

How granular should data be to compile BoP statistics?

Carla Marques¹

Abstract

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

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

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

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

1. How to define micro-data?

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

3. Advantages and disadvantages of micro-data

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

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

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

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

4. Final remarks

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