Thirty years of experience in database management: the BIS Data Bank[1]

Christian Dembiermont[2]

The dream of every policymaker, economic analyst or statistician is to have access to comprehensive, high-quality, well-documented and timely data. In the early 1970s, the Bank for International Settlements (BIS) and its major member central banks recognised the benefit of working together to share statistical data through a common platform, and established a common database known as the BIS Data Bank. This note focuses on the objectives of the BIS Data Bank and its development over time, describing its main characteristics and operational challenges, then outlining potential future developments.

Background of the BIS Data Bank

Established in 1930, the BIS is the world's oldest international financial organisation. It fosters international monetary and financial cooperation and serves as a bank for central banks. It acts as a forum to promote discussion and policy analysis among central banks and within the international financial community. The BIS is a centre for economic and monetary research, and serves as a prime counterparty for central banks in their financial transactions.

One way in which the BIS has fulfilled its mandate is by regularly organising meetings for central bank governors and other senior central bank officials. Preparing these meetings has always required the collection of a substantial amount of data from the participating central banks. The fact that BIS meetings depended on extensive international comparisons made the need for a database even greater. Central banks, for their part, were interested in data on the economies of their partners in order to conduct their own analysis of international economic and financial developments. Central banks therefore agreed to share their national data with each other through a common platform to be operated by the BIS. Economies of scale were expected, as the need for expensive bilateral data exchanges would be reduced. For the BIS itself, the Data Bank was a natural public service to its shareholders as well as a tool to facilitate its own work.

The BIS Data Bank is probably one of the oldest databases in the central banking community. The decision to set up a consolidated Data Bank of macroeconomic data was taken by the governors attending the BIS Board meeting in March 1975. At that time, many of the experts' meetings held at the BIS had their origins in the Group of Ten (G-10) initiative in the 1960s. Not surprisingly, the first countries covered by the BIS Data Bank belonged to this group of ten wealthy industrial nations.

Central banks requested that the data used for the meetings of BIS Governors follow nationally defined methodologies. To accommodate this requirement, a model of decentralised data reporting was selected, in which each central bank reports its own

---

[1] The views expressed in this article are those of the author and do not necessarily reflect the views of the BIS. The author is very grateful to Madeleine Op't Hof, Paul Van den Bergh and Kerry Wood for their comments.

[2] Bank for International Settlements, Monetary and Economic Department, Centralbahnstrasse, 2, CH-4002 Basel, Switzerland. E-mail: Christian.dembiermont@bis.org.

national data. Confidentiality requirements for certain statistical information also had to be taken into account (for some unpublished series, as well as for data made available prior to official release). Therefore, central banks decided to limit access to their common database to the central bank community. Accordingly, the objective of the BIS Data Bank is to provide a common database that is updated and controlled by the central bank community. This goal is still relevant today, despite the presence of competitive private databases.

Coverage

Currently, 41 central banks associated with the world's main economies report macroeconomic data to the BIS Data Bank. The selection of series to be reported is based on the principle of "representative series". Thus, central banks report those series used by their economists in analysing their national economies.

Over time, several additional data sets have been added to the macroeconomic data set in the BIS Data Bank. In the 1970s, the BIS international financial statistics were the second set of statistics to be developed in response to the information requirements of the various G-10 committees located in Basel. The set comprises international banking statistics, international and domestic debt securities statistics, and foreign exchange and derivatives statistics. From the outset, the G-10 committees had aimed to reach out to the financial and academic world, and to improve the dissemination of financial information outside the central bank community. Therefore, a substantial part of the BIS international financial statistics data set is publicly available on the BIS website. However, unpublished bilateral positions are available only to central banks through the BIS Data Bank. Recently, the payment and settlement systems data set became the third group of series to enrich the BIS Data Bank. Since 1980, these series have been collected by the Basel-based Committee on Payment and Settlement Systems and published in its so-called Red Book. This information is now available and downloadable at the BIS Data Bank.

Characteristics

Participation in the BIS Data Bank has always been voluntary. At its inception, the organisational agreements between the 11 participating central banks were rather informal. One founding rule was the reciprocity principle: a central bank that made its national statistics available to users via the BIS Data Bank received in return access to the information reported by the other participants. Each member was, and still is, free to withdraw from the Data Bank, and no legal arrangement made reporting to the Data Bank mandatory. Participants were expected to organise regular and timely reporting of time series on a best efforts basis. They set up a Data Bank Experts group to oversee the governance aspects of the Data Bank. This group meets regularly to decide on future developments as related to content and technical organisation.

Due to the broadening of country coverage, the need for a more formal agreement increased, though the basic principles upon which the BIS Data Bank had been built remained unchanged. This led, in 2005, to the adoption of a BIS Data Bank Policy by all members. The policy set forth clearly the common understandings and obligations of all parties involved in the Data Bank.

Given that the objective of the BIS Data Bank is to provide central bank users with high-quality statistics, its coverage is naturally driven by users' needs, under the supervision of the Data Bank Experts group. In an ever-changing world economy, the macroeconomic variables that economists use to analyse a country's conjunctural profile vary over time. Therefore, the

content of the Data Bank is constantly evolving. Data Bank users are encouraged to request any macroeconomic data needed for their analyses but not yet included in the Data Bank, as well as additional methodological information on existing series.

The BIS Data Bank benefits from having a consistent code structure in its macroeconomic data set. This enables users to navigate easily across the database in order to find the information they are seeking. Since the code structure is flexible, new series can be added individually to the BIS Data Bank. This feature gives it a strong advantage over other databases, which usually cover a pre-defined set of indicators and cannot easily be extended on a series-by-series basis. The flexible coding structure is also used to identify discrepancies in the definitions countries employ. Although several methodological manuals have been issued by international institutions in recent years, numerous statistical topics do not yet have the benefit of standardised definitions. Therefore, the BIS Data Bank provides its users the nationally defined series, and underlines methodological discrepancies through differences in the series codes.

Another interesting feature of the BIS Data Bank's macroeconomic data set is the fact that it includes a broad range of frequencies – from annual to daily. International public databases generally do not include daily or weekly frequencies. The numerous daily data transmissions to the BIS Data Bank allow smooth reporting of daily series. Participants consider it good practice to report these within 24 hours.

Challenges

Since its creation, the BIS Data Bank has faced various challenges.

Data reliability

Given the fact that 41 countries report high-frequency data to the BIS Data Bank, there is a high probability of storing inaccurate data. Continuously verifying the accuracy of data is a prerequisite for a clean database. The BIS Data Bank already benefits from the intensive checks that central banks perform on their own high-quality databases, but additional checks are performed when data are received at the BIS. The BIS Data Bank has a unique quality control feature in its Updates/Revisions (U/R) file. This is an enormous historical file that records all the vintages of each observation. The new and revised observations received during the preceding day and identified in the U/R file are validated automatically on a nightly basis.

Statistical time series are continuously being revised. In order to measure the historical impact of an economic variable on policy decisions, economists look for the vintage value of this variable at the particular point in time when the policy decision was taken. This information (real-time data) can be retrieved from the U/R file, but is not disseminated online.

Timeliness

The timeliness of updates is often the main criterion that determines users' preference for a specific database. In the case of the BIS Data Bank, updates for macroeconomic series are loaded only a few minutes after they have been received from central banks. The period between the time the data are received from central banks and the time that reported series are available to users through the end-user application is therefore minimised.

It is also important to minimise the time lapse on the side of the reporting central banks. To assist central banks in timely reporting of series, the BIS provides them with fortnightly timeliness reports. Each central bank's report contains a list of the series that should have

been updated, but for which updates were not received. It also lists the series for which updates are expected within the next month. This timeliness report is based on the series reporting calendars provided by the central banks. If the calendar is unavailable, the timeliness report is based on past reporting performance.

Motivation of reporting entities

Given that participation is voluntary, the BIS Data Bank needs to remain attractive for participating central banks. In the case of international banking statistics, central banks are very interested in becoming reporting entities, as this gives them the right to access the individual claims and liabilities data of 40 reporting countries (out of a total of 200 countries). This data set is a unique source of information that is highly useful for the compilation of external sector statistics. In contrast to the international banking statistics, the macroeconomic data set does not enjoy a monopoly position. In addition to the public databases maintained by other international institutions, numerous private commercial databases also provide economists with macroeconomic data that are helpful in analysing current developments in national economies. For the macroeconomic data set to be competitive with the other databases, it has to offer other advantages, such as coverage that can be extended on demand, and the above-mentioned emphasis on national characteristics. It also has to be on a par with private databases in terms of timely updating.

As mentioned above, serious efforts are being made to speed up the reporting process. However, the best way to differentiate the BIS Data Bank from other databases is by effectively motivating those responsible for reporting within participating central banks. The faster they increase the Data Bank's coverage in response to user requests, and the more extensively they address methodological questions, the more attractive the Data Bank becomes. Challenges may arise from the fact that the individuals doing the reporting are usually central bank statisticians working on national data, while users are central bank economists concerned with international data. Moreover, the users and those responsible for reporting do not work at the same central banks. The role of the BIS is to bridge this gap between users and those reporting the data, and to make the reporting network aware of the importance of its role in this exercise in international cooperation.

The viability of a "one-stop data shop" in the internet era

With the proliferation of websites that provide increasing volumes of data free of charge, a central hub with the type of data currently stored in the BIS Data Bank might appear an anachronistic vestige of a time when the internet was but a theoretical concept. Several international institutions associated with the Statistical Data and Metadata Exchange (SDMX) initiative are currently promoting a new model of data dissemination, known as the "pull" model. In this model, the data provider makes the data available via the internet, eg by placing a data file in SDMX-ML format on a website, from which the data collector then retrieves the data he/she needs without storing any data either in his/her own database or in a central hub. The pull model would replace the traditional push model in which the data provider actively sends data to the data-collecting party. The pull model appears very promising, but several factors may still justify the existence of a central hub based on the push model. First, the technology is not yet reliable enough to allow automatic downloading of series across several websites without interruption. Secondly, a definitive classification of internationally homogeneous statistical domains is not yet in place, and a central hub is still needed to organise national data that have been compiled under differing national definitions. Moreover, not all data can be made available on a public website. Some series are confidential, and some confidential metadata might be associated with public data.

Consequently, a "one-stop data shop" providing a local copy of data reported by different countries will remain a viable model for several years to come. The attractiveness of the

central hub is enhanced by a homogeneous code structure for different countries, centralised timeliness and data quality control, and the utility of the BIS Data Bank's help desk function.

Future of the BIS Data Bank

The BIS Data Bank has always aimed to offer users as much historical information as possible. Discontinued time series – ie historical time series that are no longer updated – are maintained in the BIS Data Bank and are available to all users, though only under special codes. The construction of "long series" by linking discontinued series with current series is left to users' discretion. Some users, however, no longer favour this approach, since they increasingly need "ready-made" long series to save work time. In response to this new demand, the BIS Data Bank will gradually be adding long series to the traditional discontinued series.

From a technical perspective, the pull model clearly represents an interesting solution and will be explored further by BIS Data Bank staff. Once internationally defined statistical domains are in place, SDMX technology will offer an opportunity to reduce overlapping storage of data in different institutions.

Conclusion

Over more than 30 years, the BIS Data Bank has provided statistical support to the central bank meetings held at the BIS under the original cooperative model. The objective of providing a central hub of data, updated and controlled by the central banks, remains relevant. Its content has been adjusted to reflect the Bank's move to a more global reach, an increased focus on financial and systemic risks, and the changing needs of users. Technical developments have also been substantial, and the potential move to a decentralised model is certainly the next challenge to be faced.