Statistical data and metadata eXchange (SDMX): a global standard for quality statistics

Jimmy McHugh[1],
Statistics Department, International Monetary Fund

Abstract

The International Monetary Fund (IMF) has a continuous and growing need for reliable, timely, and high quality data. These data are required for both macroeconomic surveillance – a key mandate of the IMF – and for effective capacity development (technical assistance and training). This paper outlines the IMF's new data collection approach. It highlights the role of its Data Dissemination Standards initiatives, which seeks to improve the quality, timeliness, and comparability of macroeconomic data. The paper also explains how the adoption of Statistical Data and Metadata eXchange (SDMX) complements these data improvement efforts and why the IMF will place SDMX at the center of its future data collections strategy.

Introduction

Data play a central role in the work of the IMF. They are required for effective macroeconomic surveillance, the development of crisis warning indicators, and to motivate and inform the institution's work on statistical capacity development among member countries.

The IMF's need for reliable, timely, and high quality data continues to grow. The financial crises of the mid–1990s and late–2000s highlighted substantial gaps in macroeconomic and financial data. These gaps obscured mounting economic and financial sector imbalances and therefore hindered an adequate policy response. In the aftermath of the crises, the Fund now places a greater emphasis on multilateral and financial sector surveillance. In turn, this has provided a powerful demand for additional data and, more generally, an incentive for improving data quality and timeliness.

This growing demand for more timely and better data poses a major challenge for the IMF. To fulfill its mandate of promoting global macroeconomic stability and fostering economic growth, the IMF needs to collect large amounts of relevant macroeconomic and financial data from all its members. With a membership of 189 countries, with widely varying degrees of statistical practices and capacity, the undertaking is quite challenging.

Member countries have, for the most part, kept up with the growing demand for high quality data. At the same time, member countries are confronted with several international and regional organizations asking for similar data using many different report forms. Therefore, the international community has started to work together to coordinate their data collection

[1] Helpful comments were received from Serkan Arslanalp, Saurabh Gupta, Andreas Hake, Gareth McGuinness, Jolanta Stefanska, and Florina Tanase. Errors remain the responsibility of the author.

strategies to reduce reporting burdens, particularly on low income countries with limited resources and low statistical capacity.

This paper describes how the IMF is reforming the way it collects data and how the institution sees the role of SDMX in addressing its data collection challenges. It reviews various long-standing IMF initiatives to encourage member countries to disseminate high quality and internationally comparable data. However, the nature of the data collection challenge continues to evolve. Many countries are reforming their data dissemination policies, increasing data transparency and now using new technologies to transmit data to outside organizations. In response to these changes, the IMF will make much greater use of machine-to-machine data collection technologies. In this regard, the IMF sees the SDMX standard as a key component in its future data collection strategy.

Data Comparability

The IMF faces a trade-off when it collects data from its member countries. It gathers data against a backdrop of widely differing statistical capacities. Some countries struggle to produce data, which often raises issues of accuracy, availability and timeliness.

At the same time, the Fund wants to gather internationally comparable data to make meaningful cross country comparisons. However, when the IMF is engaged in bilateral surveillance, data are often reported using national statistical standards, making cross-country assessments difficult. In contrast, multilateral surveillance is based on the idea of data comparability across countries. Is the data under consideration homogeneous (in terms of statistical definition) across countries? The same macroeconomic concepts should be defined and measured in the same way everywhere. For example, when making international comparisons of external vulnerability, it is important that concepts such as current account balances have the same meaning across members.

Through its work on methodology and capacity development, the Fund has made noteworthy progress towards data harmonization. It has produced, in collaboration with other stakeholders, a comprehensive library of statistical manuals, covering all the key macroeconomic sectors. The idea of comparability is also reflected in the databases maintained by the IMF Statistics Department (STA), which are based on common methodological standards. This work on harmonization has brought substantive results. For example, 138 economies currently report monetary and financial statistics using standardized report forms (Independent Evaluation Office, 2016).

The IMF's Data Collection Approach

The Fund relies on a voluntary and collaborative approach to data collection. Under Article IV of the IMF's Articles of Agreement, member states are required to "provide the Fund with the information necessary for surveillance". This is followed up with a more specific but extremely truncated list of data requirements in Article VIII. These articles oblige members to provide information on a small set of headline macroeconomic variables such as reserves, prices, exchange rates and national income. In practice, member countries provide - on voluntary basis - far more macroeconomic data than is required by the Articles of Agreement.

To address the issue of different statistical capacities, the Fund has adopted a flexible four-tiered approach.

- The Statistics Department collects data for large cross-country databases – such as the Government Finance Statistics (GFS), the Balance of Payments and International Investment Position (BOP and IIP) and Monetary and Financial Statistics (MFS). The IMF's Coordinated Portfolio Investment Survey and Coordinated Direct Investment Survey are two world-wide databases maintained by the IMF. These databases have strong methodological underpinnings. The data are validated and disseminated in a series of yearbooks and through online databases.
- IMF Area Departments also collect data from member countries to conduct bilateral surveillance. This data tends to be a mix of internationally comparable data and data based on national reporting standards.
- More recently, the IMF has developed data sharing arrangements with other multilateral institutions. These arrangements assign responsibilities for data-related issues, and reduce overlapping data requests for member countries.
- Bilateral surveillance is often complemented by using commercial data sources, which have proliferated in recent years. This trend is most evident with the IMF's bilateral surveillance of advanced economies, where data are required on financial sector issues and which is not readily available from official sources.

The first three elements of this tiered approach depend heavily on various national statistical authorities completing standardized report forms. However, many statistical agencies have indicated that they would like to move away from this resource intensive form of data reporting and rely on new machine-to-machine dissemination technologies.

Making Data More Available and Comparable – the IMF Data Dissemination Standards

To address the issue of data comparability, the IMF has, for several years, worked to develop a series of Data Dissemination Standards.[2] These initiatives aim to increase the availability and comparability of data, thus contributing to the implementation of sound economic policies and improved functioning of financial markets. The Special Data Dissemination Standard (SDDS) and General Data Dissemination System (GDDS) (see box 1) date from the 1994 Mexican crisis, when there was a wide recognition that data limitations had prevented the identification of growing macroeconomic imbalances. The Asian crisis of the late 1990s provided a further stimulus for improved data standards, especially greater data transparency and more widespread data dissemination. Most recently, the global financial crisis of 2007–08 has provided further motivation for improved data quality through the launch of the G-20 Data Gaps Initiative in 2009[3] and of the SDDS Plus in 2012.

[2] For details see http://dsbb.imf.org

[3] For details see http://www.imf.org/external/ns/cs.aspx?id=290.

A key feature of the SDDS Plus initiative is the web-based “National Summary Data Page” (NSDP) that the IMF uses to monitor compliance. These NSDPs reduce the reporting burden of adhering countries, by providing a link to files that contains time series for the prescribed components for the data category. Furthermore, countries are adapting their data production systems to automatically generate the required files in a machine-readable Statistical Data and Metadata eXchange (SDMX) format.

The e-GDDS – improving data dissemination in low income countries

In 2015, to support technically less advanced countries, the IMF launched the e-GDDS initiative. As the acronym implies, e-GDDS is an enhancement of the earlier GDDS and has three broad objectives:

- Foster improvements in data quality;
- Provide a framework for evaluating needs for data improvement and setting capacity development priorities;
- Establish a framework for member countries to disseminate comprehensive, timely, accessible, and reliable economic and financial data.

To achieve the first objective, the initiative encourages countries to share their data through NSDPs. The data are presented in a machine-readable format on the website that allows outside organizations, including the IMF, to pick up the data and populate internal databases. Over time, this initiative will allow countries to provide data without submitting resource intensive statistical report forms to multilateral agencies.

Box 1 - The IMF Data Dissemination Standards

The IMF has developed a series of data standards that are tailored to the statistical capacities of its member countries. They aim to improve the availability of timely data and ultimately, contribute to sound macro-economic policies and improved functioning of financial markets. Specifically:

- **GDDS** - The General Data Dissemination System is aimed at countries with relatively limited statistical capacity. It encourages countries to improve data quality, and identify data weaknesses, which in turn informs the IMF's work on capacity development. Its first objective is to describe the type of data that a member country produces. It offers recommendations on best practice for the production and distribution of statistics. It emphasizes sound methodological principles, timeliness, good compilation practices, professionalism and objectivity. For countries, participation in GDDS is voluntary and it imposes no requirements on the provision of data to the IMF.

- **SDDS** - The Special Data Dissemination Standard (SDDS) guides countries that have or might pursue access to international capital markets in the provision of data to the public. It has more demanding requirements than GDDS. Subscriber countries make commitments with respect to the coverage, periodicity, and timeliness of their data. They also disseminate advance release calendars. Subscriber countries also promise to follow good practice with respect to the integrity, public access, and quality of the data. Subscribers also commitment to maintaining a national summary data page that contains the actual data.

- **SDDS Plus** - This initiative builds on the SDDS, by focusing on countries with systematically important financial sectors, and which are interconnected through financial channels such as interbank and derivative markets. An SDDS Plus adherent must observe additional requirements for nine data categories,[1] in addition to being an SDDS subscriber in full observance of all its requirements. These nine data categories cover areas such as sectoral balance sheets, general government operations, and financial soundness indicators.

Data exchange with other multilateral organizations

Multilateral organizations recognize the growing data reporting burden they place on their membership. To address this concern, the G-20 Inter-Agency Group on Economic and

Financial Statistics (IAG)[4] has established the Task Force on International Data Cooperation (TFIDC). This interagency committee has examined options for streamlining data reporting using the SDMX standard, which will permit the exchange of machine readable data across institution (IMF BOPCOM, 2014).

The IMF has also engaged in data sharing and dissemination partnerships with Regional Financial Institutions. For example, the IMF is building on the SDMX-enabled Open Data Platform for Africa developed by the African Development Bank. Another example relates to the Consumer Price Index data that are collected and shared in close collaboration between the Eurostat, IMF, International Labor Organization, and the Organization for Economic Co-operation and Development (OECD). Such data sharing arrangements have avoided the unnecessary duplication of collection efforts and therefore reduced the data reporting burden on countries.

Why SDMX?

STA has set an objective goal that by 2020 all its data collection activities will use a "machine-to-machine" technology. It sees SDMX as playing a pivotal role in achieving this objective. SDMX allows countries to post their data onto a single web-based repository – the NSDP– and therefore making it available to everyone. All institutions – public and private – can draw on these data as needed. Therefore, SDMX has the potential to greatly reduce the data reporting burden for member countries and facilitate a much timelier provision of data.

SDMX5 is a collaborative initiative sponsored by major multilateral institutions such as the BIS, ECB, Eurostat, IMF, OECD, United Nations, and the World Bank. Moreover, it has garnered the support of several national statistical offices. The IMF sees this widespread acceptance of the standard as one of its most important advantages.

The SDMX standard allows an efficient exchange of data and metadata. At its essence, SDMX is a content standard, embedded in the XML and JSON markup languages. This flexibility helps to provide a common coding standard for content across all institutions making data sharing more efficient, and generally reducing collection costs. It establishes common technical and statistical standards that facilitate the production, dissemination, retrieval and processing of information across data management platforms. As such, it complements the Fund's efforts at establishing data standards to improve data comparability.

The IMF has strongly encouraged its member countries to use SDMX in their NSDPs. SDMX is the preferred standard used in the e-GDDS and SDDS Plus initiatives. So far, 12 countries are now distributing data via e-GDDS, and a similar number of advanced countries are providing data through SDDS Plus.

[4] The IAG members are the Bank for International Settlements (BIS), the European Central Bank (ECB), Eurostat, the IMF (chair), OECD, United Nations, and the World Bank.

[5] In 2013 SDMX was approved by ISO as an International Standard (ISO 17369:2013).

SDMX will enable the Fund to further modernize its data collection strategy. It is now transitioning from a data collection strategy based on standardized report forms to one that is based on data harvesting from the NSDPs.

Concluding Remarks

Reliable, accurate, and timely data is an essential input for the work of the IMF. To strengthen the quality of the data it receives, the IMF's data collection strategy has emphasized the importance of Data Dissemination Standards, which seeks to develop accessible and comparable data.

Going forward, the IMF plans to shift to a machine-to-machine collection strategy. It plans to use SDMX as the main data sharing standard underpinning its new collections strategy.

References

Independent Evaluation Office of the International Monetary Fund (IEO), 2016 "Behind the Scenes with Data at the IMF An IEO Evaluation", Washington: International Monetary Fund.

IMF BOPCOM (2014) "Twenty-Seventh Meeting of the IMF Committee on Balance of Payments Statistics" Washington, D.C. October 27–29, 2014.