BP/16/06

Data and Statistics at the IMF: Quality Assurances for Low-Income Countries

Morten Jerven

IEO Background Paper
Independent Evaluation Office
of the International Monetary Fund

Data and Statistics at the IMF: Quality Assurances for Low-Income Countries

Prepared by Morten Jerven

February 25, 2016

Contents   Page

ABBREVIATIONS

AfDB	African Development Bank
AIV	Article IV consultation report
AUC	African Union Commission
CPI	consumer price index
DQAF	Data Quality Assessment Framework
DSBB	Dissemination Standards Bulletin Board
GDDS	General Data Dissemination System
GDP	gross domestic product
GNI	gross national income
G20	Group of 20
IFS	*International Financial Statistics*
IMF	International Monetary Fund
LICs	low-income countries
LO	practice largely observed
LNO	practice largely not observed
NO	practice not observed
O	practice observed
OECD	Organization for Economic Co-operation and Development
PRGT	Poverty Reduction and Growth Trust
PPI	producer price index
REO	*Regional Economic Outlook*
RES	Research Department
ROSC	Report on Standards and Codes
SCI	Statistical Capacity Indicator
SDDS	Special Data Dissemination Standard
SNA	System of National Accounts
SSA	Sub-Saharan Africa
STA	Statistics Department
UFR	use of Fund resources
UN	United Nations
UNECA	United Nations Economic Commission for Africa
UNSD	United Nations Statistical Division
WB	World Bank
WDI	World Development Indicators
WEO	*World Economic Outlook*

EXECUTIVE SUMMARY

How does the IMF deal with the challenge of obtaining timely, high-quality data for its operational purposes? This paper examines the different ways the IMF performs quality assurances on macroeconomic data for internal and external use, with a focus on how it handles data and metadata on countries that are classified as low income. The paper concentrates on this subset of countries because they face the greatest resource constraints in producing and disseminating the high-quality macroeconomic statistics and metadata needed to fully support the IMF's surveillance and financial programs.

The paper takes up two questions that have been highlighted in previous IMF reviews on data: given that data users often cannot distinguish between IMF data and official country data, does the IMF incur reputational risks by disseminating data that may be (i) of questionable quality or (ii) inconsistent across its various databases and reports?

To explore these questions, the paper first reviews how the IMF collects and disseminates macroeconomic data on low-income countries (LICs). Of the several channels for data collection, two of the most important are: (i) reporting by the country authorities to the IMF's Statistics Department for publication in *International Financial Statistics (IFS)*; and (ii) collection by area department country desks (during missions; through direct contact with country authorities; from resident representatives, online, and commercial sources; via estimates) to form the basis for Article IV consultation reports and the *World Economic Outlook (WEO)* publication. A review of the data from these sources indicates that significant amounts of data are missing in *IFS* and that discrepancies exist among some key data reported in Article IV reports, the *WEO*, and the *IFS*.

One way to lower reputational risk is to provide sufficient qualifying information about the data and statistical systems. The IMF has created three mechanisms to provide such information: (i) the General Data Dissemination System (GDDS) and the Special Data Dissemination Standard (SDDS), both of which focus on provision of detailed metadata; (ii) the data module of Reports on Standards and Codes (ROSCs), which is based on the IMF's Data Quality Assessment Framework and provides a thorough review of a country's statistical system; and (iii) a three-tiered rating system that was put in place for Article IV reports to provide candid assessments of the adequacy of data provision for IMF surveillance.

In practice, each of these mechanisms suffers from serious problems. First, most member countries do not regularly update the GDDS, which is most relevant for LICs, and very few countries have graduated from the GDDS to the more stringent SDDS. Second, no LIC has had a data ROSC since 2009, and only two countries have had a ROSC update. Finally, as highlighted by IMF reviews, Article IV consultation reports show a lack of candor in reporting about the quality of statistics; the present study shows that even countries whose data suffer from multiple serious shortcomings are still receiving broadly adequate ratings.

I. INTRODUCTION

1.   This paper examines the processes by which the IMF collects macroeconomic data from its member countries and seeks to assure the quality of the data it uses and disseminates.[1] The paper concentrates on those countries classified as low income,[2] because these countries may face the greatest resource constraints in producing and disseminating the macroeconomic data and metadata needed to support the IMF's surveillance and lending. But income and statistical capacity are not perfectly correlated: some low-income countries (LICs) punch above their weight, providing high quality data, while others provide less reliable data than one would expect from their income level alone.[3] In the aggregate, however, there are some clear relationships that justify a special focus on LICs.

2.   The IMF plays several roles with respect to data and statistics, including collection (e.g., from member countries or commercial providers), internal use of these data, support to the member countries that produce them (e.g., through technical assistance), and dissemination to external stakeholders (e.g., through its flagship publications).

3.   Over the years, IMF Board papers on statistical issues[4] have often highlighted weaknesses in the ways that data and metadata have been collected, reported, and disseminated. The issues repeatedly raised include the potential reputational risks for the IMF that derive from (i) dissemination of data that may be of questionable quality; and (ii) lack of consistency in the data reported in the Fund's various databases and reports. Data users often cannot distinguish between IMF data and official country data. While concerns about data quality are not unique to data from LICs, they are likely to be exacerbated by the more severe capacity constraints that LICs face.

4.   To explore these issues, Section II of this paper provides a short review of known data quality problems in LICs, and Section III assesses the channels that the IMF uses to collect and disseminate data. Section IV reviews the Fund's formal and informal mechanisms

---

[1] In this paper, the terms "statistics" and "data" are used interchangeably, although in some quarters the distinction is made that statistics refer to official aggregates, such as GDP and measures of inflation, while data encompass a broader range of information. The paper focuses on economic statistics, though the issues of statistical capacity also pertain to demographic and social statistics as well as civil registration and vital statistics. For a review of these issues, see Jerven (2014a and 2014b).

[2] For the purposes of this paper, countries are classified as low income if they are eligible to borrow under the IMF's Poverty Reduction and Growth Trust (PRGT).

[3] See, for example, the World Bank Statistical Capacity Index, which reports on statistical capacity for low- and middle-income countries (World Bank, 2009).

[4] Among the most important of these papers are the regular reviews of data provision. The operational framework for the treatment of data provision to the Fund in Article IV consultations was put in place in 1995 in the wake of the Mexican crisis (IMF, 1995); a further seven reviews were issued through 2012.

for data quality assurances, both formal and informal, in the IMF's statistical structure. Section V provides concluding remarks.

II. THE STATE OF MACROECONOMIC STATISTICS FOR LOW-INCOME COUNTRIES

5.   On November 5, 2010, Ghana Statistical Services announced new and revised GDP estimates. The estimated size of Ghana's economy was adjusted upward by more than 60 percent, suggesting that previous GDP estimates had missed economic activities worth about US$13 billion (Jerven, 2012). Ghana's change in GDP was exceptionally large, but not an isolated case. On April 7, 2014, the Nigerian Bureau of Statistics declared that its GDP estimates were being revised upward to US$510 billion—an 89 percent increase from the old estimate (*The Economist*, 2014). Other LICs, many of them in Africa, are working on revisions to their GDP data,[5] which in some cases have resulted in their reclassification as lower-middle-income countries (IMF, 2013b).

6.   These well-publicized statistical events have increased the attention paid to the quality of macroeconomic data in LICs, especially those in Africa. According to the African Development Bank, the substantial data revisions have "understandably alarm[ed] many observers" (AfDB, 2013), with the World Bank's chief economist for Africa writing of "Africa's Statistical Tragedy" (Devarajan, 2013). And, in an editorial on October 28, 2013, the *Financial Times* wrote:

> Reliable data are sorely needed. The International Monetary Fund has warned that 'the quality of basic economic statistics in sub-Saharan Africa . . . is often so poor that it can lead to serious misdiagnoses. In the past, similar problems have afflicted regions such as Latin America, the former Soviet Union, and South East Asia.'

Another factor increasing the demand for high-quality, reliable data from low-income countries is that after their strong economic growth of recent years, a number of these countries are entering international capital markets.

7.   Persistent doubts about LICs' ability to provide high-quality data may in part be a true reflection of the data, but they may also stem from a perception and credibility problem. The good news is that many statistical systems in LICs are being strengthened after years of relative neglect (Jerven, 2013), and that, in most cases, the updated benchmarks show that the countries are richer than previously thought.

8.   Against this background, the international community of data users, data producers, and data disseminators looks to the IMF for advice, guidance, and leadership in the realm of

---

[5] Of the IMF member countries classified as LICs, more than half are in Africa. Recent reports from Kenya and Zambia signal increases of about 25 percent to 30 percent as benchmark years are updated across the region (Manson, 2014; *Africa Report*, 2014).

macroeconomic data. Country authorities, media, and the public not only rely on the IMF to provide technical assistance to support better data, but also turn to it as an entity that safeguards the validity, accuracy, and even the credibility of these data (Box 1).

Box 1. The IMF and Nigeria's Revisions to National Accounts Estimates

Recent large revisions to GDP estimates raise questions about IMF surveillance and technical assistance. The 2013 Article IV consultation report for Nigeria described the implications of the forthcoming revision:

> The base year for the national accounts is being updated from 1990 to 2010, a period during which the structure of the economy has changed significantly—for example, the share of agriculture in employment has fallen from 70 percent to about 30 percent. A material change in GDP would alter key performance indicators in the macroeconomic framework, including tax, investment spending, and debt ratios. For example, the tax revenue to GDP ratio of 20 percent puts Nigeria in the middle range of low-income countries, but a significant increase in GDP could push Nigeria in the low range of low-income countries. The authorities stressed the importance of credibility for the rebasing of the GDP; consequently, the release of the revised GDP data has been delayed until after technical assistance missions from the Fund, World Bank, and African Development Bank, which started in mid-December. The authorities noted, however, that the expected higher nominal GDP would have little impact on their borrowing strategy, which is based on the ratio of debt service to revenue. (IMF, 2014a)

As the GDP revision was bigger than expected, and the supply and use table was not yet ready in April 2014, the IMF's Statistics Department (STA)—which had provided technical assistance for the revision—viewed the revision as not yet complete, and thus recommended delaying the publication of numbers. STA also stressed that the IMF does not have the authority to endorse or not endorse GDP numbers.

The three agencies providing technical assistance, however, were sympathetic to the authorities' position that further delay could undermine the credibility of the Nigerian Bureau of Statistics. The IMF mission chief spoke to the media after the revision was announced and gave this statement to the press in Abuja, Nigeria:

> I'd like to make three points. First, comprehensive, quality statistics is the foundation of sound decision making. That recognition is embodied as a central objective of the System of National Accounts (SNA), which the Inter-secretariat Working Group on National Accounts (UN, WB, IMF, OECD, and EC) have been updating and refining, over the last 50 years. As of today, we are promoting SNA 2008, and it is important that one of the key improvements noted by the Statistician-General in his presentation was the use of SNA 2008 methodology in compiling the rebased GDP numbers. Further, the CME has underscored the efforts being made to improve the statistics of the Federation, as a basis of sound decision making. Let me state that we endorse this wholeheartedly and will support Nigeria in this regard.

The key statement was "we endorse this [the efforts to improve statistics] wholeheartedly." However, the Internet and blog meme became "the IMF endorsed the numbers."

---

9.   A 2011 survey of the status of GDP statistics in Sub-Saharan Africa (Jerven, 2013) found that only six out of 37 countries were following the IMF's advice to use a base year that was at most five years in the past (i.e., at that time, 2006 or later). Two years later, the IMF's *Regional Economic Outlook* for Sub-Saharan Africa (IMF, 2013b) found similar results: only four out of 45 countries were meeting the five-year rule, while 28 were using base years more than 10 years old and 13 were using base years more than 20 years old.

10.   Adherence to standards for the dissemination of data and their metadata is far weaker in LICs than in the rest of the IMF's membership. Among the LICs considered here, only two subscribe to the IMF's Special Data Dissemination Standard (SDDS), while 67 participate in the less demanding General Data Dissemination System (GDDS). Moreover, although the

GDDS recommends annual reporting of metadata, countries that subscribe to the GDDS often report irregularly and with occasional large gaps (see Section IV below).

11.   In the IMF's flagship statistical publication, *International Financial Statistics (IFS)*, significantly fewer data are available for LICs than for other member countries. Table 1 compares the availability in the *IFS* database (as of February 2015) of data from LICs on two key variables—real GDP growth and inflation—against that from a random sample of 40 non-LIC countries. The differences are striking: whereas more than 30 percent of the data for LICs are missing, non-LICs lack information on just over 3 percent of their data.

Table 1. Missing Data in *IFS* (In percent, as of February 2015)								
	2011			2012			2013		
	LICs	Non-LICs	Total	LICs	Non-LICs	Total	LICs	Non-LICs	Total
Real GDP growth	51.3	10	37.3	55.1	12.5	40.7	70.5	20.0	53.4
CPI	3.8	0	2.5	3.8	0	2.5	3.8	0	2.5
Source: IEO estimates.									

III. THE ROLE OF THE IMF IN COLLECTING AND DISSEMINATING MACROECONOMIC STATISTICS ON LICS

A. Channels of data collection

12.   The IMF collects member country data through two main channels (Figure 1). The first channel is the **direct reporting of official statistics** by IMF member countries to the IMF Statistics Department (STA), which in turn makes them available to the general public through its main publications, principally the *IFS*.[6] The IFS database often serves as a reference point for other data disseminators, such as the World Bank, which collects its inflation and balance of payments statistics from this source. Reportedly, external data users commonly believe that these statistics have the IMF's "seal of approval."

13.   STA relies exclusively on official data. There is a fixed calendar for country authorities' submission of data to STA (usually monthly or quarterly) and, for some variables, STA collects information directly from official websites. STA aims to disseminate data that, as closely as possible, follow common definitions; thus, in some cases, countries compile certain statistics in formats that differ from those they use internally. This affects the timely provision of the information.

---

[6] The other main publications of STA are *Balance of Payments Statistics (BOPS), Direction of Trade Statistics (DOTS), and Government Finance Statistics (GFS).*

14.   STA performs some light quality checks on the submitted data. These involve tests for compliance with established formats, an examination for outliers, and some broad cross-sector consistency checks to capture large discrepancies across data sets.[7] Discrepancies in these areas trigger consultation with the authorities. Updates to STA's database originate solely from official sources. There is no established mechanism to distinguish different vintages of data.

15.   The second channel involves the **Fund's area departments, which collect data from country authorities or from commercial sources** in the course of their regular bilateral surveillance (Article IV consultations) or IMF-program-related activities. These data inform the staff reports that are prepared for the IMF's Executive Board, and are also used by the IMF Research Department in the preparation of the *World Economic Outlook (WEO)* publication.[8] Staff missions to member countries, supported where applicable by IMF resident representatives, collect and evaluate real-sector, fiscal, monetary, financial, and external sector data. An abridged version of these data is published in the Article IV consultation staff reports.[9]

Figure 1. Data Flow at the IMF

Source: IEO.
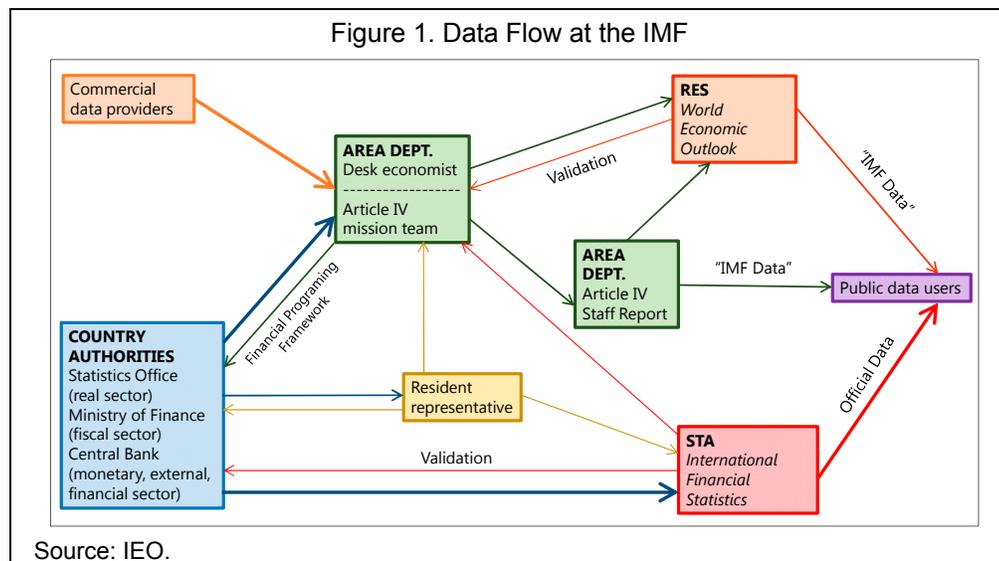
16.   Data inadequacies often come to light in the course of policy discussions between IMF mission teams and country authorities—particularly when a program to be supported by Fund resources is being designed. This can prompt corrective action. In many cases, there is a

---

[7] Reliance on official data gives no assurance of cross-sector consistency as these data often come from different domestic sources and are not necessarily mutually consistent.

[8] Occasionally, data are reported from the IMF area department to the IFS database, but the formal route is typically from country authorities directly to the IFS.

[9] In the past, Article IV consultation reports tended to be less used as a data source. However, researchers are reportedly increasingly using these data. For example, Prichard, Cobham, and Goodall (2014) fill the gaps in *GFS* data with data from consultation reports.

steady flow of communication among desk economists, resident representatives, and country sources of data, allowing the IMF desk staff to constantly update the desk's database. In other cases, missions must spend a substantial share of their time in the field collecting and double-checking aspects of the data. These tasks include verifying data in the primary sources and checking the accuracy of basic calculations and their consistency with methodological standards. The financial programming framework that underlies the IMF staff analysis plays an important role in helping to ensure cross-sectoral data consistency.

B. Dealing with data discrepancies and gaps

17.   Though the most striking data problem for LICs is missing data, another problem is that of discrepancies in the data reported. Given the different formal and informal channels of data transmission and vetting, it is not surprising that discrepancies and gaps in country data often arise in different databases. Indeed, data can appear in three different versions—country statistics, staff estimates, or *IFS* figures—that quite often differ significantly.

18.   The nature and causes of the discrepancies vary depending on country circumstances. One reason why discrepancies arise is that, for the Fund's area departments, the main goal is to have a dialogue on policy, while for statisticians in STA, the main goal is to oversee the universal application of statistical principles. The database compiled by a country desk economist or resident representative often reflects the obvious pragmatic choice of working with the same data that country authorities use. Discrepancies may also occur because of the properties of the data and the methods used to aggregate the statistics.

19.   In both program and surveillance work, agreeing upon a common set of data is crucially important. Part of the process of working with country authorities is to agree on the set of data that supports the policy analysis being conducted. Sometimes IMF staff and country authorities agree to disagree, and staff reports contain either staff estimates or the country's statistics, with the appropriate explanations or caveats included in footnotes (Box 2).

20.   On its part, STA is obligated to publish the country's statistics, seeking as much as possible to ensure that these are reported according to international reporting standards. STA has produced several manuals that specify the definitions and compilation methodologies that are to be followed for each area of the economy and have come to be accepted as the worldwide standard. However, few LICs have the resources and availability of primary data that would be needed to fully adhere to these manuals. Some countries view the manuals as universal compilation manuals (i.e., the standard that is also adopted in statistics for domestic circulation), while others view them as a reporting standard (i.e., the standard at which data are reported to the IMF).

Box 2. Staff Estimates Versus Country Statistics: The Case of Ethiopia

For Ethiopia, IMF staff have had long-standing concerns regarding the methodology used to compile national accounts. The 2013 Article IV consultation report reported that "in consultation with STA, the mission decided to use the official historical figures with a notation regarding data weaknesses and possible overestimation of GDP growth rates. This approach is consistent with the general practice of reporting official historical data in other countries with national account deficiencies. Staff will continue to have its own projections, but at the end of the fiscal year will adopt official statistics based on the national accounts outturns with the appropriate caveat" (IMF, 2013c).

This decision stood in contrast to the practice followed in the 2012 Article IV consultation report, where the staff had presented their own estimates based on a growth accounting exercise and corrected official statistics for 2009/10 and 2010/11 from 10.6 percent and 11.4 percent to 8.0 and 7.5 percent, respectively. Moreover, the tables in that report presented IMF staff estimates instead of country statistics, also for historical data. The staff appraisal noted that "Official GDP growth rates imply productivity increases that appear implausible, suggesting an urgent need to reconcile the historical difference between official GDP statistics and the staff estimates and to ensure accuracy of source data" (IMF, 2012b).

---

21.   *Fiscal and monetary data.* Regarding fiscal and monetary data, the Article IV mission and the country desk staff in most cases work with the administrative data used and provided by the country's ministry of finance and central bank. Country authorities usually aggregate these data in a way that differs from the standards that STA applies in order to make the data globally comparable. Of course, problems exist in following standards in all areas of statistics, but currently there is a perception that there are more acute methodological difficulties with government finance statistics, in particular, with the requirement that accounting be done on an accrual and not on a cash basis.[10] IMF staff has been directed to report fiscal statistics on an accrual basis, regardless of the countries' actual accounting preference. This may involve the staff using estimates to convert cash data into an accrual basis.

22.   *External sector.* Compared with government finance statistics, better established methodologies exist for external sector statistics, especially balance of payments statistics (although some countries are still following outdated *Balance of Payments Manuals*). In this area, at the country level, data are either obtained from administrative sources—which works well under more controlled capital and trade systems—or collected from surveys—which works better for high-frequency data with greater coverage. Data taken from these two sources (administrative versus survey data) will differ, and therefore central banks, ministries of finance, and national statistical agencies may all report different data values.

23.   *Prices.* Price statistics, or inflation as measured by the PPI and the CPI, are collected, aggregated, and transmitted to the IMF by national statistical agencies (sometimes via the central bank or the resident representative), and the data published nationally are the same as

---

[10] As is clear from the metadata reported in Annex 1A, very few LICs have the resources and primary data availability to adhere to the statistics manuals. In particular, the standards set by the *Government Finance Statistics Manual 2001* (IMF, 2001b) have proven unattainable for most.

published in *IFS*. Country desk economists and IMF missions will work with disaggregated data and may use them to compute "core inflation" or other metrics needed for specific analysis, but since these data are usually produced in a timely fashion one would expect relatively few discrepancies across different price-related databases. There may be disagreements about the chosen price baskets and representative samples, but these are not readily apparent in the different databases.[11]

24.   *Real sector.* Two of the most important real-sector statistics are nominal GDP and real economic growth. Typically the national statistical agency has the main responsibility for producing real-sector statistics, and the *IFS* publishes unmodified real-sector data from country authorities. However, the gap between the pertinent date and when the final annual national accounts estimates are ready can range from a year to half a decade.[12] In the interim, the ministry of finance may provide preliminary GDP estimates, and the central bank may provide its own quarterly national accounts estimates. These are often referred to as forecasts, but, in practice, for many countries they are really "nowcasts" or "backcasts" for the previous year or two.

25.   The preliminary GDP numbers that country authorities provide serve as a starting point for a discussion between country authorities and the IMF in consultations. For lack of primary data, the preliminary numbers are often just based on estimates, guesstimates, and proxies. Moreover, given the importance of data on economic activity for understanding key economic variables (e.g., taxes, monetary demand, etc.), the discussion concerns not only methods and source data but also the overall economic picture, using other economic data to triangulate what seem to be reasonable projections of future, current, and recent economic growth.

26.   Two issues may arise here. First, the domestic political debate on economic growth may mean that the data become politicized. This can stand in the way of an open exchange on policy surveillance.[13] Second, IMF staff projections may differ from those that country authorities provide. While neither set of projections are official statistics, the Fund's agreement or disagreement with the preliminary numbers may often be misinterpreted in the public domain as the Fund's endorsement or rejection of official estimates.

---

[11] The World Bank's International Comparison Project collects prices for the same type of goods and services in different countries for the purpose of creating purchasing-power-parity-adjusted metrics, but for now it is accepted that the contents and aggregation of price statistics follow domestic standards. See http://icp.worldbank.org/.

[12] Jerven (2013), in a 2011 survey of Sub-Saharan African economies, found that five countries had prepared a 2010 estimate and that the average gap among the 36 responding countries was two and a half years. Pastor (2009) surveyed 20 countries and found a lag time in reporting of just over two years.

[13] See Samuel (2015) for a discussion of such a process in Mauritania, where the IMF was involved.

27.   Here we compare the data on real GDP growth and the current account that appear in Article IV reports (AIV) with the data that are published in the *IFS*, *WEO*, and the World Bank's *World Development Indicators* (WB).[14] The discrepancies are the widest between AIV and the *IFS* data, and they are wider in the data on the current account than in those on GDP growth (Table 2). The current account data discrepancies clearly show a relatively close correspondence between the AIV and the *WEO*, thus affirming this important direction of data flow. In the current account data, the discrepancies are wider for LICs than for the non-LIC sample countries, but in the real growth data, there is no such clear difference, perhaps because growth data are missing for many LICs. Across the database, about two-thirds or three-quarters of the reported growth data roughly match, and quite large discrepancies can be seen in about a third to a quarter of the observations. Once again, the agreement between AIV and *WEO* is clear, and the discrepancy is largest between AIV and the *IFS*.

Table 2. Discrepancies Among IMF Data Sources[1] (In percent)											
Deviation Threshold[2]	AIV/*IFS*			AIV/*WEO*			AIV/WB			*IFS*/*WEO*		
Real GDP growth rate	*LICs*	*Non-LICs*	*Total*	*LICs*	*Non-LICs*	*Total*	*LICs*	*Non-LICs*	*Total*	*LICs*	*Non-LICs*	*Total*
0 – 10	70.0	64.1	66.7	74.6	75.0	74.8	71.4	64.5	68.6	61.3	77.5	70.4
10 - 30	13.3	17.9	15.9	12.7	18.8	15.1	14.2	25.0	18.6	19.4	15.0	16.9
30 +	16.7	17.9	17.4	12.7	6.3	10.1	14.2	10.4	12.7	19.4	7.5	12.7
Missing data	59.5	18.8	43.4	4.1	0.0	2.5	5.4	0.00	3.2	58.1	16.7	41.8
Current account	*LICs*	*Non-LICs*	*Total*	*LICs*	*Non-LICs*	*Total*	*LICs*	*Non-LICs*	*Total*	*LICs*	*Non-LICs*	*Total*
0 – 10	38.3	69.6	51.9	70.8	83.0	75.6	43.3	70.2	55.1	36.7	83.3	57.4
10 – 30	18.3	13.0	16.0	16.7	12.8	15.1	16.7	12.8	15.0	16.7	4.2	11.1
30 +	43.3	17.4	32.1	12.5	4.3	9.2	40.0	17.0	29.9	46.7	12.5	31.5
Missing data	18.9	4.2	13.1	2.7	2.1	2.5	18.9	2.1	12.3	18.9	0.0	11.5

Source: IEO estimates.
[1] Based on data for all 74 LICs and 48 non-LICs. See Annex 2 for more detailed description of data sources and methodology.
[2] Where data are available in both of the two sources being compared.

28.   According to information provided in interviews with STA staff, while balance of payments, monetary, and fiscal statistics are cross-checked and sometimes also disaggregated and then re-aggregated, price and real sector data are merely re-reported as and when they are made available by the country. STA does not disseminate area department data because STA does not interpret its dissemination mandate to go beyond publishing official country statistics.

---

[14] Thus, the comparison does not capture the problems of uneven updating of historical data in the different databases.

29.   Though the World Bank has a similar mandate,[15] it does disseminate the most up-to-date statistics, and the data it considers to be the best available, in its main databases. This often means that the World Bank draws upon the data available in Article IV consultation reports.[16] The data the IMF uses for operational purposes are available in staff reports, but not in readily accessible databases.

C. Conclusion

30.   Data discrepancies may be confusing for an external user, and the IMF has long been concerned that they could pose a reputational risk. Most recently, the IMF's 2012 "Review of Data Provision to the Fund for Surveillance Purposes" (IMF, 2012a) noted that "*WEO* historical data are not always consistent with *IFS*, which can be a potential reputational risk for the Fund," and—referring to the difficulty in distinguishing between IMF staff estimates and official statistics in IMF staff reports—that "this has raised the IMF's reputational risk regarding the credibility of data disseminated to the public."

31.   The extent of the reputational risk that the IMF actually faces from disseminating conflicting statistics and data of unknown quality is hard to assess directly. But it is safe to assume that reducing the discrepancies in the data and reporting better metadata will reduce reputational risk. There are indications that data users are not aware of differences in the procedures used to compile data in consultation reports and in the *IFS* and that therefore they may be unintentionally misled.[17] Moreover, the area department data may be adjusted or reported differently for specific program-related or surveillance needs, and it is not always made clear what source data or methods may have been used to derive independent staff estimates or the underlying judgment that justifies using different numbers in different reports. A final part of the problem is that because the consultation reports and the *WEO* database serve immediate data needs and include projections, there are no revision studies.

IV. Mechanisms for Data Quality Assurance at the IMF: Theory and Practice

32.   Beyond the occasional quality checks and corrections that, as mentioned above, are made during the normal process of data collection by either STA or area departments, the IMF has three formal tools to report on issues in country-specific data quality: (i) the Statistical Issues Appendix of the Article IV consultation reports—which reports on the adequacy of data for surveillance; (ii) the data dissemination initiatives—comprised of the

---

[15] The World Bank has a mandate to compile GNI per capita data that form the basis for classifying countries into different income groups.

[16] As reported in interviews, country economists of the World Bank informally consult with IMF desk economists to verify economic statistics. See Pastor (2009), for a review of the origins of recent estimates and forecasts of GDP growth.

[17] See, for example, de Las Casas and Monasterski, 2016.

Special Data Dissemination Standard (SDDS) and the General Data Dissemination System (GDDS), which involve countries' self-reporting on data methodology and standards via the IMF's Dissemination Standards Bulletin Board; and (iii) the data modules of Reports on Standards and Codes (ROSCs).

33.   But what is actually meant by data "quality"? This paper relies on how the IMF interprets data quality, as defined through the Data Quality Assessment Framework (DQAF) that the IMF developed in the early 2000s (Box 3). The latter two tools listed in the previous paragraph rely heavily on the DQAF as the measure of quality. Even the first tool—the assessment in the Statistical Issues Appendix—is expected to use the DQAF as a guide if the country has had a ROSC in recent years.

Box 3. The Data Quality Assessment Framework

Work by STA on a framework for assessing data quality began in 1997 (IMF, 1997), in the aftermath of the Asian crisis, and the Data Quality Assessment Framework (DQAF) proposal was endorsed by the IMF's Executive Board in July 2001. The DQAF serves as an assessment methodology, providing a structure for assessing data quality by comparing a country's statistical practices with best practices. In addition to a set of prerequisites for quality (such as the legal and institutional environment for data), the framework addresses five dimensions of data quality:

   1. Assurances of integrity
   2. Methodological soundness
   3. Accuracy and reliability
   4. Serviceability
   5. Accessibility

The DQAF has successfully provided a *lingua franca* for discussing statistical systems, and variations of it are used as a framework for PARIS21 countries when they design national strategies for the development of statistics. Metadata that are reported on the Dissemination Standards Bulletin Board also follow the structure and terminology described in the DQAF, and the DQAF provides the framework for the Data Module of the ROSC.

Data ROSCs involve a staff mission to the country, which collects information to check the metadata provided by the authorities in advance of the mission. The information is then organized according to the DQAF framework.

The five-part structure listed above contains two tiers of subcategories and ultimately comprises 50 different dimensions. The detailed assessment involves scoring the country on these 50 quality dimensions using the metrics of Observed, Largely Observed, Largely Not Observed, or Not Observed. A full data ROSC will conduct this scoring for the country's statistics in the real, external, fiscal, and monetary sectors.

The scorings are reviewed at IMF headquarters for internal coherence and external validity. In practice, assurance of external validity means that the scoring is compared to previously completed data ROSCs for similar countries. In a sense, the ROSC does not apply an absolute universal standard but rather creates a sort of floating benchmark for judging whether standards are observed or not. The detailed results of the review are shared with country authorities, and on the basis of this consultation, the draft report is drawn up.

The final report includes responses from the authorities on the ratings. The scores at the most detailed level are not provided in the published report nor are the ratings aggregated over the five quality dimensions.

34.   There are at least three obvious reasons why the IMF should take an interest in accurate and transparent reporting about the quality of its data and statistics. First, both policymakers and IMF staff need good data to support policy formulations and recommendations. If the data are unreliable, they will compromise the decisions being made. Second, the IMF is a disseminator of statistics, and one way of mitigating the reputational risks of reporting inaccurate country statistics is to also report metadata on the sources and methods used to produce the statistics. Third, the IMF is a provider of technical assistance—and good data diagnostics can influence its own decisions and those of other stakeholders regarding the technical assistance they provide. Even more, such diagnostics can also be used to evaluate the effectiveness of technical assistance.

35.   There are some broader, often understated, benefits of good reporting on the quality of data. A system of thorough initial diagnostics, such as that provided during a technical assistance mission or by a data ROSC, combined with subsequent rigorous and consistent quality assessments and follow-up technical assistance missions, could help generate the conditions necessary to improve and sustain statistical capacity. In the medium term, the advantage would become evident as missions would gain access to more reliable data and metadata and country desk economists could have more confidence in the data they use for day-to-day surveillance. Moreover, in countries that experience problems with data quality and availability, the time that resident representatives and staff missions currently spend chasing down and cross-checking data could be put to better use in policy discussions.

36.   The potential benefits also extend to country authorities and data users. Beyond the obvious importance of data for policy formulation and evaluation, better data-sharing procedures would directly benefit countries. Frequent and automated sharing of data would help ensure adherence to agreed dissemination practices and reduce the risk of political interference in the provision of data. Moreover, having sufficient quality checks and updated comprehensive metadata in place could strengthen data users' confidence in the IFS and other IMF data sources.

37.   In the short run, greater candor in reporting on data quality and improved transparency in reporting on statistical weaknesses may seem politically costly to country authorities and country desks. However, in the medium term, attending to statistical weaknesses may avert costly political conflict. The bottom line is that better data and better procedures for reporting on data will support better policy advice and analysis.

A.   The GDDS and the SDDS

38.   The IMF established the SDDS and the GDDS in 1996 and 1997, respectively, in the aftermath of the Mexican crisis. Participation is voluntary, but a subscribing member must meet certain requirements. As of November 2015, of the 188 IMF member countries, 111 had subscribed to the GDDS and 71 to the SDDS. The GDDS is seen as a stepping-stone to

the SDDS and is open to all IMF members, while the SDDS is meant for countries that already have access or are seeking access to international financial markets.

39.   The GDDS and the SDDS serve different purposes and have different requirements that subscribers must meet. While the SDDS is geared toward increasing the public availability of timely and comprehensive data, the purpose of the GDDS is to provide a framework for promoting and supporting improvements in statistical capacity.

40.   The SDDS and the GDDS focus on four dimensions of data:

- *Coverage, periodicity, and timeliness*. The SDDS specifies the data categories that cover the four sectors of the economy (real, fiscal, financial, and external). It also specifies requirements about the periodicity and timeliness of these data. The GDDS 7addresses the development, production, and dissemination of data on these four sectors and of data on socio-demographics.

- *Public access*. The SDDS requires advance dissemination of release calendars and simultaneous release of data to all interested parties. The GDDS focuses on policies and practices to improve the dissemination of readily accessible and reliable data.

- *Integrity of disseminated data*. The SDDS requires countries to disclose information regarding the terms and conditions under which data are produced and disseminated, including the government's access to data before release, ministerial commentary upon release, and revisions or major changes in methodology. The GDDS encourages policies and practices to enhance the integrity of the disseminated data.

- *Quality of disseminated data*. The SDDS requires dissemination of information on statistical methodology, details about data components, reconciliations with related data, and statistical frameworks that make possible cross-checks and checks of reasonableness. The focus in the GDDS is on members' plans for improving data quality.

41.   Metadata pertaining to the GDDS and the SDDS are posted on the Dissemination Standards Bulletin Board (DSBB). Subscribers are responsible for the accuracy and reliability of the metadata. Participants are asked to update their metadata at least annually.

42.   Of the 74 LICs that were covered in this study, five countries (Eritrea, Lao P.D.R, Sierra Leone, Somalia, and Uzbekistan) were not participants in either the SDDS or the GDDS. Only two countries in the sample are subscribers to the SDDS (Kyrgyz Republic, and Moldova; Box 4). The majority (67 countries) are subscribers to the GDDS.

43.   The GDDS provides countries with a systematic framework for organizing their metadata and making them public, and for planning their future statistical capacity building. Participating in the GDDS could help a country coordinate capacity building efforts not only

with the IMF but also with other organizations such as the World Bank and the U.K. Department for International Development. According to the GDDS guide, country subscribers are encouraged to appoint a coordinator who has "sufficient authority to obtain the cooperation of all national agencies involved in the compilation and dissemination of the data categories covered by the GDDS." Furthermore, "at least once a year, the coordinator should forward metadata to the relevant contact persons for each data category" (IMF, 2013a). While the GDDS guide sets out a system of voluntary self-reporting, it also suggests that "at any time, the IMF staff may take the initiative of contacting the coordinator to suggest or request improvements to the metadata" (ibid).

Box 4. Graduates from the GDDS to the SDDS

It is notable that only two of the LICs are members of the SDDS. This inertia indicates that there may be significant constraints on human capital and resources and/or that the majority of LICS do not see graduating from the GDDS to the SDDS as a high priority. Only seven countries that were GDDS members have graduated to the SDDS, of which four were LICs. There is a notable overrepresentation of Eastern European countries and former Soviet Union members in the group that has graduated to the SDSS, perhaps indicating human capital strength in the number of statisticians who were able to adopt international statistical standards. A study of reform in statistical offices in the former Soviet Union suggests that this is the case (Herrera, 2010). These countries also benefitted from STA technical assistance missions, including data ROSCS. There are no LIC graduates to the SDDS from the African region as of yet.

---

44.   Adherence to the GDDS guidelines varies considerably. Interviews with country authorities and staff do not indicate that the GDDS is a high priority, and in many cases the country coordinator position is empty. Country authorities, area department desks, and STA do not seem to see updating the GDDS as a matter of urgency. Indeed, only seven countries (about 10 percent of GDDS members) updated their metadata in 2013 or 2014 and as many as fifteen have not updated their metadata in the GDDS since 2003 or earlier. The mean updating gap is six years and the median is seven for the 62 countries for which we have data. Do the countries really have no updates in their metadata to report? The gaps suggest that some countries completely neglect the GDDS. For example, as of December 2015, the aforementioned upward GDP revision in Ghana from a 1993 to a 2006 base was still not reflected in the information published on the DSBB.

45.   Beyond the lags in updating information, the GDDS has to contend with other problems: it must depend on candor in self-reporting, and the extent to which a method has been applied may depend on interpretation. A clear example relates to self-reporting in the System of National Accounts. When the United Nations Economic Commission for Africa (UNECA) asked member countries "Do you follow SNA 1993 or newer?" as part of a survey of SNA adherence, all countries except two replied that they followed the 1993 SNA (UNECA/AfDB/AUC, 2013). In the same year, the African Development Bank (AfDB) surveyed the same membership using questions that addressed five known problem features related to adopting SNA 1993. It found that only two countries could confirm that they applied these features (AfDB, 2013). In principle, the process involved in producing data ROSCs could

address this issue by giving country authorities, the IMF, and other data users an in-depth review of a country's statistical system and the quality of the statistics it produces.

46.   In May 2015, the GDDS was reformed in order to address some of the issues that limited its progress—thus becoming the e-GDDS. Active monitoring of the countries' dissemination practices was introduced with the aim to foster dialogue during Article IV consultation missions on constraints and capacity-building needs, thereby providing incentives to graduate to the SDDS and drawing attention of policymakers to the need of statistical development (IMF, 2015).

B. Data ROSCs

47.   Data ROSCs are initiated at the request of country authorities and are intensive, expensive exercises in information gathering (Box 3 above). Since 1999, there have been 111 data ROSCs, but their heyday has passed, and after becoming increasingly infrequent (Figure 2), they have now been (at least temporarily) suspended. The decline stemmed not from lack of interest on the part of country authorities, but rather from the IMF's "downsizing" in 2007–08. From 2003–06, when twelve to sixteen data ROSCs were done annually, IMF member countries could reasonably expect to get an evaluation of their statistical systems every ten years or so. Even had the data ROSCs not been suspended, at the pace of 2013–14, it would take more than a century for this to happen.

Figure 2. Trend in Data ROSCs, 1999–2014

Source: IMF (2014b).
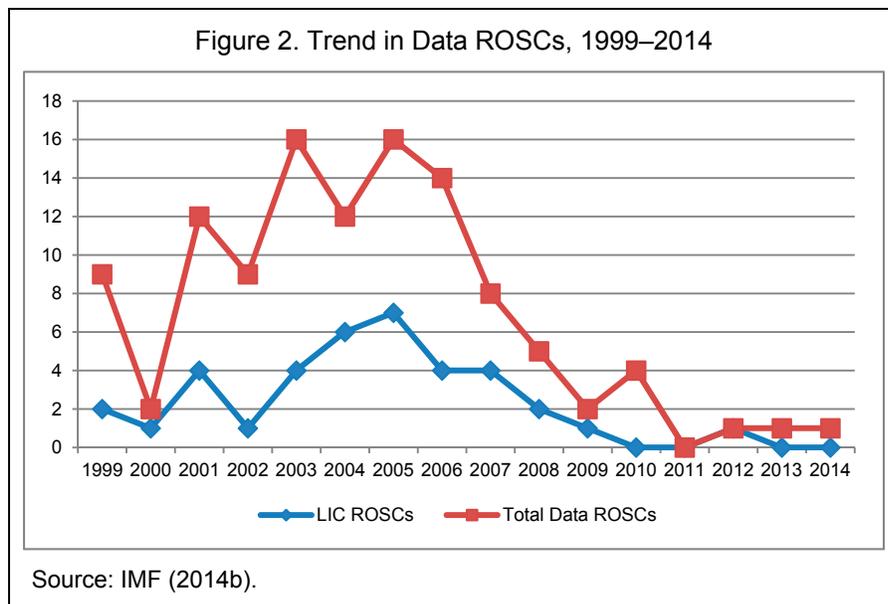
48.   In undertaking a data ROSC, the emphasis from STA and country authorities has been to assess the country's statistical system rather than the outputs of the system. In the course of the ROSC, higher-level country authorities become acquainted with the gaps and weaknesses in their own statistical system. The process is informally used as a fact-finding

mission to make the case to the country authorities—on behalf of the statistical agencies and those tasked with collecting and disseminating statistics in the central bank and the ministry of finance—that more resources and legal provisions are needed in certain areas in order to improve the functionality of the statistical system.

49.   The data module of the ROSC does not provide an evaluation of the statistical output itself.[18] It is rather a piecemeal evaluation of the inputs that are assumed to be essential for a high-quality statistical system. The reader of a data ROSC report is left to form his own judgment on whether any bias exists that is attributable to outdated benchmarks or an inadequate sampling method.

50.   Nor does a data ROSC (and this seems intentional) provide a metric with which external users could readily rank or classify countries according to their overall statistical capacity. It is, of course, debatable what merit an aggregated index, such as the Statistical Capacity Indicator (World Bank, 2009) would have,[19] but the ROSC nonetheless collects information that is potentially valuable for data users who seek a summary impression of data quality.

51.   Clarifying the intent of the ROSC exercise and the DQAF, the 2001 Review of the IMF's Data Standards Initiatives (IMF, 2001a) suggested that a framework or methodology for assessing data quality should provide "a structure and common language for data quality that could be distilled into an assessment framework" but also that "given the complexity of the assessment and the wide differences in countries' statistical systems, it should be clear that the DQAF cannot be used to rank the qualities of countries' data." Indeed, as described in Box 3 above, the baseline for ruling on whether a country is observing a particular standard is not universal but is a moving benchmark based on the practices of what are judged to be suitable peer countries for comparison. The 2001 Review also stressed that "Those applying the DQAF will need to be constantly alert to the country setting—the culture, the legal environment, the stage of statistical development. They would need to ask, when finding that a certain practice is not observed, whether the intent of the practice is achieved by some other means." This again shows that the original intent of the DQAF was to emphasize the appropriateness of the process in the light of the circumstances of the country. It also confirms that cross-country comparability of the data ROSCs is not a priority.

---

[18] As an illustration of the data ROSC, Annex 3 uses some elements of the data ROSC to evaluate aspects of the IMF's own statistical system.

[19] While the IMF may be reluctant to assign overall ratings to countries, the World Bank's Statistical Capacity Indicator (SCI) ranks countries on a scale from 0 to 100 on a range of categories. The data are drawn from readily available metadata collected by various institutions and databases, of which the SDDS and the GDDS are among the most important data sources. The SCI has become increasingly visible to a wide variety of stakeholders, including academics and country authorities themselves.

52.   Thus, in its current design, the ROSC is intended more to increase the awareness of country authorities than to provide information to external data users. The ROSC offers an evaluation of the capacity of a statistical system, but does not evaluate the quality of the data and, therefore, does not provide answers to the questions in which most data users (and IMF staff) might be interested, such as whether inflation is understated, growth estimates are biased, and so forth.

C. Article IV Consultations

53.   According to Article IV of the IMF's Articles of Agreement, member countries are required to consult periodically (typically annually) with the IMF. The reports on these consultations include a Statistical Issues Appendix in which staff must assess the adequacy of data for surveillance.[20] While the GDDS and SDDS provide an overview of metadata and data dissemination, and the ROSC evaluates the statistical system, the assessments in Article IV consultation reports are the only judgment by Fund staff of the adequacy of the output of the statistical systems for policymaking. Even these assessments are not an evaluation of the quality of official statistics as such; instead, they indicate whether the general availability and reliability of information is sufficient to conduct surveillance.

54.   The assessments in Article IV consultation reports have not been without problems, as indicated in successive reviews of data provision to the Fund for surveillance purposes. The 2004 Review "found scope for sharper assessments in staff reports of data adequacy and its implications" (IMF, 2004). The next review (IMF, 2008) reiterated these concerns and found "some upward bias in characterizations of adequacy." Accordingly, it reformulated the adequacy criteria to their present form, while stating that "Adequacy should be judged based on the aggregate impact of shortcomings in coverage, periodicity, timeliness, and other aspects of quality, on the staff's ability to analyze the key economic issues and draw policy conclusions. Key to the distinction between cases B and C would be a judgment whether data deficiencies have introduced significant uncertainty into the policy conclusions." Four years after the introduction of the new assessment process, the 2012 Review (IMF, 2012a) found that it had "resulted in greater candor in staff reports," but also that a substantial difference remained between the proportion of mission chiefs reporting important deficiencies (58 percent) and the number of cases assigned a C rating (12 percent). Thus, the 2012 review concluded "… there may be some hesitancy by teams to use the 'C' classification."

55.   Staff assessments of data adequacy for surveillance need not necessarily be consistent with the findings reported in data ROSCs. For example, often in LICs, official GDP statistics are compiled using an outdated benchmark year, or fiscal data are reported on a cash basis rather than on the accrual basis that international standards require, or there are well-known biases in the way consumer price indexes are computed (such as using price data

---

[20] Data are rated as A: "Data provision is adequate for surveillance;" B: "Data provision is broadly adequate for surveillance;" and C: "Data provision has serious shortcomings that significantly hamper surveillance."

from urban areas only). Probably such shortcomings in the source data and methods of recording would routinely give rise to a "Not Observed" or "Largely Not Observed" rating in a data ROSC. Yet in an Article IV consultation report, these data for the same countries may still be judged as "adequate" or "broadly adequate."

56. This inconsistency might be reconciled if it is the case that consultation reports evaluate the totality of information available to staff, whereas a ROSC evaluates specifically whether official statistics conform to international standards. These are quite different questions. Staff may feel that, on the basis of the information available to them, they are able to judge economic trends and evaluate economic risks, even though the actual data might be deemed outdated, inappropriate, and inadequate for cross-country comparisons. To put it bluntly: you can still have a useful conversation about policy although data are of low quality. That is, staff assessments are an indication of whether data quality became an issue in policy surveillance—but not of whether statistics are compiled in accordance with international standards.

57. The analysis in an Article IV consultation may not necessarily be invalid if the mission considers that a country's GDP is underestimated. An underestimate might bias calculations of the country's debt-to-GDP or tax-to-GDP ratio, but as long as the mission members feel reasonably well informed about the direction and the size of bias, they can offset the bias in their own estimates or keep in mind the margins of error when they consider risks. This is a pragmatic choice—again geared toward evaluating the country context and not necessarily in line with actions that would safeguard the universal comparability of macroeconomic statistics.

58. Staff who rate the adequacy of data for surveillance face incentive problems. Giving a "C" ranking to a country's data may provoke resistance from the country authorities and even from markets. Such a ranking undermines the credibility of the entire staff report. It also creates additional work, as the guidelines related to this ranking require the staff to give specific advice on remedial action, and to follow up the advice in subsequent mission and staff reports (IMF, 2013c). In sum, it is not surprising that, by and large, data inadequacies that on pure quality considerations should indicate a "C" rating are often awarded a "B" rating.

59. Among the 74 low-income countries, two are ranked "A" in their latest consultation reports, 58 as "B," and only 13 as "C."[21] These rankings are broadly comparable with the scores given in the World Bank's Statistical Capacity Indicator (SCI): the average SCI score for the "A" countries is 78, the average for the "B" countries is 63, and the average for the "C" countries is 50. But the average scores hide a lot of country variation. The lowest country score on the SCI for the "B" countries is 38—which is lower than those of all the "C" countries except one (whose SCI score is 33). Meanwhile, Uganda, one of the only two

---

[21] One (Somalia) is not evaluated.

LICs ranked "A," receives an SCI score of 70, while 17 of the "B" LICs score higher than 70 on the SCI.

60.　　The Statistical Issues Appendix to Article IV consultation reports includes a table of "Common Indicators Required for Surveillance." This table mostly describes issues related to the frequency of data provision, but for countries that have had a data ROSC,[22] it also includes a report on two of the DQAF data quality elements: methodological soundness and accuracy/reliability. These two elements are certainly important for the immediate work of the surveillance mission, but the Statistical Issues Appendixes do not cover other quality attributes important for many data users or country authorities, and they do not give the Fund the full information it needs to prioritize technical assistance and capacity building.

D.　Conclusion

61.　　Thus, there are reasons to be concerned that the data adequacy assessments in Article IV consultations leave out important data quality dimensions and that these assessments are not detailed enough to provide a basis for gauging when staff estimates conflict with those provided by country authorities. Such detailed assessments would need to be conducted using the template of the DQAF, but reports published in the ROSC format have been too infrequent and, as mentioned, they focus on the system that provides the data, not on the quality of the data themselves.

V.　CONCLUDING REMARKS

62.　　For the IMF, the quality of the macroeconomic data it uses and disseminates is a major concern. The analysis of economic reality, the policy discussions with member countries, and the integrity of the data it disseminates are predicated on a solid and reliable base of statistical information. Over time, the Fund has developed a number of formal and informal mechanisms and arrangements to obtain assurances about the data it uses, but significant weaknesses and deficiencies remain in this regard. In practice, bits and pieces of a data quality system are in place at the IMF, but because the system is incomplete in coverage, unevenly adopted, and incoherent at different levels, it does not perform as might be desired.

63.　　The data module of the ROSC comes closest to shedding light on the quality of the data that countries provide, but its focus is more on the functionality of the statistical system and on the process of statistics production than on the statistical output itself. Nor is the DQAF that lies at the core of the data ROSC designed to judge the quality of a given datum but rather to provide a framework to help identify weaknesses in the country's statistical setup. Of course, a well-functioning statistical system is more likely to produce good data,

---

[22] Only 22 of the 74 LICs have had a data ROSC.

but there is no certainty in this regard. Only a minority of member countries have had a ROSC, and the ROSCs have now been suspended.

64.   The IMF's initiatives to promote the dissemination of statistical information have positively affected quality inasmuch as they require participating countries to focus on metadata, and—by publishing these—expose them to the scrutiny of their peers and the markets.

65.   The collection of data by the Fund, whether by STA or by area departments in the course of their surveillance or lending activities, involves some validation of data that occasionally leads to quality improvements. The interaction of STA with its country counterparts and the hands-on work by staff missions or resident representatives with country authorities make additional contributions to data quality assurances and improvements.

66.   Perhaps the most significant contributions of the IMF to the improvement of statistics overall are its work on statistical methodology and the technical assistance it offers. These contributions, however, work gradually and their results become apparent only over the medium term. The methodological manuals developed by the IMF have become the world standard that countries seek to adopt and implement, while technical assistance is the effective force that bit-by-bit, country-by-country, works for the improvement of data.

67.   All these efforts to strengthen quality assurances, nevertheless, fall short of reducing the uncertainties that still surround the quality of the data used and disseminated by the IMF. Undoubtedly, the Fund needs to strengthen its work in all these areas, but the realm of the possible is limited. Ultimately, the quality of data will depend on the attention and resources that member countries devote to it.

REFERENCES

African Development Bank (AfDB), 2013, "Situational Analysis of the Reliability of Economic Statistics in Africa: Special Focus on GDP Measurement," June.

Africa Report, 2014, "Gambia GDP Adds 25 Percent after Rebasing Economy," *Africa Report*, February.

Alexander, W. E., J. Cady, and J. Gonzalez-Garcia, 2010, *The IMF's Data Dissemination Initiative after 10 Years* (Washington: International Monetary Fund).

de Las Casas, Miguel, and Chris Monasterski, 2016, "How Well Is the IMF Doing on Data? Evidence from Surveys," IEO Background Document No. BD/15/01 (Washington: International Monetary Fund).

Devarajan, Shantayanan, 2013, "Africa's Statistical Tragedy," *Review of Income and Wealth* 51 (Supplement S1): S9–S15.

Economist, The, 2014, "Africa's New Number One," April 12.

Financial Times, 2013, "Africa at Dawn," October 28.

Herrera, Yoshiko, 2010, *Mirrors of the Economy: National Accounts and International Norms in Russia and Beyond* (Ithaca, New York: Cornell University Press).

International Monetary Fund (IMF), 1995, "Strengthening Fund Surveillance: Provision of Statistical Data by Members," SM/95/59, March (Washington).

_____, 1997, "Provision of Information to the Fund for Surveillance: Progress Report" SM/97/269, November (Washington).

_____, 2001a, "Fourth Review of the Fund's Data Standards Initiatives," SM/01/208, July (Washington).

———, 2001b, *Government Finance Statistics Manual 2001* (Washington).

_____, 2004, "Review of Data Provision to the Fund for Surveillance Purposes," SM/04/56, April (Washington).

———, 2008, "Review of Data Provision to the Fund for Surveillance Purposes," SM/08/76, March (Washington).

———, 2012a, "Review of Data Provision to the Fund for Surveillance Purposes," SM/12/229, August (Washington).

———, 2012b, "The Federal Democratic Republic of Ethiopia—Staff Report for the 2012 Article IV Consultation," SM/12/230, August (Washington).

———, 2013a, *The General Data Dissemination System: Guide for Participants and Users* (Washington).

———, 2013b, *Sub-Saharan Africa: Building Momentum in a Multi-Speed World* (Washington).

———, 2013c, "The Federal Democratic Republic of Ethiopia—Staff Report for the 2013 Article IV Consultation," SM/13/249, August (Washington).

———, 2014a, "Nigeria: 2013 Article IV Consultation: Staff Report, Press Release, and Statement by the Executive Director for Nigeria," Country Report No. 14/103 (Washington).

———, 2014b, "Standards and Codes ROSC Participation Database," April (Washington).

_____, 2015, "Ninth Review of the International Monetary Fund's Data Standards Initiatives," SM/15/80 (Washington), April

Jerven, Morten, 2012, "Briefing: For Richer, For Poorer: GDP Revisions and Africa's Statistical Tragedy," *African Affairs* 112/446: 138-147.

———, 2013, *Poor Numbers: How We Are Misled by African Development Statistics and What to Do about It* (Ithaca, NY: Cornell University Press).

———, and Magnus Ebo Duncan, 2012, "Revising GDP Estimates in Sub-Saharan Africa: Lessons from Ghana." *African Statistical Journal* 15, August.

———, 2014a, "Measuring African Development: Past and Present", *Canadian Journal of Development Studies* 35(1): 1-8.

———, 2014b, "Benefits and Costs of the Data Revolution." Copenhagen Consensus, Assessment Paper.

Manson, Katrina, 2014, "Kenya to Make Wealth Leap with New GDP Statistics," *Financial Times*, April 9.

Pastor, Gonzalo C., 2009, "A Framework to Assess the Effectiveness of IMF Technical Assistance in National Accounts," IMF Working Paper (Washington).

Prichard, Wilson, Alex Cobham, and Andrew Goodall, 2014, The ICTD Government Revenue Dataset (September 1, 2014). ICTD Working Paper 19, (Brighton, U.K.: International Centre for Tax and Development).

Samuel, Boris, 2015, "Economic Calculations, Instability and (In)formalization of the State in Mauritania, 2003–2011," in Morten Jerven (ed.) *Measuring African Development: Past and Present* (London: Routledge).

UNECA/AfDB/AUC, 2013, "Statistics for Good Economic Governance, Regional Integration, and Sustainable Development in Africa—African Project on the Implementation of the 2008 System of National Accounts: Phase I," Draft, May.

World Bank, 1999–2013, Bulletin Board on Statistical Capacity, 1999–2013.

———, 2009, "Note on the Statistical Capacity Indicator."

ANNEX 1A. OVERVIEW OF STATISTICAL DATA FOR LOW-INCOME COUNTRIES

Country	Last ROSC[1]	AIV Rating[2]	SCI[3]	G(S)DDS GAP[4]	Base Year		Fiscal Accounts Accrual/Cash Basis
					GDP	CPI	
Afghanistan		B	50	8	2002	2004	Cash
Bangladesh	2005	B	66	6	2005	2005	Cash
Benin		B	54	12	1985	1996	Cash
Bhutan		B	72	N/A	N/A	2003	Cash
Bolivia	2007	B	67	9	1990	1991	Cash
Burkina Faso	2004	B	72	13	1999	2008	Cash
Burundi		C	54	3	2005	2013	Accrual
Cambodia		B	68	7	2000	2006	Cash
Cameroon		B	60	6	1989	1993	N/A
Cape Verde		B	66	11	1980	N/A	N/A
CAR		B	57	10	1985	1981	Cash
Chad	2007	B	58	12	2005	1994	Cash
Comoros		C	44	1	2007	1999	N/A
Côte d'Ivoire		B	63	13	1997	2008	Accrual
Djibouti		B	46	2	1986	1999	Cash
Dominica	2007	B	53	8	2006	2010	Cash
DRC		B	43	10	1987	1998	Cash
Eritrea		C	32	N/A			
Ethiopia		B	63	11	2011	2011	Cash
Gambia, The	2005	B	68	N/A	2004	2004	Cash
Ghana		B	59	9	2006	1997	Cash
Grenada	2007	B	48	12	2006	2010	Cash
Guinea		B	58	12	1986	1991	N/A
Guinea-Bissau		C	51	11	2003	2008	Cash
Guyana		B	53	N/A	2006	2010	N/A
Haiti		C	39	5	1986	2004	Cash
Honduras	2005	B	76	9	2000	1999	Cash
Kenya	2005	B	58	5	2001	1997	Cash
Kiribati		B	38	1	2006	1996	Cash
Kyrgyz R.	2003	A	84	SDDS			
Lao P.D.R.		C	68	N/A			
Lesotho		B	69	3	N/A	2010	Cash
Liberia		C	43	1	2008	2005	Cash
Madagascar		B	66	10	1985	2000	Cash
Malawi	2005	B	74	N/A	2010	2012	Cash
Maldives		B	67	3	2003	2012	N/A
Mali		B	68	11	2006	1996	Cash
Marshall Island		B	53.3	0	2003/04	2003	Accrual
Mauritania		B	58	10	1985	2002	Cash
Micronesia		C	37.8	0	2003/04	2008	Accrual
Moldova	2006	B	88	SDDS	1996		Cash
Mongolia	2008	B	78	0	2005	2005	Cash
Mozambique	2003	B	78	4	2009	2004	Cash
Myanmar		C	49	1	2010	2006	N/A
Nepal		B	68	3	2000	2005	Cash

Country	Last ROSC[1]	AIV Rating[2]	SCI[3]	G(S)DDS GAP[4]	Base Year GDP	CPI	Fiscal Accounts Accrual/Cash Basis
Nicaragua	2005	B	69	9	2006	2006	Cash
Niger	2006	B	69	11	1987	1996	Cash
Nigeria		B	76	9	1990	2009	Cash
Papua New Guinea		B	38	2	1998	1977	N/A
Republic of Congo		B	50	11	2000	2005	Cash
Rwanda		B	77	3	2011	2014	N/A
Sao Tome and Principe		B	53	10	2001	1997	N/A
Samoa		B	51	N/A	2002	2008	Cash
Senegal	2002	B	63	5	1999	1996	N/A
Sierra Leone		C	54	N/A	2006	2010	Cash
Solomon Islands		B	41	N/A	2004	2005	Cash
Somalia		N/A	27	N/A			
South Sudan			N/A	N/A	2009	2009	
St. Lucia		B	63	10	2006	2008	Cash
St. Vincent and The Grenadines	2007	B	50	3	2006	2010	Cash
Sudan		B	44	11	1981	2007	Cash
Tajikistan	2005	B	76	1	N/A	N/A	Cash
Tanzania	2004	B	69	0	2001	2010	Cash
Timor-Leste		C	49	2	2010	2012	Cash
Togo		B	57	10	2000	1996	Cash
Tonga		B	59	8	2010	2002	Cash
Tuvalu		C	N/A	1	2005	2011	Modified accrual
Uganda	2006	A	70	6	2009/10	2005/06	Cash
Uzbekistan		C	61	N/A			
Vanuatu		B	49	10	2006	1998	Cash
Vietnam		B	73	2	2000	2009	Cash
Yemen		B	43	4	2000	1999	Cash
Zambia	2005	B	62	8	2010	2009	Cash
Zimbabwe		B	53	1	2009	2012	Cash

_____

[1] Last ROSC: the year when the most recent ROSC was conducted.
[2] AIV: Adequacy for Surveillance rating of the country's statistics in 2012, or latest available prior to 2012.
[3] SCI: World Bank's Statistical Capacity Index in 2012.
[4] GDDS: the gap between the most recent update year and 2014. For example, 8 means the most recent update was in 2006.

ANNEX 1B. STATISTICAL OVERVIEW OF SELECT NON-LOW-INCOME COUNTRIES

Country	Last ROSC[1]	AIV Rating[2]	SCI[3]	G(S)DDS GAP[4]	Base Year GDP	Base Year CPI	Fiscal Accounts Accrual/Cash Basis
Algeria	N/A	B	N/A	N/A	2001	2001	Cash
Antigua	N/A	B	N/A	12	1990	2001	Cash
Australia	2012	A	N/A	SDDS	2013/14	2011/12	Accrual
Albania	2006	B	78	7	N/A	2001	Cash
Angola	2006	B	50	11	2002	2014	Cash
Armenia	2009	B	87	SDDS		2010	Cash
Azerbaijan	2003	B	74	4	2005	2005	Cash
Brazil	2005	A	78.9	SDDS	2000	1993	Cash
Canada	2008	A	N/A	SDDS	2010	2002	Accrual
Colombia	2006	A	84.4	SDDS	2005	2008	Accrual and cash
Botswana	2007	B	63	1	2006	2006	Cash
China	N/A	B	70	2	2010	2000	N/A
Costa Rica	2010	B	78	SDDS	1991	2006	Accrual
Egypt	2005	B	87	SDDS	2011	2010	Cash
Equatorial Guinea	N/A	C	N/A	N/A	1985	N/A	N/A
Fiji	N/A	B	N/A	12	1989	2008	Cash
Finland	2005	A	N/A	SDDS	2000	2010	Accrual
Germany	2006	A	N/A	SDDS	2005	2010	Cash
Georgia	2003	B	93	SDDS			Cash
Greece	2005	C	N/A	SDDS	2005	2009	Accrual
Hungary	2004	A	84	SDDS	Previous year	Previous year	Cash
Indonesia	2005	B	82	SDDS	2000	2007	Cash
India	2004	B	81	SDDS	2011/12	2011/12	Cash
Ireland	N/A	A	N/A	SDDS	2003	2006	Accrual & cash
Israel	2006	B	N/A	SDDS	2010	2012	Accrual & cash
Jamaica	N/A	C	81	7	1996	2004	Cash
Kazakhstan	2008	B	N/A	SDDS	1994	1995	Cash
Malaysia	N/A	B	N/A	SDDS	2010	2010	Cash
Mexico	2013	A	88	SDDS	2003	2010	N/A
Morocco	2003	A	70	SDDS	1998	2006	Accrual & cash
Namibia	2005	B	N/A	5	2010	2010	Cash
New Zealand	N/A	A	N/A	N/A	2010	2006	Accrual
Paraguay	2006	B	68	0	1994	2007	Accrual
Pakistan	2004	B	64	11	2005	2007	Cash
Poland	2003	A	81	SDDS	2005	Previous year	Accrual
Portugal	N/A	A	N/A	SDDS	2006	2012	Accrual & cash
Saudi Arabia	N/A	B	N/A	6	1999	1999	Cash
Slovenia	N/A	A	N/A	SDDS	Previous year	2005	Accrual
South Africa	2001	A	71	SDDS	2005	2012	Cash
Sri Lanka	2002	B	78	7	1998	1996	Cash

Country	Last ROSC[1]	AIV Rating[2]	SCI[3]	G(S)DDS GAP[4]	Base Year		Fiscal Accounts Accrual/Cash Basis
					GDP	CPI	
Korea	2010	A	N/A	SDDS	2010	2010	Accrual
Spain	N/A	A	N/A	SDDS	2008	2011	Accrual
Sweden	2002	A	N/A	SDDS	Previous year	1980	Accrual
Thailand	2006	B	N/A	SDDS	2002	2011	Accrual
Turkey	2009	B	82	SDDS	1998	2003	Accrual
Ukraine	2003	B	89	SDDS	2007	2010	Accrual & cash
Uruguay	2014	B	92	SDDS	2005	1997	Accrual
United States	N/A	A	N/A	SDDS	2009	1982	Accrual

----

[1] Last ROSC: the year when the most recent ROSC was conducted.
[2] AIV: Adequacy for Surveillance rating of the country's statistics in 2012, or latest available prior to 2012.
[3] SCI: World Bank's Statistical Capacity Index in 2012.
[4] GDDS: the gap between the most recent update year and 2014. For example, 8 means the most recent update was in 2006.

ANNEX 2. DATA SOURCES AND METHODOLOGY FOR TABLE 2

Table 2 summarizes the degree of inconsistency in the data for real GDP growth and the current account balance (in US dollars) across three IMF data sources—Article IV staff reports, the International Financial Statistics (IFS), and the World Economic Outlook (WEO)—and the World Bank's World Development Indicators (WB). The summary is based on all 74 low-income countries (LICs) and 48 randomly-chosen emerging market and advanced economy countries (non-LICs). The data sources and comparison methodology are described below:

Data sources:

- Article IV consultation reports: Data are gathered from selected macroeconomic data tables of the most current Article IV consultation reports for the countries. For cases where the Article IV reports are not recently updated but there are recent program reviews, the most current program review is used as a substitute.

- IFS: Data from IFS are from the February 2015 version of the online database: http://elibrary-data.imf.org/FindDataReports.aspx?d=33061&e=169393

- WEO: Data from WEO are from the October 2014 *WEO* published report, which was the most recent version of the WEO database when this exercise was done: http://www.imf.org/external/pubs/ft/*WEO*/2014/02/*WEO*data/index.aspx

- WB: Data from the World Bank are from the February 2015 edition of the World Bank's World Development Indicators database: http://databank.worldbank.org/data/views/variableSelection/selectvariables.aspx?source=world-development-indicators

Comparison methodology

The data from Article IV staff reports are used as the reference source to check for data coherence: Article IV versus IFS, Article IV versus WB, and Article IV versus WEO. By the same token, the data from IFS are used as the reference source to check for data coherence between IFS and WEO. Data discrepancies are assessed by computing the percentage differences of the data from the other sources compared to those of the reference source. For example, the percentage differences between the real GDP growth data of the World Bank's database and those of the Article IV reports are calculated as $(GDP_{AIV} - GDP_{wb})/GDP_{AIV})*100$. The results are converted into absolute numbers to avoid negative statistics. The discrepancies are then classified into three groups: (1) smaller than 10 percent, (2) from 10 percent to 30 percent, and (3) greater than 30 percent, interpreted as minor acceptable discrepancies, discrepancies, and significant discrepancies, respectively.

ANNEX 3. USING THE DATA QUALITY ASSESSMENT FRAMEWORK TO EVALUATE IMF STATISTICS

This evaluation of the IMF's statistical system using the DQAF is solely the author's personal opinion, based on evidence gathered, to the extent possible, from document research and interviews with IMF staff and country authorities. The evaluation is only a partial application of the DQAF, as it is difficult to assess some aspects based on the available evidence.

The evaluation covers 74 LICs, but many of the assessments here apply to the IMF's statistical system as whole. It should be noted that the IMF does not have a statistical system in the sense of a national statistical system, and it inherits some of the problems related to data sources and methodology that occur at the country level. Nevertheless, it is useful to see to what extent the IMF is able to conduct quality checks on the statistics it collects, uses and disseminates, much as a national statistical office would when it depends on reports from district offices. (The key difference is that the IMF has a very limited mandate regarding statistics and does not have the same legal provisions that some of the national statistical offices have through statistical acts.)

This evaluation uses the same codes as the ROSC Data Module uses: Observed (O), Largely Observed (LO), Largely Not Observed (LNO), or Not Observed (NO), using boldface abbreviations. Where necessary, a short explanation is provided in italics.

0. Prerequisites of Quality

0.1. Legal and Institutional Framework

> 0.1.1. The responsibility for collecting, processing, and disseminating the statistics is clearly specified.
>
> > **NO**: *The IMF has multiple systems for collecting and disseminating data, varying across and within departments. For example, country desk economists have ad hoc and unspecified duties and different perceived responsibilities for collecting and reporting data and assuring consistency of data in processing.*
>
> 0.1.2. Data sharing and coordination among data-producing agencies are adequate.
>
> > **NO**: *There are no clear procedures for how the many datasets within the IMF are shared among economists within or across departments. There are conflicting observations on the same countries in the same years for the same data series for data that are central to the IMF's surveillance, as found in the IFS, the WEO, the REO, and AIV staff reports.*

0.1.4. Statistical reporting is ensured through legal mandate and/or through measures to encourage response.

LO: *The IMF has a mandate to collect "minimum necessary" information, but largely relies on voluntary data submissions from its member countries for most of the data required for surveillance. The IMF has developed data dissemination standards that have helped to encourage the timely submission of such data from the membership.*

0.3. Relevance

0.3.1. The relevance and practical utility of existing statistics in meeting users' needs are monitored.

NO: *Preliminary findings from a survey of IMF staff indicate that they do not primarily (if at all) use IMF-provided statistics but rely on data directly from member countries or commercial data providers for analysis.*

0.4. Other Quality Management

0.4.1. Processes are in place to focus on quality.

LNO: *By the IMF's admission, it does not have procedures in place to vouch for the quality of its statistics. Some instruments do exist (simple validation checks, data ROSCS, evaluation of adequacy of data for surveillance), but these are not consistent or universally applied.*

0.4.2. Processes are in place to monitor the quality of the statistical program.

LNO: *The IMF's statistical system has not yet been fully reviewed. Reviews of parts of the program are done on an ad hoc basis.*

1. Assurances of Integrity

1.1. Professionalism

1.1.1. Statistics are produced on an impartial basis.

LO/LNO: *LO refers to the statistics "produced" by STA, as these are member countries' own national statistics, with some objective validation checks applied. Thus, such statistics are impartial in so far as the national statistics are impartial. LNO refers to statistics in area departments, in that concerns have been raised at times about the objectivity of statistics at the country level. Staff estimates may not be grounded in source data and statistical methodology and may be subjective*

corrections of country estimates that could be deemed to be biased. Staff estimates are much more frequently required for LICs to fill in for missing statistics.

1.1.2. Choices of sources and statistical techniques as well as decisions about dissemination are informed solely by statistical considerations.

LNO: *Which sources and techniques are used are determined by availability. In area departments, economic considerations might outweigh statistical considerations in making adjustments to data.*

1.1.3. The appropriate statistical entity is entitled to comment on erroneous interpretation and misuse of statistics.

LO: *While statistical entities are entitled to comment on or challenge IMF-provided statistics, procedures for consultation between staff and country authorities are not consistent. Earlier reviews have called for clearer explanations of why staff estimates may differ from country statistics.*

1.2. Transparency

1.2.1. The terms and conditions under which statistics are collected and disseminated are available to the public.

LNO: *Findings from interviews and surveys indicate that external users of IMF-provided data are typically unaware of the reasons for the differences between the data in the IFS, WEO, and Article IV databases. Furthermore, it is often impossible to distinguish between official statistics and staff estimates in the WEO and AIV data.*

1.2.2. Internal governmental access to statistics prior to their release is publicly identified.

NO: *The extent to which these numbers are discussed and negotiated between the IMF and country authorities before they are released is not recorded.*

1.2.3. Products of statistical agencies/units are clearly identified as such.

LNO: *In the IFS, the number clearly comes from the authorities, but other databases often do not clearly identify whether it is an official country statistic or a staff estimate.*

2. Methodological Soundness

2.1. Concepts and Definitions

2.1.1. The overall structure in terms of concepts and definitions follows internationally accepted standards, guidelines, or good practices.

> **O:** *The IMF Statistics Department, in cooperation with other agencies, maintains and defines these internationally accepted reporting standards.*

2.2. Scope

2.2.1. The scope is broadly consistent with internationally accepted standards, guidelines, or good practices.

> **O:** *The IMF Statistics Department, in cooperation with other agencies, maintains and defines these internationally accepted reporting standards.*

2.3. Classification/Sectoralization

3.3.1. Classification/sectoralization systems used are broadly consistent with internationally accepted standards, guidelines, or good practices.

> **O:** *The IMF Statistics Department, in cooperation with other agencies, maintains and defines these internationally accepted reporting standards.*

2.4. Basis for Recording

2.4.1. Market prices are used to value flows and stocks.

> **LNO:** *The mean observation in country data ROSCs of LICs.*

2.4.2. Recording is done on an accrual basis.

> **LNO:** *The mean observation in country data ROSCs of LICs.*

2.4.3. Grossing/netting procedures are broadly consistent with internationally accepted standards, guidelines, or good practices.

> **LNO:** *The mean observation in country data ROSCs of LICs.*

3. Accuracy and Reliability

3.2. Assessment of Source Data

3.2.1. Source data—including censuses, sample surveys, and administrative records— are routinely assessed, e.g., for coverage, sample error, response error, and non-sampling error; the results of the assessments are monitored and made available to guide statistical processes.

LNO: *The IMF does not evaluate or collect source data comprehensively.*

3.5. Revision Studies

3.5.1. Studies and analyses of revisions are carried out routinely and used internally to inform statistical processes.

NO: *The IMF does no revision studies in its databases.*

4. Serviceability

4.1. Periodicity and Timeliness

4.1.1. Periodicity follows dissemination standards.

LNO: *Flagship publications follow a standard, but datasets may be updated in an irregular fashion.*

4.1.2. Timeliness follows dissemination standards.

LNO: *Flagship publications follow a standard, but datasets may be updated in an irregular fashion.*

4.2. Consistency

4.2.1. Statistics are consistent within the dataset.

NO: *Sources and methods used in country statistics vary considerably*

4.2.2. Statistics are consistent or reconcilable over a reasonable period of time.

LNO: *Sources and methods used in country statistics vary considerably.*

4.2.3. Statistics are consistent or reconcilable with those obtained through other data sources and/or statistical frameworks.

NO: *There are many gaps, inconsistencies, and discrepancies.*

5. Accessibility

5.1. Data Accessibility

5.1.1. Statistics are presented in a way that facilitates proper interpretation and meaningful comparisons (layout and clarity of text, tables, and charts).

O.

5.1.5. Statistics not routinely disseminated are made available upon request.

NO: *Such service is not available to individual users.*

5.2. Metadata Accessibility

5.2.1. Documentation is available on concepts, scope, classifications, basis of recording, data sources, and statistical techniques, and differences from internationally accepted standards, guidelines, or good practices are annotated.

LNO: *This varies considerably across databases and countries.*

5.2.2. Levels of detail are adapted to the needs of the intended audience.

LNO.

5.3. Assistance to Users

5.3.1. Contact points for each subject field are publicized.

NO: *Such information is not directly available.*

5.3.2. Catalogues of publications, documents, and other services, including information on any charges, are widely available.

O: *Such information is available.*