Key points of the Porto Workshop on Integrated Management of Micro-databases

João Cadete de Matos, Aurel Schubert and Paul Van den Bergh

On 20 and 21 June 2013, the Irving Fischer Committee on Central Bank Statistics (IFC) co-sponsored a Workshop on Integrated Management of Micro-databases with Banco de Portugal and the European Central Bank. The meeting was addressed by Mr Carlos da Silva Costa, Governor of the Banco de Portugal who noted that *"…, as central bankers, our responsibilities in this area have now increased significantly, given the fact that IT systems are providing us with new capabilities to draw on the existing micro-databases and to put their data to good use. … If we want to do macro-supervision, if we want to understand the behaviour of companies and individuals, if we want to realize what is happening as regards financial reality, we need to know how diverse reality is and we need to look further than the aggregate level and try to see what is going on at the level of the different groups of individuals or firms and to look for the reasons that are behind this diversity."* Mr João Cadete de Matos, Director of the Statistics Department at the Banco de Portugal, presided over the Workshop in cooperation with Aurel Schubert, Director General Statistics at the ECB, and Paul Van den Bergh, Head of the IFC Secretariat at the Bank for International Settlements. The workshop was attended by 142 officials from 44 central banks from all regions. A total of 42 papers were presented.

The workshop yielded the following insights:

1. Central banks generate and have access to huge amounts of micro-data. This includes raw data underlying statistical data collections (based on censuses and surveys); administrative data sets such as security-by-security databases, credit registers, central balance sheet registers, payment and settlement systems, trade repositories; and third-party data sources (institutional and commercial). Finally there are Big Data from the Internet.

2. Micro-databases have been around for quite a long time but central bank statisticians have been constrained to fully exploit their potential. IT solutions are now becoming available that facilitate the processing of huge quantities of data not only from the same source but also from separate sources.

3. There are clear benefits in using micro-data, such as:

    *   reducing the respondents' reporting burden,

    *   enhancing data quality and quality control procedures,

    *   increasing the efficiency of the statistical production process,

    *   improving responsiveness to users' requirements, by expanding the range of available statistics and providing higher flexibility and agility, in particular in the context macro-economic analyses and risk-based financial stability and micro-supervisory analyses,

    *   providing an excellent tool for economic analysis and research.

4. There are also various constraints to using micro-data, in particular those dealing with the problems of record matching, data coherence, under-coverage, missing variables, reference data consistency over the various

domains, differences in concepts and statistical units and data updating equally reflected across domains.

5. Business Intelligence tools can be used to integrate micro-databases. Such tools include data warehouses, registries, central repositories or dictionaries; building smart cubes; ETL tools; using aliases; anonymisation of data; digital signatures for accessing data. These are not new but are becoming increasingly important for statistical IT applications.

6. Business Intelligence tools are not by themselves the solution to managing and integrating micro-databases. As important are:

    -   appropriate governance of statistical functions in the central bank,

    -   improving the skill mix of data managers (eg registration of data, data validation),

    -   good management of key common identifiers (eg business registers) and reference data (eg bank identification numbers),

    -   good cooperation between economists, statisticians and IT experts,

    -   appropriate solutions to legal, confidentiality and privacy constraints to facilitate access by users,

    -   use of data standards such as SDMX (statistics), XBRL (accounting), DDI (surveys),

    -   education of users, including through online tools.

7. Maximising the benefits while reducing the costs of using micro-data will take time and should be carried out stepwise. New opportunities will be discovered when integrating micro-data bases, including meeting demands from policy makers and data analysts that were previously difficult to accommodate.

The papers of the workshop are available at http://www.bis.org/ifc and will be compiled into an IFC Bulletin. The findings of the workshop will be reported to BIS Governors as part of the yearly IFC Annual Report.