A model data producer: the importance of sound metadata management – Botswana's case

Sabata Legwaila[1]

Introduction

The availability of quality statistics is a prerequisite for the formulation of appropriate macroeconomic and financial policies. Therefore, the importance of statistics cannot be overemphasised. Activities throughout the modern economy, in the government and private business sectors, as well as other spheres of the broad economy, increasingly require up-to-date and accurate statistics, especially given that credibility is the key to policy effectiveness. Like all modern and up-and-coming economies, Botswana has come to recognise the importance of quality and timely and transparent statistics, the production of which is guided by internationally accepted standards.

Because of demand for quality statistics by users, statistics producers now emphasise producing timely, internationally comparable and transparent statistics. For there to be transparency, there is a need to document the processes of data production through the use of metadata. Metadata are simply a description of how the data were produced, be it the production processes, description of concepts and terminology, classification and standards, statistical methods and software used, as well as the creators of the data.

One of Botswana's statistical objectives has been to produce timely, accurate and transparent statistics. This has taken the form of the country taking an active part in international data improvement initiatives, in particular that of the International Monetary Fund (IMF), through the United Kingdom Department for International Development (DFID)'s financial assistance. Such efforts have helped give a broad range of official statistics in Botswana, compiled across several agencies, backbone and discipline towards improvement efforts, through a focus on metadata issues.

The purpose of this paper is to outline Botswana's experience of attaining modern data producer standards through metadata management. In addition, the paper also discusses the benefits that the country has achieved with reference to the aforementioned statistics, and the way forward for the production of statistics in Botswana, especially in utilising modern data and metadata management systems. The main focus of the paper is on the country's experience with the production of monetary and balance of payments statistics, as this is where the author has the most experience. But this is not to downplay the effort made in other areas to overcome the many and diverse challenges.

An overview of models of modern data and metadata systems

Data are public goods; hence their production cannot be limited to use by nationals. With the world's economies operating in a global arena, the need for quality data is at the forefront, to enable timely and accurate policy decisions, especially with the advent of recessions and

---

[1] Bank of Botswana, Research Department, Plot 17938, Khama Crescent, Gaborone, Botswana. E-mail: Legwailas@bob.bw

financial turmoil. Ultimately, the value of statistics depends upon their quality[2]. The producers of statistics and the methods and standards employed in statistics production must have credibility. With the advent of technology and global competition, it is becoming more apparent that national, international, and transnational data are needed for policy formulation and evidence-based decision-making at national and international levels. As such, national data enter the international statistical systems, where standardisation is necessary to produce consistent, international datasets. Hence, models of national data are, to a large extent, influenced by international data systems and models.

The international community, such as the IMF and the World Bank, therefore plays an important role in statistical development by promoting and implementing internationally agreed standards, methods and frameworks for statistical activities. Since the international community has a stake in national statistical systems, it ensures that statistics are compiled using common methods which enable international comparisons of statistics and enhance coherence of statistics across subject areas through the use of standardised concepts and definitions. By providing a common basis for comparison, international guidelines make statistics more useful and offer more transparent reporting of outcomes. They also help to establish benchmarks or standards against which national statistical systems can measure their performance and through which they can strive to reach even higher standards. This has led statistics producers to adopt models of data to meet modern data needs, and to put in place metadata systems that not only enhance transparency, but also increase ease of usability.

Metadata are essential in all stages of the statistical production process in that they assist the user in understanding the data and all the processes involved in their production. Users can utilise sources of information more effectively if they know the definition of concepts and classifications behind the data and how they have been produced. There is, therefore, a wide variety of metadata systems being developed by specialists in various fields, some simple and others more sophisticated (Dragon, 1995–2000) in relation to their compilation and management[3].

A quest for quality in statistics: Botswana's case

Background to statistical developments in Botswana

To be effective, statistical systems must be backed up by legislation that provides both safeguards of confidentiality for the providers of raw data and assurances of integrity and accessibility for users. The nature and organisation of national statistical agencies vary from country to country. In decentralised systems, separate agencies have independent mandates to compile and disseminate statistics in particular areas. But even in highly centralised systems, responsibilities may be divided. Central banks, for example, are usually mandated with the collection of data on money and banking and other areas such as the balance of payments. Thus, a national statistical system is often a network of agencies whose activities are usually coordinated by legislation, administrative practices and professional standards, which is essential for avoiding discrepancies in data.

In Botswana, the main agency responsible for producing national statistics is the Central Statistics Office ((CSO), which is due to be transformed from a government department into an autonomous agency, the National Statistics Office). However, the Bank of Botswana has

---

[2]   See World Bank (2002).

[3]   See T J Lukhwareni et al (2005).

responsibility for balance of payments and monetary statistics, while the Ministry of Finance and Development Planning provides government finance statistics. Necessary intra-agency coordination is provided through various working groups and a high-level Statistics Producers Committee (SPC). The overall statistical objective is to promote good governance in statistics, which entails transparency and accountability, and hence leads to well-guided policy formulation and decision-making by stakeholders, including government, households and businesses.

While, through the CSO, Botswana has for many years enjoyed a reputation for the production of relatively reliable statistics, the quest for a coordinated approach to adopting an overall framework of good statistical processes started in 2001 with the preparation by IMF staff of a data Report on the Observance of Standards and Codes (ROSC)[4]. The ROSC carried out in Botswana assessed statistics covered by the General Data Dissemination System (GDDS), namely real sector, prices, government finance, monetary and balance of payments, against benchmarks set out in the IMF's Data Quality Assessment Framework (DQAF). The report noted that Botswana's statistics were of generally good quality, but that there were shortcomings in terms of both compilation and, in particular, periodicity and timeliness of dissemination. To help address these, since 2001 Botswana has participated in the GDDS Project for Anglophone Africa, through which technical assistance is provided jointly by the IMF and the World Bank, and funded by the DFID. Initially, the project was intended to last for three years, but was subsequently extended to a second phase, due to end in October 2009; a third phase is then scheduled to commence. While the length of the project indicates the extent of the work that is needed, the continuing financial support from the DFID is itself a clear indication of the positive impact across the range of countries covered by the project.

The following major recommendations were made under the various statistical areas:

- Monetary statistics: expansion of data coverage to include other depository corporations that were not covered in the then monetary survey, in order to produce a full depository corporations survey; proper application of the concept of residency by banks; banks to submit data electronically; and there was a need for the development of statistical database software;

- Balance of payments: improving periodicity by moving to quarterly as opposed to annual data as well as improving coverage and/or methodology for various balance of payments items, most importantly trade in services;

- Government finance statistics: adopting the 2001 GFS Manual, timely dissemination of budgetary central government accounts as well as compilation of consolidated data of general government (both local and central government);

- Real sector: acceleration of work programme for implementing the System of National Accounts (SNA) 1993 with respect to national accounts and producing and disseminating the producer price index , as well as regular updates of the consumer price index (CPI) weights, every five years.

With respect to documentation, prior to 2001, at the Bank of Botswana, except for internal working documents, there was no publicly available documentation on how the data were produced. Moreover, while the data were disseminated on a timely and regular basis (the Bank's *Botswana Financial Statistics* publication is produced on a monthly basis, and the annual report contains an extensive range of statistical tables; both are available on the Bank's website), there was no publicly available metadata documentation in this respect. Thus, apart from experience, data users had no well-based expectation about when they

---

[4]   See Appendix 1 for a summary of the mission.

could expect to receive data updates, what the policy is regarding revisions, or confidence that other users did not have more favourable access.

Implementing plans for improvement

Participation in the anglophone GDDS Project has led to the development of metadata outlining appropriately prioritised plans for improvement, which were first published in the IMF Dissemination Standards Bulletin Board (DSBB) in August 2002[5]. To initiate the actual implementaion of plans for improvements, the country's authorities requested technical assistance from the IMF through the GDDS Project, both under phase 1 and phase 2, which adopted a modular approach. The main purposes of the technical assistance (TA) missions were to assist with implementation of plans for improvements in specific areas.

With regard to monetary statistics, the first TA mission took place in 2003. During the follow-up missions in August 2004 and 2005, the actual expansion of monetary data took place. The monetary statistics framework recommended by the IMF's Monetary and Financial Statistics Manual (MFSM) was adopted and data were published using standardised report forms. The missions also addressed, among others, issues of instrument classification and proper definition of institutional units as laid out in the monetary and financial statistics framework recommended by the MFSM and the MFS Compilation Guide.

With regard to BoP, the major objective was to compile and publish the quarterly balance of payments accounts, which was achieved in January 2008. The challenge remains to investigate the large errors and omissions, an exercise that started with a technical assistance mission during February 2009.

Efforts over the last few years have focused on ensuring that the monetary and financial statistics and the balance of payments statistics at the Bank were in conformity with the 2000 Monetary and Financial Statistics Manual (MFSM) and the 1993 Balance of Payments Manual, fifth edition (BPM5), which provide practical guidance for compiling monetary and financial statistics and balance of payments statistics, respectively, as well as dissemination issues. Because data improvements are an ongoing process, efforts are still continuing to achieve further improvements, through phase II of the GDDS project.

Achieving data quality through metadata management

As important as it is to document data production processes, it is worth noting that production of metadata is not a one-off activity. It has to be maintained and managed to meet the demands of modern-day data requirements, for ease of use by both users and producers.

As mentioned before, Botswana was one of the very first countries to meet the prerequisites of the GDDS by producing and posting its metadata on the IMF DSBB. This act was followed up with deliberate steps to implement the plans for improvements. Focusing on metadata issues has helped to highlight what needed to be done to effect improvement, and also has added backbone and discipline to Botswana's statistical effort, including improved data quality, dissemination and coordination among agencies and between datasets.

---

[5] Botswana made noteworthy progress in this regard. It is not an exaggeration to say that prior to the ROSC mission, metadata was a largely unknown concept to compilers of statistics in the country. Nevertheless, Botswana was the first country, ahead of the other 14 countries in the project, that subscribed to the GDDS to produce and publish its metadata on the IMF DSBB.

Data quality improvements

By implementing plans for improvements with seriousness, Botswana was able to achieve a remarkable improvement in the quality of its data. The country, in particular the Bank of Botswana, was able to publish survey data for the monetary statistics according to the framework for monetary and financial statistics recommended in the MFSM within three years of subscribing to the GDDS project. With regard to the balance of payments, quarterly balance of payments estimates are now published, although there are some concerns with respect to the accuracy of the data, as indicated by the persistently large errors and omissions.

Improved dissemination

With regard to dissemination, the monetary statistics are submitted on a monthly basis for publishing in the IMF's *International Financial Statistics* (IFS), in the standardised report formats.[6] Data are also published on the Bank of Botswana website on a monthly basis. Metadata are published on the DSBB and currently updated on an annual basis or when there are changes to data. A link to the metadata is provided on the Bank of Botswana website, which is currently under review to make further improvements to data accessibility issues.

Coordination among data producers

Concentrating on effecting improvements in data has also resulted in coordination among data-producing agencies, namely the Bank of Botswana, the CSO and the Ministry of Finance and Development Planning, through forums such as the SPC and a national GDDS committee, which both comprise representatives from the three statistics agencies. This has also resulted in accountability in following through with agreed goals in terms of what data need to be produced and when.

Lessons learnt and challenges

Lessons

Although the Bank has made tremendous progress in improving its data, it has become evident that while the country needs to be vigilant in taking action where statistics improvements are concerned, there is a need to avoid being overambitious. Plans for improvement should set targets, but they should not be so demanding as to be unachievable. Statistics development and improvements are long continuous objectives, which cannot be achieved overnight, such as with regard to the graduation to Special Data Dissemination Standards (SDDS). This involves issues regarding both institutional capacity and policy practices; for example, the issue of timeliness of reporting of monetary data that uses the central bank balance sheet (which has a bearing for the reserve template, a requirement for graduation to SDDS).

It is also worth noting that achieving certain standards can compromise others. For instance, modern data models and data standards, such as the GDDS, emphasise timelines, which risks putting quality in second place (ie get the data out on time whatever the costs). There

---

[6] Botswana was among the first few countries that published data in the IFS supplement, using the standardised report formats.

are also practical issues, especially for countries with limited manpower. For example, cooperation between data producers and users makes sense, but there are risks of violating rules prescribed by the data standards, in particular, privileged access to data by some users. An example is cooperation between the Bank of Botswana and the CSO in producing national accounts estimates; this means involvement of data users, which violates the rule against privileged access for certain users prior to publication.

Regular reviews of progress that come with data standards are an asset to any data producer. For example, the second ROSC mission in 2006 was most useful for both the Bank of Botswana and the IMF because it enabled stocktaking and refocusing, which brought to the fore the realities of how much progress had been and could be expected to be made.

Documenting plans for improvement based on metadata is good for guiding prioritisation for technical assistance. For Botswana, this has promoted clearer prioritisation within and among the three data-producing agencies. This also, in a sense, creates peer pressure for Botswana to benchmark itself against countries that have achieved higher data quality standards, hence instilling discipline in ensuring perseverance in data improvement efforts.

Challenges

The next expected step for GDDS subscribers is to graduate to SDDS. One of the major requirements for subscription to SDDS is to produce and publish the country reserve template. This has been a challenge for Botswana due to confidentiality issues. The authorities are however, studying practices in other countries for guidance on the matter.

Other stumbling blocks emanate from the fact that other statistics producers in the country lag behind in terms of meeting the basic requirements, even for the GDDS, for example, the country's government finance statistics are still produced on the basis of the 1986 GFS manual instead of the latest one. This has been stalled on their part by capacity challenges, and this will in turn hamper the whole process of graduating to SDDS.

On other improvements, such as the project on compiling and disseminating other financial corporations (OFCs) data, there are challenges as the project is a joint effort between the Bank of Botswana and the newly formed Non-Bank Financial Institutions Regulatory Authority (NBFIRA). Due to the infancy of the institution, there are still capacity challenges which will slow the progress on the project.

Recommendations and conclusion

Botswana has gone a long way in improving its statistics, and has done it right the first time by being disciplined in using metadata to focus on issues of data improvements. The coordination among data-producing agencies is commendable, and is something that has to be given momentum going into the future. It will be wise for the country to keep pushing despite the challenges, to develop and improve data further.

Even though there are challenges to issues such as graduating to SDDS, it is worth the effort to continue being involved with international agencies, which have gone the extra mile in helping with the development of the country's statistics in general. Despite the fact that other agencies in the country are lagging behind in some statistics, there have been marked improvements. The country can learn from others in terms of adopting and utilising modern data models, and develop world-class metadata systems to ensure that data production and use are done with ease.

Appendix 1:
Summary on ROSC[7] mission

The ROSC report on Botswana

The IMF ROSC mission commenced in early October 2001 and lasted for two weeks. It focused on the following data categories: national accounts; prices (consumer and producer); government finance; monetary statistics and balance of payments. Neither socio-demographic nor labour market statistics were included at this stage, although the subsequent GDDS project has covered some of these areas through the provision of TA.

The final report of the mission was presented to the government in early December at the same time as the workshop to launch the regional GDDS project, which was held in Gaborone. The report included an assessment on the quality of Botswana statistics together with key short- and longer-term recommendations for improvement. After discussions with the Government and the Bank of Botswana, the report, together with "The Response of the Authorities", was posted on the IMF website. Overall, the quality of Botswana statistics was generally seen as good, and in some cases very good.

For example, the CPI, on which inflation measures are based, was assessed as being of SDDS quality, and the report emphasised that a concerted effort could quickly move Botswana within striking distance of SDDS requirements more generally. The major deficiency in terms of quality was in the area of producer prices, which the CSO candidly admitted was a concern and that assistance would be required to effect necessary improvements.

Another major area where improvements were found to be needed was that of dissemination. This was in terms of both timeliness and ready availability to all users. Some indication of this can be seen from Table 1, which divides the main ROSC recommendations according to sector. Of the 31 recommendations, nine (about 30 per cent) were in the "general" category and, of these, eight dealt with dissemination issues ranging from the need to establish advance release calendars to providing more extensive information on methodology, including on data limitations. (The ninth was a recommendation regarding the need for training.) Six of these recommendations were identified as achievable in the short term.

TABLE 1: BOTSWANA ROSC RECOMMENDATIONS

	General	National Accounts	Prices	Government Finance	Monetary Statistics	Balance of Payments	Total
Short term	6 (2)	1 (1)	2 (1)	2 (1)	4 (2)	2	17
Med term	3	2	2	4	1	1 (1)	13
Long term			1				1
Total	9	3	5	6	5	3	31

Note: figures in brackets are the number of recommendations which were given high priority
Source: Botswana ROSC Report

Some of the sector-specific recommendations also dealt with dissemination issues, notably timeliness. The relatively high number of recommendations for government finance was due in part to the difficulties associated with incorporating local government finances fully into the framework of regular reporting. A further important conclusion was that, in some areas, production of statistics could be facilitated by improved coordination between agencies. The report noted discrepancies in methodology, delays in communicating necessary information (for instance, the balance of payments relies on various inputs from the CSO, such as trade data, while in turn the national accounts require timely balance of payments data), and the potential confusion caused by the various reporting "years" used across the different sectors.

_____

Source: Bank of Botswana *Research Bulletin*.

---

[7]   ROSC refers to the Report on the Observance of Standards and Codes.

References

Bank of Botswana (2003): *Research Bulletin*, vol 21, no 2, December.

Lukhwareni, T J, S F Madonsela, D E Mokhuwa and L M Podile (2005): "Management of metadata in national statistical agency", Data Management and Information Delivery Project, Statistics South Africa, Pretoria, 5–9 September.

Matjaz, J et al: "Using metadata in statistical processing cycle – the production tools perspective".

World Bank (2002): *Building statistical capacity to monitor development progress: information paper prepared for the World Bank Board of Executive Directors*, Washington, DC, October.