Person Identification Number). Standard unique identification codes are fundamental for sharing and integrating information from different sources. The ISIN code accomplishes this purpose for securities. Unfortunately, there is no corresponding code at an international level for entities.

3. Opportunities and challenges

SIET was designed to meet all of the Banco de Portugal's needs with regard to statistical information on securities. Developing and implementing the system was a major challenge, and the outcome has been very positive.

During the preparation of the project, reporters were contacted and the new reporting scheme was discussed. One might think that asking reporters to transmit individual information on transactions and positions would represent a burden. However, most reporters were also developing their own information systems, and sending individual information was easier and less expensive than aggregating it according to different statistical criteria.

Compilers at the Statistics Department were required to deal with vastly increased amounts of data. Gradually, methods were developed to rapidly identify errors and discrepancies. As in all new systems, there was a learning curve.

Having an integrated system with individual data provides a number of advantages from the perspective of statistical compilation:

- Classification of information for statistical purposes is handled by statistical experts within a common methodological framework.

- Calculations are performed locally according to internally defined algorithms.

- Valuation adjustments follow uniform criteria.

- Consistency between transactions and positions is monitored in detail.

- Issued and held outstanding amounts are compared at the security level.

- Outputs are compiled according to multiple criteria without reporters being required to do so.

SIET promotes consistency across the statistics produced by the Banco de Portugal, and securities issues statistics are a SIET output. In addition to analysis and release at the national level, these statistics are reported to international organisations – namely the ECB and the Bank for International Settlements (BIS). Moreover, SIET's information is used as input for the compilation of a broad set of statistics produced at the Banco de Portugal, which are also disseminated at the national and international levels:

- In the field of monetary and financial statistics, SIET enables the validation and detailing of the information reported in MFIs' and non-monetary financial institutions' balance sheets, as regards their securities issues and own portfolios. The data are also used to produce investment fund statistics.

- In the area of balance of payments and international investment position statistics, SIET data on external transactions and securities positions are used to generate the portfolio investment item (assets and liabilities).

- Public finance statistics are produced using SIET information on issues of the various general government sub-sectors, and on the share of these securities represented by general government entities (for data consolidation purposes).

- Values reported to the central balance sheet data office on the activity of non-financial corporations are also compared with SIET data for quality control of the information collected by both systems on securities issued and purchased by these entities.

- Finally, for the national financial accounts, SIET data are used to compile assets and liabilities items regarding debt securities and equities, broken down by type of security and maturity, for the economy's various institutional sectors.

SIET is an open system, in the sense that new components may be developed and integrated. The component involved in estimating missing information is currently being enhanced, and in the near future, information on foreign securities will be obtained from the CSDB (see below), taking advantage of improvements in the quality of this information.

In our experience, developing an integrated system for securities statistics led to major improvements and opportunities in this area. Representing more than a gradual change in the production of statistics, it constitutes a very important change in the way things are done, and in that sense could be called a revolution.

4.   The near future

Compiling statistical information on securities entails a number of difficulties, both in terms of classification and valuation and as regards identifying holders of securities. The problems are not due to lack of information sources, since several commercial databases provide information on individual securities, and several ESCB central banks maintain their own

databases. In some cases, however, there are gaps, and in others information from different sources is inconsistent. These are the main reasons the ECB is developing a reference securities database with information on a security-by-security basis – the CSDB.

The purpose of the CSDB is to provide a database with complete, consistent, validated and updated information on all securities relevant to the ESCB's statistical objectives. The existence of a single database should promote consistent results and efficient data collection and compilation. The database draws on information from commercial databases and other sources, including national central banks (NCBs) that maintain s-b-s databases. Data quality management will benefit from the cooperation among ESCB members, the BIS and a number of national statistical institutes.

From a statistical viewpoint, the CSDB serves two purposes: supplying information for the compilation of euro area aggregates (eg securities issues statistics), and supplying reference information on securities and issuers to facilitate the collection of statistical information on an s-b-s basis, improving the production of aggregate statistics.

The CSDB is being developed gradually: in phase 1, completed in May 2005, the ECB implemented the system; in phase 2, currently in progress, mechanisms will be implemented to provide online access and application-to-application communication for NCBs.

Several countries are already collecting information on portfolio investment on an s-b-s basis for the production of balance of payments and international investment position statistics. In the near future, all euro area members will be using this approach, employing the CSDB to classify information on securities. In this way, the CSDB will be a major contributor to achieving more efficient production of harmonised statistics.

The Banco de Portugal has actively collaborated in this project since its beginning. Information on Portuguese securities from SIET is being sent to the ECB on a monthly basis. Also, monthly extracts of the CSDB are being used for data quality checks.

Finally, it should be reiterated that recent changes in financial markets as a result of globalisation and innovation have created new statistical challenges and demands. At the same time, technological change continues to create opportunities for developing more highly integrated systems based on "item-by-item" data. These factors, together with increasing collaboration among institutions nationally and internationally, are definitely contributing to the production of more accurate, reliable and comparable statistics.