

The Current State of Play

11. The Fund has taken steps to address some of the challenges regarding data, with significant recent efforts in this regard. But will these be sufficiently holistic and well-funded to be sustainable? Will they support the organization in being proactive, not just reactive, in the realm of data?

12. To address such questions, the following sections present the evidence gathered for this evaluation on: (i) data issues pertaining to the Fund’s core strategic functions of surveillance and lending; (ii) efforts to address information gaps; (iii) the quality of data; (iv) internal data management practices; and (v) data dissemination and international cooperation. The key theme throughout is how data support the Fund’s strategic operations, and the sections are prioritized accordingly; for example, good data management is a means, not an end in itself, to better enable data to play its role in surveillance and lending.

A. Meeting the Fund’s Core Operational Needs

The global financial crisis changed the Fund’s approach to surveillance and created a surge in demand for more data in new areas.

13. The crisis underscored the importance of mainstreaming macrofinancial analysis into bilateral surveillance and better integrating bilateral with multilateral surveillance. It thus provided considerable motivation for revamping the IMF’s toolkit for detecting macrofinancial risks and risks associated with global interconnectedness.

14. The resulting, more integrated conceptual frameworks all depend heavily on data, making it increasingly difficult to disentangle the data issues related to the three main branches of surveillance—bilateral, multilateral, and financial¹²—and to lending. Each of

them, to varying degrees, face the fundamental data dilemmas of trade-offs between: accuracy versus timeliness, granularity versus aggregation, international comparability versus country specificity, and confidentiality versus transparency. And according to the IEO’s survey of IMF staff, each of these core operations is adversely affected by data deficiencies (Figure 1).¹³ Despite considerable overlap, the following discussion considers separately the three types of surveillance and also lending, as each poses some unique data issues.

(i) Bilateral surveillance

The fundamental question is whether data are adequate for surveillance . . .

15. Bilateral surveillance¹⁴ is the cornerstone of the Fund’s operational work. The workhorse accounting structure underlying this surveillance is the financial programming framework, an integrated macroeconomic framework that demonstrates how the data for a country’s various economic sectors—real, monetary, fiscal, external—are interlinked, allowing the Fund to construct a picture of the overall economy.¹⁵ But any analysis based on this framework can only be as good as the data supporting it, which will also reflect the approaches used by IMF staff to address data gaps and inconsistencies.

is technically not an independent, third “branch” of surveillance, but rather, as articulated under the Integrated Surveillance Decision, an integral part of both bilateral and multilateral surveillance. Nevertheless, in practice, the IMF has often treated financial surveillance as a separate entity. See, for example, IMF (2012c).

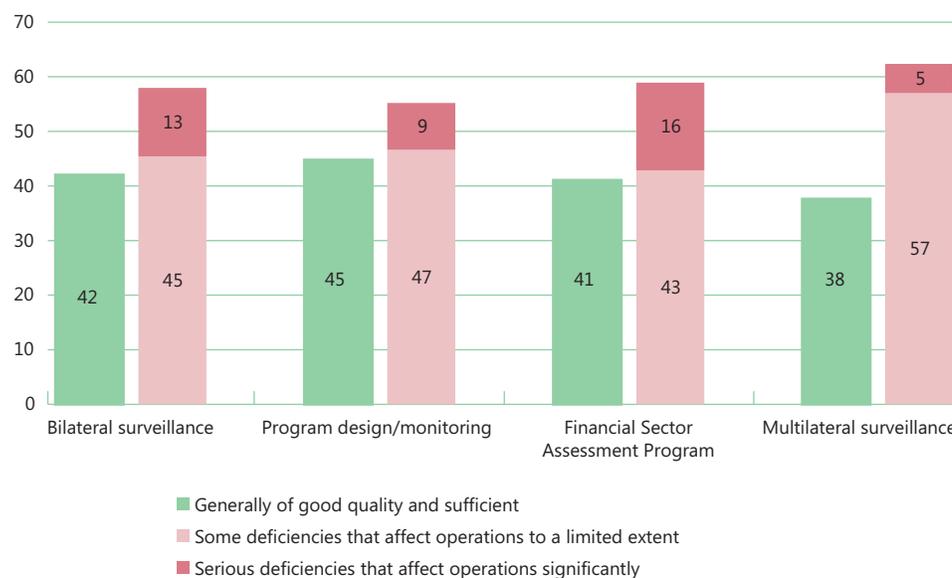
¹³Although this report focuses on surveillance and lending, data deficiencies also can have a bearing on other important areas of Fund work, such as calculating quota shares to guide decisions regarding relative size and distribution of members’ actual quotas.

¹⁴Most notably, the Article IV consultations that the IMF conducts (typically) on an annual basis with each of its member countries.

¹⁵A shortcoming of the financial programming framework is that the financial sector is still not fully integrated into the framework.

¹²The Articles of Agreement only recognize two forms of surveillance—bilateral and multilateral. Thus, financial surveillance

Figure 1. Staff Perceptions of Data Quality and Availability for Operations
(In percent)



Source: IEO Survey.

16. The Greek crisis provides a compelling illustration of the importance of high-quality data for IMF surveillance—and for global economic stability. Greece’s debt crisis erupted in late 2009, when a new government revealed that the projected fiscal deficit and government debt had been grossly understated by the previous government. This disclosure alarmed financial markets, ultimately precipitating a “sudden stop” of financial inflows and the need for a bailout. But this was not the first time that Greece’s data had been found wanting: a 2004 Eurostat report showed that Greek government deficit and debt figures had been misreported as far back as 1997, and that the deficit had not been below the Maastricht limit of 3 percent in any of these years (Eurostat, 2004). How did the Fund miss the warning signals of problematic data?

... and the IMF staff’s answer is a qualified “maybe.”

17. Greece has not been alone with respect to flawed data. Data deficiencies have adversely affected the bilateral surveillance of all categories of countries—advanced, emerging, and low-income countries (LICs), albeit to different degrees (Figure 2)—with almost 60 percent of IMF staff survey respondents noting such deficiencies regarding their primary country assignment. Lack of data or inadequate quality were each cited by about 90 percent of these survey respondents.

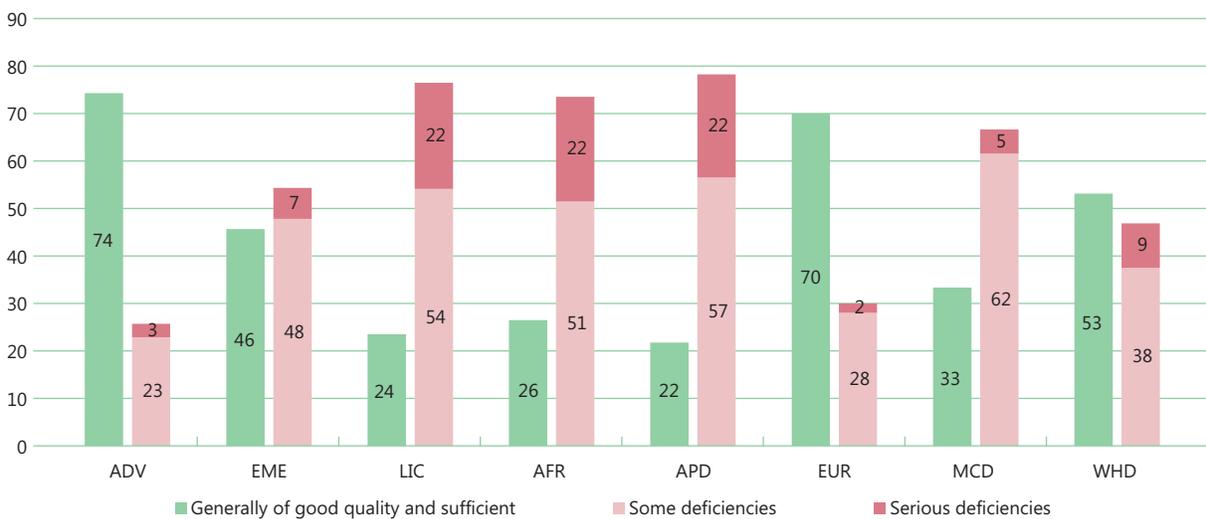
These findings on data deficiencies and the adverse impact on surveillance reinforces those of the Fund’s 2014 and 2011 Triennial Surveillance Reviews and various past IEO evaluations (Annex 4).

18. A number of cases have been documented in which problematic data reporting hampered the Fund’s conduct of surveillance and led to faulty analysis. Reichmann and Monasterski (2016) discuss about a dozen such country cases that have arisen since the 1990s.¹⁶ It is also highly likely that many data-induced shortcomings have left no traces in Fund documentation, and that in most such cases, the Fund could not have detected data problems that might affect its analysis, absent the explicit admission of the member country. Australia’s Bureau of Statistics, generally considered among the best, is one notable example. It admitted to a benchmarking issue in its 2010 and 2011 official employment figures, overstating the strength and the weakness, respectively, in the labor market. This led to perceptions that the Reserve Bank’s decision to push up rates in 2010 and to reverse course the following year could have been influenced.¹⁷

¹⁶Most cases where the Fund has documented data that have undermined analysis have occurred in the context of Fund-supported programs, reflecting the much greater attention the Fund gives to data when its own financial resources are at risk.

¹⁷*Sydney Morning Herald*, July 3, 2012. See also on this issue: <http://www.abs.gov.au/AUSSTATS/abs@nsf/Previousproducts/6202.0>Main%20Features2Apr%202012?opendocument&tabname=Summary&prodno=6202.0&issue=Apr%202012&num=&view> and <http://www.rba.gov.au/publications/smp/2012/aug/box-e.html>.

Figure 2. Staff Perceptions of Data Quality and Availability for Bilateral Surveillance
(In percent)



Source: IEO Survey.

19. The most common reason for data deficiencies, according to the survey of staff, is a country's limited capacity (including cost constraints), but a more troubling reason, cited by close to 20 percent of staff survey respondents, is the authorities' unwillingness to provide the data. While in some cases non-provision was due to cost considerations, more than half of such instances were due to confidentiality concerns about how the IMF would handle the data. The survey of data providers also indicated a strong regional component, with about 40 percent of respondents from Asia and from Middle Eastern oil-exporting countries expressing concerns about confidentiality. Worse still, 10 percent of IMF staff (with higher numbers for those working on emerging markets) claimed that intentional manipulation of data was responsible for data inadequacies.¹⁸

The IMF has a broad-ranging toolkit to address data deficiencies . . .

20. What instruments does the IMF have at its disposal to question official statistics and to address data deficiencies during the conduct of bilateral surveillance? In addition to staff judgment and experience,

¹⁸Intentional manipulation is often a case of Goodhart's Law, the popular formulation of which is "When a measure becomes a target, it ceases to be a good measure." Goodhart's Law (named after an economist who was a member of the Bank of England's Monetary Policy Committee) refers to the vulnerability of a statistical indicator to manipulation once it is used to define a policy target.

data inconsistencies are often discovered through the use of the IMF's financial programming framework.¹⁹ Problems can also be detected by checking flow data against stock data. In about half of country cases with data deficiencies, staff survey respondents said that they had to come up with their own estimates for the problematic data.

21. In the context of Article IV consultations, IMF staff are expected to candidly assess the adequacy of member countries' statistics for surveillance (IMF, 1995a), with major deficiencies discussed in the main text of the Article IV staff report, along with a more detailed review in a Statistical Issues Appendix (SIA). This guidance is aimed at raising the profile of data issues in surveillance and to prompt corrective action if warranted, with staff proposing remedial measures or technical assistance, if needed.

. . . but doesn't always deploy it effectively

22. The Fund's regular reviews of data provision had identified a number of problems with the SIA arrangement (Box 1):²⁰ (i) lack of candor in staff's discussion of data deficiencies, with an "upward bias in the

¹⁹Until recently, financial programming was typically not applied to advanced economies, a factor which may have contributed to the undetected buildup of the large imbalances prior to the financial crisis.

²⁰The system currently in place was approved and reviewed, respectively, during the 2008 and 2012 reviews of data provision to the Fund for surveillance (IMF, 2008 and 2012b).

Box 1. A, B, or C? Grading a Country's Data Adequacy for Surveillance

The IMF's current framework for data provision for surveillance was first defined in 1995 (IMF, 1995b), with some amendments since then. One of the elements that was added to Article IV consultations was a Statistical Issues Appendix (SIA), which includes an overall judgment on the adequacy of data provision for surveillance and, where relevant, a discussion of the implications of data deficiencies and recommendations for improvement. How candid have these judgments been?

Initially, assessments in SIAs included only two categories: adequate or inadequate. In 1995, 59 percent of a sample of 50 member countries were deemed to have "adequate" data provision.¹ This number climbed to 70 percent by 2003. In 2005, an intermediate category was added, allowing the following options for assessing data provision:

- Adequate for surveillance (A)
- Broadly adequate for surveillance, but with some shortcomings (B)
- Inadequate for surveillance (C).

By 2007, the sample percentage judged as either category A or B jumped to 90 percent, with only 4 percent

¹Note that the Article IV consultations for those deemed to have inadequate data provision were still completed.

given a category C rating (another 6 percent were unclear as to rating). Yet in a survey of mission chiefs that was conducted at the time, more than half of the respondents noted problems with data provision that hampered surveillance, with 40 percent reporting that their teams routinely had to prepare estimates for key data. Why this discrepancy between SIA assessments and mission chief responses?

The IMF's 2008 Review of Data Provision (IMF, 2008) concluded that the relatively rare use of Category C may have stemmed, in part, from concerns that it would undermine the relationship with country authorities or would raise questions about how surveillance can be conducted at all if data are "inadequate." This led to more changes in the framework, including elimination of the term "inadequate" from category C, recasting it as "Data provision has serious shortcomings that significantly hamper surveillance."

Did this change improve candor? By the time of the 2012 Review of Data Provision (IMF, 2012b), the percentage of countries classified as C had increased threefold to 12 percent. But a supporting survey of mission chiefs still indicated a huge discrepancy, with 59 percent of the respondents stating that important data deficiencies had hampered surveillance, thus suggesting continued reluctance to use the lowest rating. In response, new guidelines were issued in 2013 (IMF, 2013a) to encourage staff to use more candor and to provide clearer instructions to staff on the classifications.

characterization of data adequacy;" (ii) excessive workload on staff, deriving from the requirement to document and propose remedial measures to address data issues; (iii) poorly focused SIAs, with limited coverage; and (iv) lack of attention by the Executive Board.²¹ These problems were confirmed by the evaluation's survey and interviews; for example, according to the staff survey, formal data adequacy assessments are softened, as only 46 percent of cases in which data are perceived as inadequate are reported as such in SIAs. In response to the 2012 Review of Data Provision to the Fund (IMF, 2012b), the Fund issued a guidance note (IMF, 2013a), updating and clarifying how staff are to address any data shortcomings in the Article IV report. The updated guidance note aimed, in part, at improving compliance with the intent of the SIA, but—as discussed in [Annex 5](#)—little appears to have changed since it was issued.

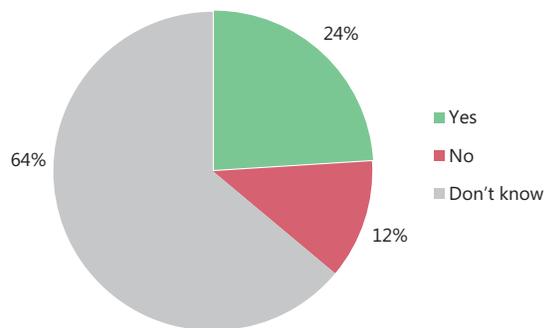
²¹As senior IMF staff members pointed out to the evaluation team, the Board's "lack of attention" to data quality issues at times reflected peer protection and political considerations.

23. Perhaps the most serious indictment of the SIA is its relative obscurity. Neither the Board nor IMF staff pays much heed to the SIA,²² with more than half of staff survey respondents noting that country teams lacked the resources and time to make thorough assessments. More worrisome, though, are the survey results of country authorities (i.e., data providers), fewer than one-quarter of whom were familiar with the SIA for their own country ([Figure 3](#)). This implies that the SIA does not provide the intended incentive for countries to improve their data.²³

²²As an example, for the 2007 United States Article IV consultation, the SIA noted that "Coverage of international capital flows in external sector statistics has been improved, with the June 2007 releases of BOP and IIP data on financial derivatives." This identical statement, highlighting 2007 data, appeared in the SIAs from 2008 until 2014, when an attentive staff member finally changed the date to June 2014. Of course, the U.S. SIA was not alone in conveying incorrect information. This evaluation found errors in a number of SIAs, as confirmed by country authorities during interviews.

²³Interviews with country authorities showed that a major reason for their lack of familiarity with the SIA was its issuance in a separate supplemental document for the Board meeting. Most of the authorities only read the main section of the Article IV report.

Figure 3. Did Your Country's Last Article IV Report Include a SIA?
(In percent)



Source: IEO Survey of Country Authorities and Data Providers.

24. The IMF has some legal scope to question countries on issues of data accuracy and availability, as embodied in the policies on misreporting and breaches of obligations under Article VIII. Potential breaches of obligations in the context of surveillance have occurred with some frequency; the 2012 Review (IMF, 2012b) noted that, in the preceding four years, “sustained concerns were raised with eight members about their willingness to share data required for Fund surveillance to the best of their ability.” Seven of those cases referred to nonprovision of data and were resolved within a year, while one related to provision of inaccurate data and resulted in the Board issuing a decision of censure, calling on the member to adopt remedial measures. This case—Argentina—remains unresolved (Box 2).²⁴

²⁴The most recent Board meeting on this breach of obligations was held in May 2015, with no change in the stance adopted by the Fund.

Box 2. Argentina and the Breach of Obligations

In January 2007, the Argentine government changed the personnel in charge of producing the consumer price index (CPI) at the National Statistics and Census Institute (INDEC). Concerns about the integrity of the CPI started to be voiced soon thereafter.

These concerns led several private sector entities and provincial governments to compute their own indices that showed consumer prices growing at significantly higher rates than those reported by INDEC. For example, Cavallo (2013), using data collected by the Billion Prices Project at MIT—such as prices in major supermarkets available on line between October 2007 and March 2011—replicated the components and weights of the CPIs in five Latin American countries (Argentina, Brazil, Chile, Colombia, and Venezuela). He found that, while the online price indices for the other four countries tracked well both the level and dynamic behavior of inflation, in Argentina they exceeded the official index by a factor of about three.

The apparent underreporting of CPI has implications for other key variables of significant importance for economic analysis. Inasmuch as the official CPI enters their calculation, measures of poverty or of the real effective exchange rate would be underestimated while the real growth of the economy would be overestimated. Underreporting would also have notable financial implications given Argentina’s issuance of inflation-linked peso bonds.

Failure to provide information to the Fund (i.e., a breach of obligation under Article VIII, Section 5) is defined to cover both non-provision of information and provision of inaccurate information (IMF, 2004b). The latter failure is considerably more difficult to substantiate than non-provision

of required information. This caused the Fund to take a measured approach to Argentina’s case. Eventually, in July 2011, Argentina was found in breach of its obligation due to inaccurate reporting of official data for the CPI and GDP. Technical assistance was offered, which resulted in several recommendations to correct the known deficiencies. However, delays in the implementation of key remedial measures led the Executive Board in September 2012 to issue a Statement of Concern, followed by a Declaration of Censure in February 2013. The latter established a timeframe for the adoption of the remedial measures, noting that failure to follow this timeframe could result in a declaration of ineligibility to use the general resources of the Fund.

Over the following two years, Argentina took some measures to address the difficulties, for example, introducing a new *national* CPI (the previous one was limited just to Buenos Aires) and revised GDP data (now with base year in 2004) in early 2014. These actions, however, did not fully assuage the Fund’s concerns, and further actions—related to the transparency of the process—were required before the Executive Board could withdraw the Declaration of Censure.

Argentina was an early subscriber to the SDDS and never lost that status. While the SDDS is a dissemination standard, not a quality standard, the Fund does issue an annual observance report as a form of monitoring, covering, among other metadata dimensions, the integrity and quality of the data. Argentina’s reports for 2012 through 2014 still indicated that Argentina met the integrity dimension, while the discussion of quality relegated the information on the Declaration of Censure to a footnote.

25. Greece is a timely case study as to how well (or how poorly) the IMF used its toolkit to deal with data deficiencies. In its assessment of Greek statistics in the run-up to the disclosures of false data, the Fund had been rather sanguine, with only occasional expressions of mild concern (“... data are adequate for surveillance but should be strengthened” was a common refrain). Not only was surveillance inadequate in this regard, but Greece had engaged in several important statistical milestones with the Fund: Fiscal Transparency ROSCs in 1999 and 2006, an SDDS subscription in 2002, and a full data ROSC in 2003—none of which brought to light the seriousness of the data problems. In 2010, in conjunction with Board approval of the Fund’s initial IMF-supported financial program with Greece, the Managing Director issued a report to the Board (IMF, 2010) on a breach of obligations under Article VIII, Section 5. The Board determined that Greece had taken sufficient remedial actions, including enacting a new law granting independence to the national statistical agency (ELSTAT). Yet—as discussed in [Box 3](#)—the independence of ELSTAT remains a concern five years later.

(ii) Multilateral surveillance

The perennial dilemma for multilateral surveillance data is international comparability versus country specificity . . .

26. Multilateral surveillance, always an important component of the Fund’s operations, took on an even larger role with the Fund’s adoption of the Integrated Surveillance Decision in 2012. This decision made the Article IV consultations a vehicle for both bilateral and multilateral surveillance, and helped to push forward work on policy spillovers and interconnectedness. In the context of this wider scope, data needs have grown markedly.

27. Multilateral surveillance (and cross-country analysis even more so) poses a special challenge for data, as it is predicated on comparability across countries²⁵—that is, on the same concept being defined and measured in the same way everywhere. But global standards do not necessarily suit local conditions. Particular country circumstances unavoidably result

²⁵In contrast to cross-country analysis, multilateral surveillance, which often focuses on spillovers and interconnections, does not always necessitate perfectly standardized cross-country datasets.

in different definitions, measurements, or coverage of economic variables, implying that concepts can be homogeneous across countries only to a certain degree. How can the IMF ensure that it is not “comparing apples and oranges” in its multilateral and cross-country work? And what does the analysis mean if the data are not fully comparable?

28. The IMF’s work on methodology and capacity development in the area of statistics has gone a long way to strengthen comparability. This is particularly true for the databases maintained by STA, which emphasize data that meet methodological standards. But the main sources of data for much of multilateral surveillance are area departments, where data are more likely to conform to country specificities or be based on staff estimates.

29. IMF staff recognize the challenge posed by lack of comparability. According to the IEO’s survey of staff, almost two-thirds of those engaged in multilateral surveillance claim that data deficiencies hamper surveillance to some degree, with lack of comparability across countries overwhelmingly cited as the main reason.²⁶ In sharp contrast to IMF staff views, the IEO survey of external data users indicated that almost 90 percent believed IMF data are comparable across countries,²⁷ a misperception that could pose a reputational risk to the Fund.

30. Problems with non-comparability have been highlighted in some IMF work. A notable example is Dippelsman, Dziobek, and Gutiérrez Mangas (2012), which underscores how failure to follow international guidelines for reporting of public sector debt (arguably one of the most important macroeconomic indicators) or inadequate documentation of data definitions “can lead to major misunderstandings in the fiscal policy debate.”²⁸

31. The present evaluation also considered comparability of data by examining the definitions of government deficit that were used for performance criteria in the 48 IMF-supported programs approved from January 2011 to April 2015 ([Annex 6](#)). The combination of different components resulted in nine different definitions in terms

²⁶The importance of comparability was confirmed by the 2014 Triennial Surveillance Review (TSR) survey of IMF mission chiefs; when asked to check those factors most important for increasing the use of cross-country studies in surveillance, 85 percent chose greater availability of *comparable* cross-country data.

²⁷By a slight margin, *World Economic Outlook (WEO)* data are (wrongly) believed to be more comparable than those of *International Financial Statistics (IFS)*.

²⁸The authors use Canada as an example to illustrate how different definitions of the public sector give rise to very different debt levels, with debt-to-GDP ranging from 38 percent on a narrow budgetary definition to 104 percent, using the consolidated general government.

Box 3. Greece: Policy-Based Evidence-Making and the Perils of Statistics

In October 2009, the Greek authorities disclosed to Eurostat that government deficit and debt data for 2005–09 needed to be revised. The revisions, completed in November 2010, were of an exceptional scale and resulted in the forecast deficit for 2009 moving from 3.7 percent of GDP to 15.4 percent of GDP, while the government debt moved from 99.6 percent of GDP to 126.8 percent of GDP.

The revisions reflected methodological weaknesses and unsatisfactory technical procedures in the Greek statistical system, but also inappropriate governance as exemplified by lack of clear responsibilities between institutions, diffuse personal responsibilities, and opaque empowerment of officials “which left the quality of fiscal statistics subject to political pressures and electoral cycles” (European Commission, January 2010). The contemporaneous Fund report on Breach of Obligations under Article VIII, Section 5 (IMF, 2010) stated that “the institutional setting at the time failed to ensure the independence and accountability of the National Statistical Service of Greece and other services involved in the production of fiscal data and public debt data.”

The problems reported in October 2009 were not new; in fact, both Eurostat and (to a far lesser degree) IMF staff had repeatedly indicated that Greek statistics were notoriously weak and plagued with problems. A 2004 report by Eurostat triggered “the first Greek data crisis” by showing that Greek government deficit and debt figures had been misreported since as far back as 1997, and that in none of these years had the deficit been below the Maastricht limit of 3 percent per year. Subsequently, as noted in Eurostat’s 2010 report, Greek government deficit and debt statistics were the subject of “continuous and unique attention for several years.”

On its part, Fund staff took a generally approving stance with only occasional expressions of mild concern. Congratulations were offered on the occasion of Greece completing the Fiscal Transparency ROSC in 1999, subscribing to the SDDS in 2002, and completing a data ROSC in 2003. On the latter, staff observed (IMF, 2003a) that: “. . . Statistics-producing agencies in the main have a legal and institutional environment that supports statistical quality. . . . All agencies demonstrate professionalism and are transparent in their practices and policies. In particular, the strong laws protecting confidentiality, rules for civil servants, and internal regulations of the central bank provide a clear set of ethical standards for staff. . . .” By 2006, in the Fiscal Transparency ROSC that was specifically prepared after the 2004 data crisis, staff was still maintaining a positive line: “Greek budget processes give assurances of

integrity about fiscal data through independent audit and recently strengthened statistical reporting.” In most consultation reports, staff took the general line of “. . . data are adequate for surveillance but should be strengthened” (e.g., the 2006 and 2007 Article IV consultations). Notwithstanding staff’s generally accommodating attitude, muted concerns about data weaknesses and calls for “further improvements” were an almost constant feature of consultation reports. Only by 2009, on the eve of the government’s acknowledgment of data deficiencies, did staff take a more forceful line, with the Article IV consultation for that year including a quite specific and detailed list of failings in Greek statistics.

Admittedly, even in the best statistical systems, it can be difficult to uncover truth when those in charge are bent on hiding it. Moreover, analyses may be unduly obstructed by insufficient financial sector data—as bemoaned in the 2005 consultation report—by differences across sectors in the coverage or definition of variables, by the complexity of intergovernmental fiscal relations in Greece, or by opaque financing activities—such as the off-market swaps in which Greece frequently engaged—but a more thorough application of the financial programming framework should have allowed staff to get an inkling of the sizable ongoing irregularities.

IMF staff had on several occasions (viz., the 2005 and 2006 consultations) called for granting independence to Greece’s national statistical service. This finally came about when the creation of an independent new office, ELSTAT, was made a condition of the 2010 program and part of the remedial action proposed by the authorities subsequent to the May 2010 report on breach of obligations. Under a new chief statistician, Greek government finance statistics were accepted by Eurostat without reservation in 2011–15, in contrast to the repeated reservations of the previous years—indicating a marked improvement in the quality of Greek statistics. However, ELSTAT’s independence continued to be challenged by vested interests (e.g., a criminal investigation was launched in 2013 against the chief statistician regarding revisions to historic data on public finances and debt), raising doubts about the underlying commitment of the country to truly independent statistics and pointing to risks of re-politicization in the future. Five years after being set up, ELSTAT’s independence was still not assured, as suggested by the Euro Summit of July 12, 2015 when “. . . Given the need to rebuild trust with Greece . . . safeguarding of the legal independence of ELSTAT . . .” needed to be included among the required measures.

of coverage, a heterogeneity that was further magnified by measuring the criterion on a cash or accrual basis, and in above- or below-the-line terms. This wide variety of concepts often carried over to the data reported in the *WEO*, thus putting paid to the notion that the numbers included in *WEO* are strictly comparable.^{29,30}

... highlighting the importance of having countries adopt standard data templates and—in the meantime—providing clear metadata.

32. These findings underscore the importance of providing clear metadata for all IMF-disseminated data. However, notwithstanding the IMF's guidelines to staff to this effect, an examination of a large sample of Article IV consultation reports showed that none of them complied with this requirement (*Annex 5*). An inattentive or impatient economist could simply download the desired data, compare apples and oranges, and draw the wrong policy conclusions.³¹ Nevertheless, even with excellent metadata, the diversity of definitions can greatly impair the ability to do cross-country work.

(iii) Financial surveillance

Data issues for financial surveillance are among the most challenging ...

33. Financial sector surveillance, in the aftermath of the global financial crisis, has become even more central to the Fund's core operations. However, data issues are particularly challenging here, given the sensitive (and often confidential) nature of the data, the need for granularity and comprehensiveness (e.g., “off-balance-sheet” exposures), and lack of consistency. Data are often

²⁹Nominal GDP provides another example of comparability issues in *WEO* data. While most countries still measure GDP using the 1993 System of National Accounts (SNA), some, including most of the advanced economies, have now moved to the 2008 SNA. Typically, GDP, as measured under the 2008 SNA, is larger than that under the older system (e.g., U.S. nominal GDP was almost 4 percent larger, while it is estimated that, were China to move to the newer system, its economic size could be as much as 16 percent larger).

³⁰The *WEO* makes adjustments to some data to improve comparability. For example, the *WEO* has migrated balance of payments data to the methodology used in the sixth edition of the *Balance of Payments and International Investment Position Manual (BPM6)*, even though many countries still submit data under the previous *BPM5* methodology. The *WEO* also converts data on a fiscal-year basis to a calendar-year basis.

³¹In interviews with external data users, many admitted that they use multiple (noncomparable) IMF data sources (*IFS*, *WEO*, country reports, Working Papers) to fill in missing data for cross-country studies.

nonexistent or opaque in some critical areas, particularly on cross-border linkages and the shadow banking sector.

... due, in large part, to the often market-sensitive nature and need for granularity of data.

34. Financial surveillance is constantly struggling with the tension between granularity and aggregation. Aggregate data can mask critical vulnerabilities—that granular data might reveal—and may not be usable with some of the Fund's new analytical tools. For example, network analysis (used to examine issues of interconnectedness) needs quite granular data. But the Fund's hands are essentially tied by its Articles of Agreement, as it cannot require countries to provide institution-specific data.

35. The Financial Stability Assessment (FSA), a component of the FSAP and a key instrument of the IMF's surveillance, illustrates the data challenges facing the IMF. According to this evaluation's survey of staff, the data collected for FSAPs are perceived as the most problematic.³² Just under a third of the survey respondents from the Monetary and Capital Markets Department (MCM) believed data were sufficient for conducting an FSAP exercise. Notably, almost 90 percent said data problems had hindered the conduct of stress tests, while about three-quarters said the analysis of potential cross-border spillovers was hampered by data problems. Poor quality data was cited by 40 percent of respondents.

36. A 2014 review of the FSAP (IMF, 2014d) made clear the role of data—in particular, availability and quality for stress testing—in underpinning (or undermining) the program's effectiveness. While many countries voluntarily provide these data to the FSAP team—subject to strict confidentiality protocols—the FSAP review noted that this practice is not universal (as confirmed by this evaluation's survey results), with advanced countries the least likely to share supervisory data. And even when the needed data are available, FSAP teams are typically not equipped to assess their accuracy or the quality of underlying assets.

... underscoring the importance of building trust, yet being candid about data limitations.

37. Considering the reasons why country authorities are loath to share data, about a third of MCM survey

³²Also, compared with staff working on the *WEO*, staff involved with the *Global Financial Stability Report* were much more likely to note problems with lack of data, comparability, and uncertain quality.

respondents cited legal constraints,³³ but another third cited issues of trust. Data providers who were interviewed regarding access to market-sensitive financial data noted that banking supervisors tend to trust, in order, other supervisors, central banks, the Bank for International Settlements (BIS), and only then the IMF.³⁴ Further, as noted by the 2014 Triennial Surveillance Review, “some of the Fund’s counterparts have become less willing to share data as the crisis has subsided.” This is clearly problematic for the Fund’s FSAP, which in such cases must rely on publicly available data and/or on stress tests conducted by supervisors and the banks themselves. According to some interviewees, the results from the FSAP team’s stress tests differed at times from those of the stress tests conducted by the authorities or the banks, largely because of differing access to data.

38. These findings highlight two issues: (i) there is a tension between the mandatory character of FSAs and the voluntary provision of the data they require,³⁵ and (ii) the limitations of the associated risk assessment need to be clearly communicated by the FSAP team. The evaluation survey of MCM staff is revealing in this regard, with 40 percent of respondents advocating mandatory data provision to help address data deficiencies, and only half agreeing that the Financial System Stability Assessment (FSSA) report had clearly noted the problems with data quality or access. To help address the first issue, the Fund could clarify its confidentiality protocols to the membership to encourage the voluntary provision of the needed information.³⁶ On the second issue, the 2014 FSAP review noted that the standard disclaimer on all FSSA reports should be expanded to highlight any data limitations. But this evaluation found no change in the standard disclaimer in the most recent FSSA reports, including some with serious data access and quality issues.

39. The Fund has made notable strides, nonetheless, with respect to data needed for financial surveillance.

³³Some countries with legal constraints find ways to allow the FSAP team to “access” the data without actually violating the law (e.g., letting the FSAP team into the room to watch the conduct of supervisory stress tests).

³⁴The 2013 IEO evaluation, *The Role of the IMF as Trusted Advisor*, also found that country authorities placed more trust in the BIS than the Fund in the handling of confidential data (IEO, 2013).

³⁵In September 2010, the Executive Board decided to make the Financial Stability Assessment (FSA) mandatory for systemically important financial sectors in response to the shortcomings revealed by the financial crisis. Previously, all FSAs, as part of an FSAP exercise, were conducted on a strictly voluntary basis.

³⁶See “Confidentiality Protocol—Protection of Sensitive Information in the Financial Sector Assessment Program,” IMF, *Selected Decisions*, Thirty-Second Issue, p. 108.

Substantive progress has been made since the global crisis on collecting data on Globally Systemically Important Financial Institutions (G-SIFIs), the nonbank financial institutions (NBFIs), shadow banking, and Financial Soundness Indicators (FSIs).^{37, 38} The Fund’s efforts to collect data on NBFIs and shadow banking is particularly important to allow the Fund to expand its coverage of stress tests to the nonbank sector (an increasingly important player in many countries’ financial sectors), and to help member countries limit regulatory arbitrage, a potential precursor for a future crisis. The Fund has also developed new analytical tools that benefit from the expanded set of financial data.

(iv) Use of Fund resources

Data deficiencies can affect program design and monitoring . . .

40. Data quality and availability are also extremely relevant for IMF lending. Staff must be able to count on information adequate to allow the design of a program fit for the intended purpose. This has usually been the case, but in some instances, staff has indicated that policy programs would have been formulated differently if more accurate information had been available (Reichmann and Monasterski, 2016) (Box 4). From 2000 through March 2015, there were 62 cases of misreporting vis-à-vis data in the context of Fund-supported programs,³⁹ up sharply from the nine cases in the previous 15-year period from 1985 to 2000. Occasionally, inaccurate or incomplete information about a member country’s observance of a program performance criterion may give rise to a “noncomplying purchase” and the issuance of a misreporting notification to the Executive Board.

41. Even when data allow for adequate diagnosis and formulation of policies, the specific design of performance criteria is influenced by considerations of data accuracy, availability, and timeliness. Trade-offs are unavoidable among these factors, and the resulting criteria will seldom be totally homogeneous across time

³⁷These datasets are part of the Data Gaps Initiative.

³⁸The improvement in the collection of FSIs is especially noteworthy, with 101 countries currently providing at least the core indicators as of mid-2015, compared with 57 in 2007. Nonetheless, FSIs notably suffer from lack of comparability across countries, as they are based on very heterogeneous definitions of capital, nonperforming loans, etc.

³⁹Of these 62 cases of provision of incorrect data, 11 were considered “de minimis,” 38 received waivers, and only 13 required corrective actions, usually involving early repurchase or repayment.

Box 4. Faulty Data and Faulty Analysis: Past Examples

Instances of data that subsequently prove to be wrong or incomplete are probably frequent, but are usually of little consequence and therefore go unreported. However, a number of cases of data-induced faulty analysis were documented in reports on breaches of obligations under Article VIII, Section 5 or misreporting in programs (Reichmann and Monasterski, 2016). The following are examples of the type of cases that can occur:

Hungary (1982–89)

In November 1989, the government revealed that both domestic and external debt had been underreported since the mid-1970s. The misreporting involved a misspecification of the net credit to the government and the consequent misreporting of monetary and balance of payments statistics as well as the public debt. In the February 2000 review of misreporting cases (IMF, 2000), staff stated that “. . . Hungary’s widespread, systematic and substantial misreporting of data clearly resulted in a fundamentally distorted view of the program by the staff. . . . Had the staff been aware of actual [developments] the program would not have been submitted for Board approval with the same quantified criteria. Had correct data been known, it would have at least affected the staff’s assessment of the size of corrective actions needed. . . .”

Jordan (1996–97)

Under an extended Fund arrangement, the authorities provided staff with erroneous information on national accounts and fiscal data. Revisions provided in mid-1998 indicated that GDP growth had been substantially lower

than first reported—around 1 percent per year instead of 5 percent—and, consequently, fiscal revenues had been substantially lower than reported. As a result, Jordan’s budget deficit had been higher and had to be financed by recourse to nonbank sources. In the 1999 Article IV consultation report, staff indicated that “. . . the data set that [had been] available had portrayed a fundamentally distorted picture of the state of the Jordanian economy and performance under the extended arrangement . . .” and “. . . staff [had been] working on the basis of a wrong view of economic developments in Jordan, which had a major impact on the assessment of performance. . . .”

Ukraine (1996–98)

Ukraine misreported the level of its international reserves continuously during 1996–98 and in the negotiations on a follow-up extended arrangement. The misreporting involved multiple transactions that impaired the liquidity of the foreign assets involved and, more egregiously, two “round-tripping operations” which artificially inflated the reserves. Days before the Board meeting on the requested arrangement, it was revealed that almost \$700 million of reserves was illiquid, leaving usable reserves of less than \$300 million—forcing an impromptu redesign of the program. The corresponding staff paper stated “. . . With the new information on Ukraine’s external reserve position, and the pressure in the market, the authorities have had little choice but to move the exchange rate band . . . the staff has reluctantly accepted the reimposition of the export surrender requirements . . . [and] further modifications of the program might be unavoidable. . . .”

or countries. Usually, the wider the coverage of a performance criterion, the better it reflects policy aspects that have a bearing on the program’s objectives. But wider coverage may run afoul of the availability and timeliness of the required data, forcing an inevitable narrowing of the criterion’s scope.⁴⁰ Over 60 percent of staff acknowledged the influence of data conditions in the formulation of performance criteria.⁴¹

⁴⁰This narrowing of the scope can have a critical impact on policy implications. For example, based on interviews with the relevant country authorities, the Fund missed about 25 percent of GDP in public debt, in a recent financial program, by failing to include data on public-private partnerships and state-owned enterprises.

⁴¹The same percentage of staff noted that the program included undertakings to improve data provision or quality.

... and performance criteria must often be tailored to fit the availability of data ...

42. Even the variables that data are intended to measure may differ across countries, reflecting the particular historical and political developments that determine a country’s institutional organization and hence the definition and scope of a given economic sector or instrument. This is particularly the case in regard to the concept of government or the public sector (Annex 6). The resulting differences in definition mean that a balance must be struck between the Fund’s need to treat members evenhandedly in the application of conditionality and its need to tailor performance criteria to fit the circumstances of each case.

... but an IMF-supported program can also help improve data quality and availability.

43. Often the existence of a program can have a reciprocal effect on the quality, timeliness, and availability of data. The due diligence that staff is required to do before including data in a performance criterion can result in the correction of data that are found wanting or in efforts to develop and provide the data needed. Such positive effects, plus the intersectoral consistency checks provided by the financial programming framework, have been felt more by developing and emerging economies than advanced economies,⁴² as the former have been more frequent users of Fund resources.

B. Addressing Information Gaps

The IMF's efforts to address data gaps have resulted in a significant expansion in data ...

44. By and large, the collaborative arrangements in place for data provision have served the Fund well, with most member countries providing data that far exceed those required under Article VIII. Even more so, since the global crisis, there has been a notable rise in the amount and breadth of data (much of which is in the financial realm) that member countries provide to the Fund. For example, 138 economies currently report monetary and financial statistics according to the IMF's standardized report form (SRF),⁴³ up from 83 as of end-2007.⁴⁴

45. Much of this strengthening of data provision is due to concerted efforts—on the part of the IMF (especially STA), other members of the IAG, and member countries—to address data gaps identified by the global crisis.⁴⁵ In particular, significant progress has been made in implementing the recommendations of the G20 Data Gaps Initiative (DGI); all G20 members and many

⁴²In fact, until the global economic and financial crisis with its origin in advanced countries, many desks on such countries did not use the financial programming or other macroeconomic framework to check for intersectoral data consistency. This became particularly evident when some member countries of the European Union (EU) came to the Fund for financial programs in the aftermath of the crisis.

⁴³Among STA's many databases, the SRF data are the most used by area department staff.

⁴⁴Notwithstanding this impressive progress, several G20 countries and other economies with systemically important financial centers still do not report with the SRF.

⁴⁵The Managing Director's Global Policy Agenda (IMF, 2015d) noted that closing data gaps should be a key area targeted by the Fund's capacity development activities.

non-G20 economies have enhanced their data provision to the IMF (IMF, 2014e), and efforts to collect a broader array of financial data (including FSIs) are also proceeding apace. Most of the associated conceptual work for the DGI has been completed, and more generally, the number and types of data-based analytical tools have expanded significantly.

... but do the benefits outweigh the costs?

46. After a crisis, data suddenly become a forethought, rather than an afterthought. This raises the question: were data gaps a core reason or a scapegoat for missing the recent global economic and financial crisis? The answer to this question is an important one, as it can help determine the direction for future surveillance. In fact, the failure to foresee the impending crisis cannot be attributed to lack of data (Box 5). With hindsight, it became clear that a substantial amount of existing data had pointed to growing vulnerabilities in several key areas.⁴⁶ Failure to foresee the crisis stemmed more from ignoring or misinterpreting these warning signals than from the absence of signals, a view shared by many of this evaluation's interviewees.

47. Nevertheless, filling in key data gaps could substantially strengthen surveillance. But this also comes with costs, particularly for those responsible for collecting or providing the new data. Almost three-quarters of the respondents to our survey of data providers believe that the benefits of the Fund's new data initiatives outweigh their costs, yet 40 percent felt the IMF was asking for too many data and almost half said the initiatives would pose a very heavy burden on reporters. This was particularly the case with the respondents from advanced countries, who are most affected by the new demands under the DGI. Among the respondents from low-income countries, only one in five indicated concern in this regard.

48. The proliferation of data and analytical tools also risks the possibility of the Fund failing to strike the right balance between collecting information and being able to process it efficiently and analyze the results. Indeed, while two-thirds of staff respondents to the survey indicated that the additional data from the new

⁴⁶IEO (2011a) notes, for example, that had the IMF conducted the Vulnerability Exercise for Advanced Countries prior to the crisis, using data that were available in 2006 would have pointed to the United States, United Kingdom, and Iceland as being at high risk of financial crisis.

Box 5. You Don't See What You're Not Looking For

The global economic and financial crisis generated a surge in the demand for new and better data. Yet were lack of data or inadequate data key factors behind the Fund's and others' failure to foresee the crisis? The answer would seem to be "no," based on the following:

- The Fund largely ignored some core data in the key financial centers that could have helped to signal a forthcoming crisis, including such traditional mainstays as broad measures of credit growth, leverage (household, government, corporate), and the growth of high-risk financial instruments.¹
- Despite lack of data supporting such a view, the Fund was overly enthusiastic about the soundness of U.S./U.K. financial systems and the risk-dispersing properties of financial innovation (including "exotic" mortgage products).
- The Fund believed it was sufficiently well-equipped with data to highlight the risks and vulnerabilities in emerging markets and developing countries, but did not use similar data trends to see similar risks in advanced countries.

¹In the U.S. staff reports, the standard table on "Indicators of External and Financial Vulnerability" did not include market-sensitive and publicly available data such as an ABX index, a composite index of CDS spreads for key financial institutions, the TED spread, leverage ratios of the largest commercial and investment banks, aggregate and sectoral credit expansion, debt trends across major sectors (government, corporate, household), and pertinent information (as available) of shadow banking assets as a share of total assets, the maturity structure of shadow banking liabilities, and financial institution exposure to credit derivatives.

- The very nature of financial innovation is to stay ahead of the regulators and hence their data collection efforts as well. For this reason, the Fund would always be "behind the curve" if the Fund could only see the risks when the data are finally available.
- "You don't see what you're not looking for." Neither the U.S. nor the U.K. authorities, despite presumably having much greater access to data, saw the crisis coming. Indeed, the Fund's views on financial sector soundness were very much in line with those of country authorities. Furthermore, once the crisis was evident, the Spring 2008 *Global Financial Stability Report* was able to provide a remarkable estimate of expected financial sector losses, without any additional access to data. But now they knew what to look for!
- To quote from the *Economist* (January 15, 2010), "In the run-up to the crisis, policymakers and supervisors, like most other people, managed to rationalize bad things that were plain for all to see, such as inflated house prices and some banks' rock-bottom capital levels." As Claudio Borio of the Bank for International Settlements put it, "The main reason why crises occur is not lack of statistics but the failure to interpret them correctly and to take remedial action" (Borio, 2012).

In sum, gathering more and more data is not a substitute for the effective use of available data or for willingness to challenge mainstream thinking.

Source: This box is based on the findings in IEO (2011a).

initiatives would help their work at the Fund, half of the respondents believed that the Fund currently lacks the capacity to effectively use all the data that ideally would be gathered under these initiatives. Prioritization is thus key to ensure that the Fund has the data needed to strengthen its surveillance of an increasingly complex global economy, yet avoids placing an excessive burden on member countries and on its own ability to absorb the information.

A growing body of work helps to point the way forward.

49. Some of the recent literature has been critical of focusing primarily on ever more financial and/or

market data to sound early warning of crises.⁴⁷ Several authors argue that some macroeconomic indicators are better at crisis prediction than are financial sector and market indicators, concluding that using available data in a different way may be at least as fruitful than the never-ending quest for more data (Borio and Drehmann, 2009; Eichner, Kohn, and Palumbo, 2010; Borio,

⁴⁷For example, a number of FSIs often continue to suggest soundness even as conditions are deteriorating. Even more timely data may perform poorly as early warning indicators. For example, market indicators might fail to indicate problems on the horizon—risk and volatility indicators were at historic lows just prior to the recent global crisis. This does not imply that collecting these data serves little purpose. Some of these data may not serve well as early warning indicators, but could prove extremely useful in responding to crises.

2012; Drehmann and Juselius, 2013; and Alessi and Detken, 2014). Indeed, Haldane (2012), in a speech at the 2012 Jackson Hole conference, stressed that the more complex the system, the greater the need to keep it simple, echoing findings of the BIS, U.S. Federal Reserve, and others that sometimes “less is more.”

50. On the Fund’s part, some recent work on balance sheet analysis (BSA) provides a good example of how collaboration between Fund statisticians and economists can shed light on the way forward for more effectively identifying and using data to support the Fund’s strategic work (IMF, 2015b). Both the global financial and euro crises might have been better foreseen through rigorously applying BSA.⁴⁸ A full set of balance sheet matrices is also a primary starting point for understanding macrofinancial linkages, and complemented with a global flow of funds,⁴⁹ forms the basis for the analysis of interconnectedness and spillovers. The use of BSA to strengthen surveillance was a running theme throughout the 2014 TSR and IEO (2011a).

51. But the recent global crisis was not the first to shed light on the usefulness of this approach.⁵⁰ The Asian crisis was the catalyst for work on the BSA at the IMF (Allen and others, 2002), and the 2004 “Review of Data Provision for Fund Surveillance Purposes” (IMF, 2004a) was already pushing for its use in Article IV consultation staff reports. Yet BSA was used only sporadically pre-crisis and typically for emerging markets. It was rarely employed for low-income countries (largely due to lack of data) or for advanced economies (where at least partial, and in some cases, like the United States, fairly complete data were available).

52. Why was BSA used so sparingly pre-crisis? Lack of analytical tools (and staff training on those available) hindered its use in bilateral surveillance. But the primary reason was that very few countries, even today, provide the full set of sectoral balance sheets. The IMF, particularly in the context of the DGI, has become more

proactive in encouraging the compilation of balance sheet data by its member countries, and now there is reason to expect that data availability will not be as significant a hindrance as it had been in the past (Box 6).

53. Much more remains to be done, however, especially on data for the corporate, household, and shadow banking sectors.^{51,52} Against a background of fiscal austerity in many countries, the demand for more complete balance sheet data might run up against other, perhaps more urgent, needs. Nevertheless, a compelling case could now clearly be made that the benefits, not only to the IMF but to the member countries themselves, outweigh the costs.

C. Data Quality

The Fund is not just a passive recipient of data; it runs some validation checks and promotes data quality.

54. The Fund has a number of mechanisms to obtain some assurances about the quality of the data it uses. With STA playing the pivotal role, it has developed methodologies for the proper compilation of economic and financial statistics, and works to support high-quality data through capacity-building—technical assistance and training. The Fund also performs some validation checks in the course of its operational work and prior to dissemination, with these checks varying by department and purpose of the data.

55. STA relies mostly on official data reported directly by countries. It checks these data for their compliance with established formats, examines them for outliers, and performs some routine consistency checks to capture large discrepancies across data sets. STA

⁴⁸A study on the United States using balance sheet analysis concluded: “Detailed analysis of aggregate sectoral balance sheets could have been helpful in identifying pressure points for the U.S. economy pre-crisis Balance sheet data for [households] and [other financial centers] were indicating a build-up of vulnerabilities, while standard vulnerability (financial soundness) indicators for the U.S. were not recording ‘red flags’ pre-crisis.” (IMF, 2015c).

⁴⁹In addition to its work on balance sheets, STA is also pushing forward with cutting edge work on a framework for the global flow of funds.

⁵⁰A key difficulty is that statistics are often produced with considerable delay. Ideally, forward-looking indicators would be the preferred means of detecting emerging risks, but these are difficult to come by. In their absence, macroeconomic stocks data (e.g., balance sheet data) could better indicate a buildup of pressures due to their “sticky” nature (the slow rate of change of stocks).

⁵¹In many countries, the shadow banking sector is the fastest growing segment of the financial sector, and in some cases, is larger than the banking sector.

⁵²Latin American Shadow Financial Regulatory Committee (2015) and Reinhart (2015) raise concerns, in the context of the expansion of shadow banking, about data on the extent of leverage in emerging markets and whether international reserve positions may overstate available resources. For example, reserve availability may be overstated when (i) central banks intervene by issuing dollar-linked debt, (ii) third parties (e.g., sovereign wealth funds, special status banks, state-owned enterprises) intervene in forex markets on behalf of the central bank, (iii) swap arrangements are not adequately captured in reserves data, and (iv) lines of credit extended by Chinese development banks to emerging markets are not included in external debt data. In general, recent Article IV reports for the affected emerging market economies have not covered these potential data shortcomings or have done so very tangentially. On occasion, issues such as the treatment of certain types of interventions have been raised, but have not been viewed as key areas for concern.

Box 6. Does Lack of Data Still Prevent the Use of Balance Sheet Analysis?

In October 2015, to encourage Fund economists to utilize balance sheet analysis (BSA) more frequently, the IMF posted an Intranet article entitled, “Five Things You Need to Know About Balance Sheet Analysis.” The following excerpts from three of those “five things to know” indicate that the authors of the article believe that lack of data is no longer the inhibiting factor that it once was:¹

“1. Balance sheets matter a lot. Balance sheet analysis captures the role played by financial frictions and mismatches in creating fragility, amplifying shocks, and generating spillovers. The boom, bust, and recovery associated with the global financial crisis can all be viewed through the prism of balance sheets. The boom was associated with increased private sector debt, and the bust created a decline in wealth that was propagated across the world through balance sheet linkages, even as debt remained elevated. Recovery has featured deleveraging, as the private sector restructures its balance sheets by increasing savings, cutting spending, and repaying debt. In turn, governments have responded by expanding the fiscal or central bank balance sheets to buttress demand.

2. Data no longer pose major constraints to analyzing balance sheets. While balance sheet data were hard to

¹This conclusion may be too sanguine, in light of the fact that fewer than two dozen Fund members were able to provide complete annual and quarterly sectoral balance sheets as of mid-2015.

come by in the past, country coverage and granularity of data have improved. This is in part due to Fund-supported initiatives such as the collection of data on financial sector balance sheets through standardized report forms (SRFs), and information on cross-border financial interlinkages through international investment positions (IIP) and coordinated surveys on direct and portfolio investments (CDIS and CPIS). Such initiatives have helped to better capture the state of the balance sheets of key sectors of the economy and how they are linked to each other as well as to the rest of the world. Even where balance sheet data are not fully available for some sectors, it is possible to make pragmatic assumptions or use supplementary information—including national sources, micro data, and surveys—as a workaround.

3. A lot of the data can be exploited using off-the-shelf techniques. Successive waves of crises have driven innovation in the Fund’s analytical toolkit, including macro models, techniques for macrofinancial stress testing, the balance sheet approach (BSA), and debt sustainability analyses. The new Board paper and note also develop some new empirical tools, including: illustrations of how to construct balance sheet matrices using Fund-collected data, and analyze them to get an aggregate sense of the key vulnerabilities in the economy; tools to dig deeper into these identified vulnerabilities using micro data; and general equilibrium and reduced form approaches to improve macro forecasting by incorporating balance sheet variables.”

is also working on implementing some intersectoral consistency checks,⁵³ which could prove an important additional tool for quality control.

56. Though many of the Fund’s area department country teams obtain much of their data directly online from national sources or from commercial databases (such as Haver Analytics), many staff missions, particularly in low-income countries, still spend considerable time collecting data in the field, with Fund staff often “getting their hands dirty” working on data with their counterparts during missions.⁵⁴ The policy discussions between mission teams and country authorities often reveal data inadequacies, potentially prompting

⁵³Official data are typically sourced from several agencies within the same country (e.g., national statistics office, central bank, ministry of finance) and are thus often inconsistent on an intersectoral basis, as these agencies often do not cross-check their respective data.

⁵⁴While this is often among the most appreciated contributions of IMF staff during missions, staff often consider it among the least rewarding parts of mission work.

corrective action and improvements in data quality. Staff estimates, which are often discussed first with country authorities, are frequently used to fill in the blanks from missing or problematic data.⁵⁵

57. Implicit in IMF country teams’ collection of data are validation activities that involve some verification of data at the primary source, checking the accuracy of basic calculations, and assessing overall consistency within a macroeconomic framework. According to staff interviews, country teams have traditionally been more active in checking and validating data in emerging markets and low-income countries, while tending to accept without question the data from advanced economies. The IMF’s flagship publications use a combination of the data collected by area departments and those from commercial databases as inputs. At this stage,

⁵⁵Jerven (2016) uses the example of Ethiopia to illustrate the lack of clear procedures as to the use of staff estimates in place of official data that are questioned by staff.

additional validation checks are performed, attending, to some degree, to issues of global cross-country consistency.

58. Nonetheless, all these validation activities still fall short of fully addressing deficiencies in source data and disparities in definitions and coverage. Errors can be, and have been, missed by the standard validation checks. This can have real consequences for member countries, as exemplified by an incident with the April 2009 *GFSR*. The IMF presented the external debt refinancing needs, as a ratio to foreign exchange reserves, of the central and eastern European countries, but the figures presented were, in some cases, more than twice the real ratios. The IMF corrected the errors, but not before they had (according to authorities) adversely affected market confidence. This incident prompted MCM to tighten its validation procedures, including by assigning a dedicated “fact checker,” sending the data to area departments for review,⁵⁶ and improving country desk inputs to the *WEO* database.

Ultimately, though, the quality of data depends on the member country . . .

59. Inevitably, there are limits to what the IMF can do to correct the shortcomings of data provided by member countries. Member countries show a wide disparity in the degree of development and independence of their national statistical compiling agencies, in their institutional structure, and in the resources allocated to statistical activities. While there is still room for the IMF to strengthen its work on quality assurances, ultimately, the quality of data depends on the attention and resources that member countries devote to it. Indeed, the accuracy of the information disseminated by the Fund is the sole responsibility of countries.⁵⁷

60. Concerns about quality are particularly relevant for the more resource-constrained statistical compiling agencies in low-income countries, which face greater difficulties in collecting primary source data, tend to employ outdated methods, and struggle to retain qualified personnel.⁵⁸ In response to a survey question, IMF

⁵⁶According to some interviewees, this step is very time-consuming (and at times, impossible) for area department country desks, as the *GFSR* heavily uses data from commercial sources (including for some macroeconomic data) which might diverge from those used by the country desks.

⁵⁷Article VII, Section 5 notes that it is the member’s obligation to provide accurate data to the Fund, to the extent of its ability.

⁵⁸The binding nature of resource constraints was clearly evident in recent years when, in many low-income countries, the emphasis on the Millennium Development Goals forced authorities to give

staff, particularly those working on emerging market and low-income economies, overwhelmingly pointed at the limited capacity of countries as a main reason for data deficiencies that hamper surveillance.⁵⁹ But limited capacity was also mentioned for almost a quarter of advanced countries with data deficiencies.

. . . but STA has played a major—and much appreciated—role in supporting high-quality data from the membership.

61. Although the assurances of quality that the IMF can provide in the short run are limited, STA’s initiatives to strengthen data quality over the medium and longer term are significant. The methodological manuals developed by STA have become the world standard that countries seek to adopt and implement,⁶⁰ while the technical assistance (TA) and training provided by STA are effective forces for the improvement of data.⁶¹

62. Training in statistics is highly appreciated by recipients, with 90 percent or more of survey respondents agreeing that the training is of high quality, aligned with the recipients’ priorities, feasible to implement, and has helped improve the quality of data. Appreciation for TA is even stronger than that for training (Figure 4), with views on its quality, relevance and feasibility almost unanimously positive. Some reservations were expressed, though, on follow-up and support subsequent to TA—partly in response to STA’s revised approach to committing follow-up assistance, which involves setting specific benchmark actions for implementation, together with evidence of compliance.⁶²

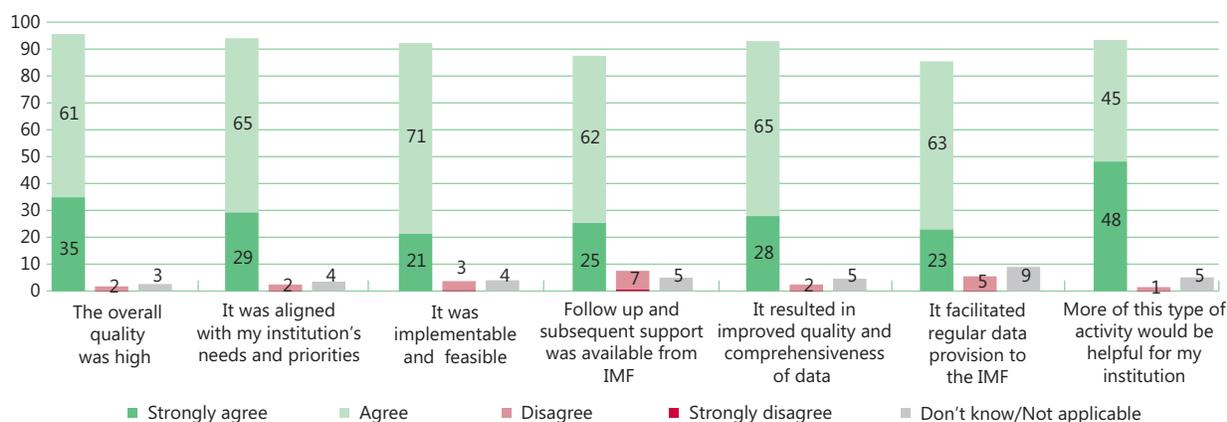
precedence to social indicators to the detriment of data on economic growth or employment (Jerven, 2013).

⁵⁹Jerven (2016) notes, as examples, huge changes in some low-income countries’ GDP statistics due to rebasing after years of using out-of-date baselines, calling into question the validity of surveillance based on numbers that could change so markedly. Nigeria’s GDP, for example, increased by 89 percent in 2014 after the base year was changed from 1990 to 2010, instantly vaulting Nigeria to the top of the GDP chart in Africa.

⁶⁰Data providers in member countries, both in interviews and surveys, expressed highly favorable views on the associated manuals and guides, with respondents agreeing that they are both practical and helpful (almost unanimous), as well as easy to understand and feasible to implement (85 percent).

⁶¹IMF staff, nonetheless, noted that the effectiveness of TA is sometimes undermined by the fundamental tension between weak governance and transparency, as opacity and lack of data preclude accountability.

⁶²This change in approach includes a move to a Results-Based Monitoring Framework and is due, in part, to the demand from the donor community to ensure effective allocation of resources. See also IEO (2014c).

Figure 4. Perception of Statistics Technical Assistance*(In percent)*

Source: IEO Survey of Country Authorities and Data Providers.

63. A significant development in recent years has been an increase in the share of TA financed by donors. This doubled between FY2011 and FY2015 to 60 percent (35 percent excluding Regional Technical Assistance Centers). The increasing reliance on donor financing has led, at times, to a less than optimal allocation of resources, when donors' priorities have not been fully aligned with those of the Fund.⁶³

64. In general, the IMF explicitly avoids providing assessments of the quality of member countries' statistics.⁶⁴ However, the data modules of ROSCs come closest to a comprehensive assessment of data quality. The Data Quality Assessment Framework (DQAF), which lies at the core of the data ROSC, provides a structure for assessing the extent to which countries meet the prerequisites of data quality—such as independence of, and adequacy of financing for, the compiling agency—or follow international best practices in regard to established standards.⁶⁵ However, the DQAF is more focused

⁶³In this regard, STA has recently developed statistical scorecards for a large share of the Fund's membership. The scorecards provide country-specific snapshots of data methodology and provision in a heat map format, so as to provide country teams and reviewers a quick reference tool to help determine capacity development needs and underpin surveillance dialogue on data issues. These scorecards seem a promising approach to better prioritization of TA needs and could also promote more candid assessments of data adequacy for surveillance.

⁶⁴When the dissemination initiatives were first discussed at the Executive Board, "... Directors emphasized that the Fund should avoid making direct public assessments of data quality ... to avoid the implication that ... the Fund was certifying good practice with respect to quality and other characteristics of the data." (IMF, 1996b).

⁶⁵That is, integrity, methodological soundness, accuracy and reliability, serviceability, and accessibility.

on statistical processes than on passing judgment on the quality of the statistical output itself.⁶⁶

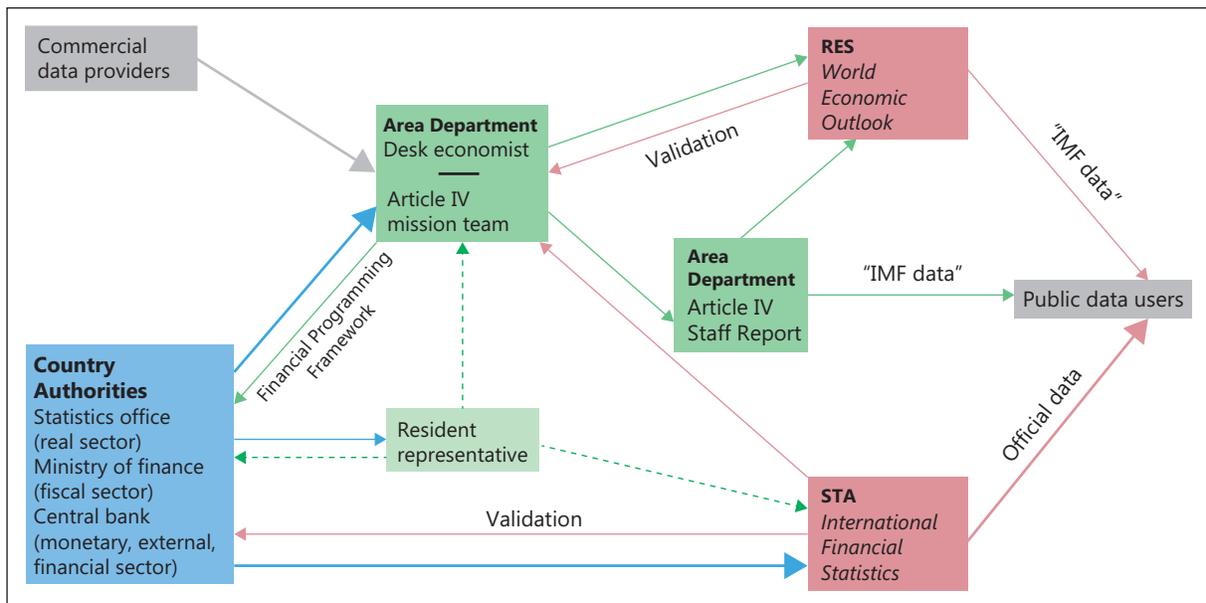
65. This evaluation's interviews and survey of data producers suggested that almost all of those who had been involved in their countries' data ROSCs considered them very useful for improving data quality and implementing best practices. Some authorities indicated in interviews that data ROSCs had the additional effect of strengthening the hand of national statistical offices in their quest for more resources. More than three-quarters of respondents believe that conducting these exercises on a periodic basis would be helpful. However, in recent years, due to their high cost, data ROSCs have become ever fewer and far apart and have now been (at least temporarily) suspended.

66. Recent problems with the reporting of fiscal and debt statistics in some countries, together with the phasing out of data ROSCs, have led the Fund to revamp its Fiscal Transparency Evaluation (FTE), including the addition of an important data pillar. This pillar replicates for the fiscal realm the categories of the data ROSC, including those of data quality and integrity. In contrast to the data ROSCs, the revamp of the FTE focuses on outputs rather than processes, thereby placing greater emphasis on the quality of published information.⁶⁷ It clearly presents strengths, weaknesses, and reform priorities through summary heat maps, making the FTE more accessible to policymakers, civil society,

⁶⁶Of course, it might be expected that a well-functioning statistical system is more likely to produce quality data.

⁶⁷STA has noted that it plans to revise the data ROSC to increase its efficiency and effectiveness, including by covering statistical outputs.

Figure 5. Data Flows at the IMF



Source: IEO.

and other stakeholders (in contrast to the relatively impenetrable data ROSCs, largely accessible only to statisticians).

67. The IMF is also cautious about explicitly assessing the prerequisites of quality, perhaps most importantly, that of a well-funded and autonomous national statistical office. That is, the IMF does not typically emphasize the need for member countries to better equip their national statistical offices, notwithstanding the evident benefits this would bring to the countries' own policymaking. In particular, weak statistical offices can fall prey to political pressures and inadequate funding, undermining the reliability, accuracy, and unbiasedness of their output. The Fund seldom places a priority on establishing an active dialogue on data issues with country authorities at the policymaking level, such as their needs for capacity-building or, when relevant, on including such issues in Fund conditionality.

D. Internal Data Management

Data management problems are deep-rooted . . .

68. The usefulness of data for IMF operational purposes also depends on the Fund's internal data management practices and, in particular, on the staff's ability to access a wide range of good quality, consistent data

on a timely basis. The evolution of the Fund's data requirements and activities has led to a highly decentralized approach to data collection, management, and dissemination (Figure 5 is a stylized representation of data flows to, within, and from the IMF). As a result, there are now about 180 *cross-country* databases in the IMF, of which about half are internally produced and the remainder externally provided, and more than 180 country-specific area department databases. At the same time, databases have become commensurately larger and more complex, implying a greater premium on efficient management and documentation. Departments have long been expected to adopt guidelines for data management.⁶⁸

69. Decentralization and the associated proliferation of databases have created a number of deep-rooted problems, all of them closely intertwined.⁶⁹ First, the Fund's fragmented and uncoordinated approaches to

⁶⁸For example, a November 1995 memo from the then-First Deputy Managing Director stated, "All departments that maintain economic databases will be expected to establish and implement data management guidelines in accord with the Fund-wide guidelines."

⁶⁹In addition to the proliferation of databases, there has also been a proliferation of interfaces for accessing data—Economic Data Sharing System (EDSS), Economic Data Warehouse, Joint Library (which manages commercial databases), Data Management for Excel (DMX) Data Navigator, Economic Outlook Suite (EcOS), etc., adding to the complexity and confusion for the user in finding data.

data collection, validation, and management have contributed to data inconsistencies. Second, internal data sharing has been burdensome and inefficient, a problem aggravated by lack of incentives for proper data management and transfer of knowledge. Third, many of the Fund’s databases have been poorly structured and documented, without sufficient metadata for proper use outside the specific unit managing each database.

70. Decentralized data collection and management has also indirectly resulted in isolating STA from the rest of the Fund, increasingly leading STA to focus its efforts on data dissemination outside the Fund and on the external provision of statistical services—with its outputs largely disconnected from the Fund’s core operational work. One reason for this disconnect is that economists and statisticians have different approaches to data, with the former emphasizing pragmatism, usability, and timeliness, while the latter focus more on accuracy and methodological purity. Timeliness versus accuracy remains an unresolved question.⁷⁰

71. These problems have been amply documented in the past. [Annex 7](#) lists 17 of the many studies on the Fund’s data management problems over the past 50 years, most of which highlighted these same recurrent themes.⁷¹ In the wake of an Office of Internal Audit report on data management (IMF, 2007), the Fund launched in April 2010 the Economic Data Management Initiative (EDMI), the third in a series of attempts within the last decade to strengthen data management. The EDMI concluded that: (i) the Fund was at the earliest stages of data management maturity,⁷² with technology driving the approach rather than analytical needs; (ii) there were no clear guidance strategies; (iii) the Fund data arrangements were characterized by weak governance bodies; (iv) data procedures were poorly executed; and (v) there was no holistic view, with the approach to data management being excessively “bottom up.” More specifically, the EDMI highlighted the existence of two parallel paths for data compilation at the Fund, reflecting the differences in the mandate and priorities of area departments and STA. The report recommended extensive changes, including in the data

governance structure and a move to a Fund-wide structured database.

... and although progress has been made in several areas ...

72. In response to the starkly-stated EDMI findings, IMF Management created a new economic data management governance structure. This began operations in May 2012, with three key bodies: the Economic Data Steering Committee (EDSC), the Economic Data Governance Group (EDGG), and the Economic Data Team (EDT)⁷³ ([Figure 6](#)).

73. Substantial progress has been achieved over the last few years. The creation of the new governance structure—while still on a temporary basis—is a milestone and could contribute to overcoming organizational resistance and breaking down silos and associated data fiefdoms. Area departments’ data have been moved from Excel spreadsheets to structured databases, with associated gains in organizational clarity, use of metadata, more consistent processes, data sharing, and ease of transfer of knowledge. The Economic Data Registry—a single access point for all IMF internal databases and some external ones—is being developed, and the Common Surveillance Databases (CSD) are already in use ([Box 7](#)). These achievements address some of the problems described in the paragraphs above and provide a stepping stone for future and more ambitious actions.

... some of the adopted measures have been subject to criticism ...

74. Staff’s assessment of the new data governance structure is mixed. Among those surveyed staff familiar with it, a majority believe it has been helpful in improving data management practices, but many think it “just adds another layer of bureaucracy.” In interviews, members of the EDSC and EDGG expressed concerns regarding the new governance structure, including an excessive focus on technical and budget issues, rather than strategy, and the dominance of the EDT, together with its organizational location and that of the CSD.

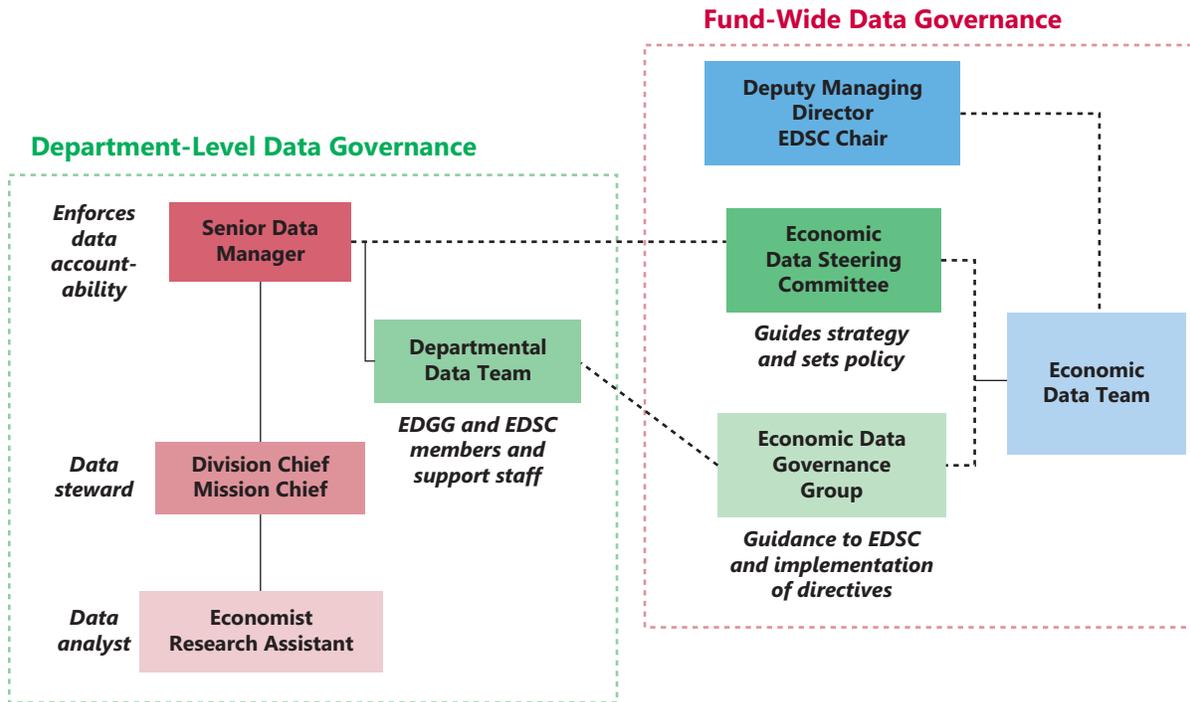
⁷⁰The trade-off between timeliness and quality was well expressed at the IMF’s Second Statistical Forum, with speakers’ views ranging from “speedy rubbish is of no value” to “put the data users first.”

⁷¹See also De Las Casas and Pedraglio (2016).

⁷²Gartner Consulting, hired as part of the EDMI, determines data management maturity levels by grading six dimensions (vision and strategy, metrics, governance, organization, processes, and technology infrastructure) and comparing practices with industry standards. The Fund scored particularly low on vision and strategy.

⁷³The EDSC is supposed to be comprised of “Senior Data Managers” at the Deputy Director level from 15 departments, while the EDGG consists of mid-level managers, with the chair of the EDGG heading up the EDT.

Figure 6. New Data Governance Structures at the IMF



Source: IMF Economic Data Team.

Box 7. The Common Surveillance Databases and the Quest for Better Data Sharing at the Fund¹

Arguably the number one problem in the area of data management at the IMF is the lack of proper systems and procedures for efficient and consistent data sharing within the Fund. Data sharing has traditionally been done manually and on an ad hoc basis, with virtually nonexistent consistency controls. These issues are increasingly worrisome as cross-country analysis and multilateral surveillance gain relative weight among the Fund's operations.

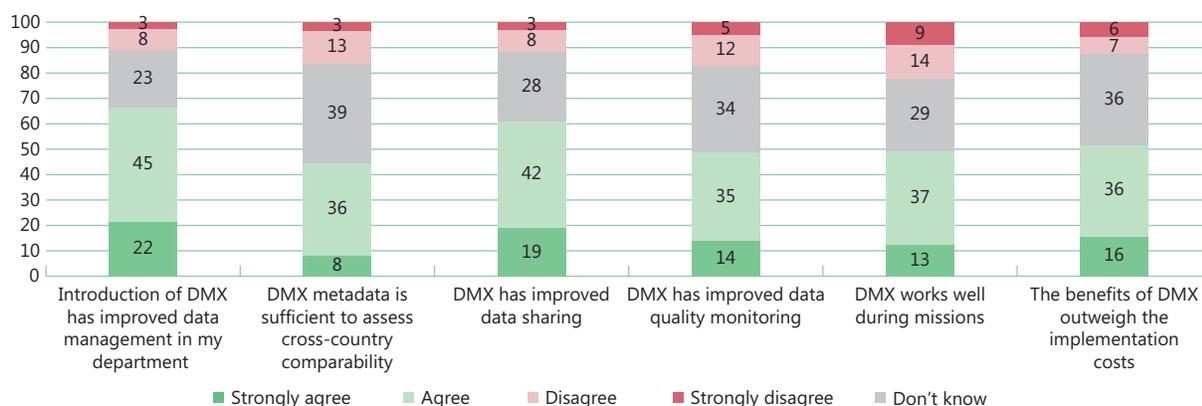
In the post-EDMI context, with the migration to structured databases completed, the EDT has turned to the creation of the Common Surveillance Databases (CSD)—with the explicit aims of facilitating data sharing, integrating data used for bilateral and multilateral surveillance exercises into one common database, reducing reputational risk associated with data inconsistencies by improving metadata documentation, and avoiding excessive proliferation of databases. The CSD is composed of two Fund-wide accessible databases: (i) the Forecast CSD, which will contain all variables included in desks' macro-frameworks and

all data required for desk-based cross-country exercises, and (ii) the Historical CSD, made up of all historical desk data and all data required for desk-based cross-country historical databases.

The success of the CSD is based on the systematic collection and storage of data and metadata, together with the implementation of new processes for data flows, revisions to historical series, and validation checks. Therefore, responsibility is shared among country teams and their departments' data manager, cross-country database managers, and the Fund's Data Management Governance Structure. The operational work of the CSD has been assigned to RES, building on existing processes and expertise associated with the *WEO*. The October 2015 *WEO* is the first for which all country teams submitted their data via the CSD.

¹Based on EDT 2014, "Proposal for the Establishment of Common Surveillance Databases."

Figure 7. Staff Perceptions of DMX
(In percent)



Source: IEO Survey.

75. While it might seem that STA would be the logical home for the EDT,⁷⁴ interviewees felt that long-standing concerns regarding STA's ability to manage data for operational purposes would have undercut support for the EDT, had it been initially located there. Nevertheless, many interviewees admitted that STA should be the natural long-term location, provided that STA undertakes the necessary reforms—particularly in the area of timeliness.

76. Concerns were also raised by interviewees regarding the location of the CSD in the Fund's Research Department,⁷⁵ arguing that STA might be better suited for this task, given its core expertise.⁷⁶ Another line of argument is that the CSD should be managed by dedicated staff outside any department, for example, the EDT. While the CSD's potential to improve data flows in the Fund is generally recognized, some interviewees questioned the lack of strategic thinking behind the development of the CSD and the consideration of data sharing as a goal in itself, rather than as a means to strengthen the Fund's economic analysis. It was

⁷⁴The EDM's recommendation was that the EDT be located in the Office of the Managing Director (OMD), but at first it was placed in an area department. More recently, it has been relocated to the OMD.

⁷⁵Minutes of the relevant EDSC meeting indicated that all but one of the EDSC members preferred RES as the CSD location. However, in interviews of EDSC members, a number of them thought that STA could be an appropriate location.

⁷⁶The CSD, together with the Economic Data Registry, have a clear precedent in the Economic Data Warehouse (EDW), a STA-led initiative to create a single point of access to all data available at the Fund. However, under its current configuration, the CSD would not contain STA's databases. While the development of the EDW is now suspended, the experience illustrates the complexity of data management issues at the Fund (see IMF (2007), which supported the EDW and its management by STA).

highlighted that the CSD perpetuates some of the data management problems by adding an additional platform, when a unified system for all data management at the Fund is what the institution needs.

77. More broadly, the problems with the new governance structure stem from a lack of engagement by Management and insufficient interest on the part of EDSC/EDGG members.⁷⁷ Management involvement in statistical matters has also been hindered by the split in responsibilities between two Deputy Managing Directors—one of them chairing the EDSC and another one in charge of STA.

78. Staff regard the move to structured databases fairly positively (although one-third declined to provide an opinion), believing it has improved data management and sharing. But the assessment of other dimensions is more nuanced (Figure 7),⁷⁸ particularly on the technical front, as they see deficiencies in DMX (Data Management for Excel) as its chosen platform. A significant number of respondents raised concerns regarding its "black box" nature, its operational complexity, the quality of metadata, and the coding system. Moreover, DMX, as an internally developed tool, might prove less adaptable, state-of-the-art, and cost-effective than commercial solutions in the long run. Furthermore, some departments use alternative, externally developed

⁷⁷Indeed, many of the EDSC and EDGG members stressed that they did not volunteer for this position and had no deep interest in data issues. In fact, many of the members were reluctant to be interviewed, noting that they knew very little about such issues.

⁷⁸An important caveat regarding the survey results is that the Fund's data management system has been evolving rapidly since the survey was conducted in February–March 2015 (e.g., the CSD became operational after the survey was completed).

platforms (e.g., EcOS in FAD, RES, and STA), complicating the information technology environment for effective institution-wide data management.

... and some fundamental issues remain unaddressed ...

79. A clear data strategy is the crucial missing element in the efforts underway. The EDT has provided a set of intermediate targets, some tools, and a roadmap but, according to interviews with EDSC/EDGG members, a holistic strategy—that clearly sets medium-to-long-term goals, defines the business case, and establishes the value attached to data as an institutional asset—is still lacking. Arguably, this may be a consequence of the continuation of an excessively bottom-up approach to data management. This consensus-based, process-oriented style slows progress and hampers the adoption of broader, more innovative solutions with the potential to yield more sustainable outcomes over time.

... not least the role of STA ...

80. The issue of STA's disconnect from other departments, with its outputs not integrated with the Fund's core operations, has been largely dropped from the broader agenda. STA's data are perceived by IMF staff as primarily useful for research and historical analysis, but not for policy-oriented and operational work, mainly for lack of timeliness and coverage.⁷⁹ Adding to the lack of integration of STA's outputs, some departmental data management guidelines explicitly favor the use of other sources over STA and raise questions regarding the usability of STA's data.^{80,81}

⁷⁹On coverage, Jerven (2016) notes that the February 2015 *IFS* was missing 2011 data on real GDP growth for almost 40 percent of countries. By comparison, the October 2014 *WEO* database was missing the same data for only 8 percent of countries.

⁸⁰From the data management guidelines of an area department: "Country teams should maximize electronic data collection from national statistical bureaus and central banks, as well as from commercial sources. . . . Use of STA economic and monetary data, where relevant and feasible, including the Integrated Monetary Databases (IMDs), is encouraged in cases where country data are not available from commercial sources. . . . However, delays in STA data processing, and the limited scope of data available may make this not possible."

⁸¹Staff working on advanced and emerging market countries strongly prefer Haver Analytics over STA (the number of IMF staff using Haver exceeds 1,000), on the grounds that data are easier to find and better access tools are provided, and despite the fact that Haver Analytics feeds intensively on official data sources (largely the same sources used for STA's macroeconomic data) and draws directly on some STA data series.

81. Moreover, it could be argued that the new CSD, together with Management's decision to assign responsibility for oversight of data management and of STA to different Deputy Managing Directors, institutionalizes the existence of two parallel data collection and management systems in the Fund and isolates STA further.⁸² During interviews, EDSC and EDGG members expressed doubts about the current and future role of STA regarding internal data management in general and managing the CSDs in particular.

82. Yet the survey results show staff's clear desire for a centralized provision of statistical services (seemingly an obvious role for STA), in line with the practices in most peer institutions (Box 8). For example, three-quarters of staff think a centralized data unit should be in charge of managing a common database for IMF staff to access all data. And about two-thirds of respondents think this unit should monitor the consistency of internal databases and collect and provide the bulk of the data for surveillance operations in a timely manner. At the same time, however, the survey reveals staff's dissatisfaction—and lack of knowledge—regarding the current performance of STA in providing internal services (Figure 8).

83. These problems are well-known among STA management and staff, who indicated during interviews a strong commitment to undertake the necessary reforms to turn the department into a service provider to the rest of the Fund. In fact, STA has already put in place some initiatives that increase collaboration with other departments in support of surveillance activities.⁸³ Successful examples include the recently published joint work on balance sheets, the integrated monetary database, and the work of STA on the DGI.

84. Nevertheless, developing an ability to provide "on time" data would require a major cultural shift—at least on the part of STA staff—from a focus on methodology to that of timeliness. Peer statistical units in the IAG disseminate data (at least internally) as they are produced by member countries—recognizing that analysts need to have immediate access to data that move the markets or affect the countries' policymaking—and subsequently clean and adjust them to methodological standards.

⁸²While STA is formally represented in the EDSC and EDGG, it is treated like all other represented departments, with no special status, inputs, or additional responsibilities within the governance structure.

⁸³In March 2015, STA established a new division to serve as a focal point for coordinating STA's activities with area and functional departments.

Box 8. Data Management Practices in Comparable Institutions

While models differ across institutions, and data management frameworks must adapt to the needs of each organization, the 2005 Towe Report and the interviews conducted for this evaluation identified a number of successful practices in organizations comparable to the Fund.¹ These practices imply higher levels of centralization and coordination than those currently in place in the IMF:

- A single unit is responsible for the institution's database that provides inputs to all or most publications.
- This unit is responsible for collecting, validating, and documenting the data, and providing tools to access data for official publications.

¹The Towe Report studied the Asian Development Bank, the Federal Reserve, the OECD, and the World Bank. The IEO has extended the analysis to the BIS, ECB, and Eurostat.

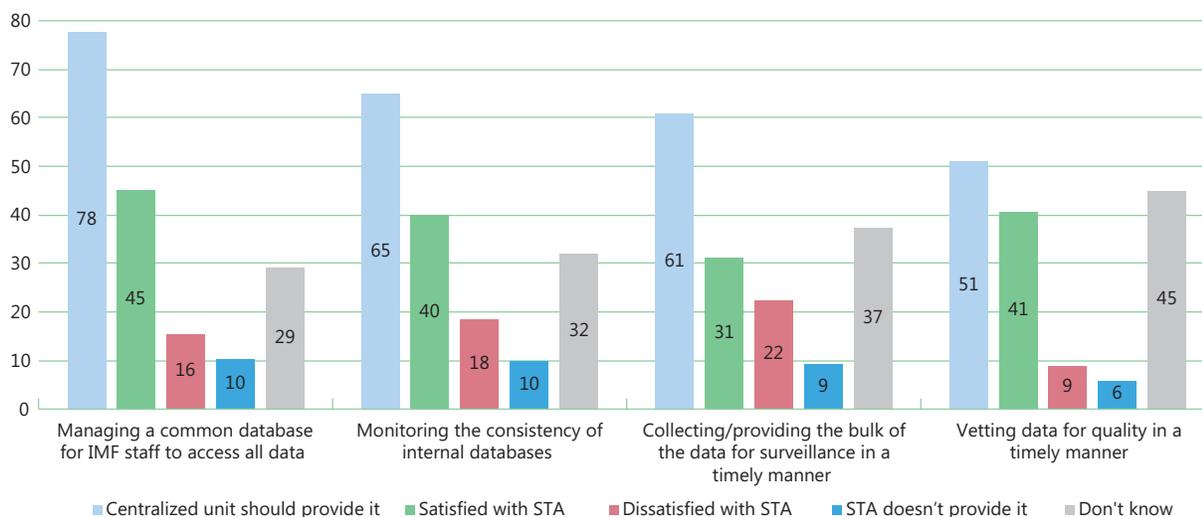
- This unit also ensures that (preliminary) data are available to analysts with minimal delay.
- A common nomenclature is used across all series stored in official databases, and this nomenclature is maintained by the centralized data unit.
- Desk economists use the institution's database because they are mandated to do so, and—more importantly—because they receive the array of tools and the support to access the data.

The Towe Report also highlighted how, in contrast with the other institutions, data management initiatives in the Fund depend largely on unrewarded work. This, of course, hampers their effectiveness, sustainability, enforcement, and standardization.

Sources: IMF (2005) and IEO interviews.

Figure 8. Staff Perceptions of Centralized Provision of Data Services

(In percent)



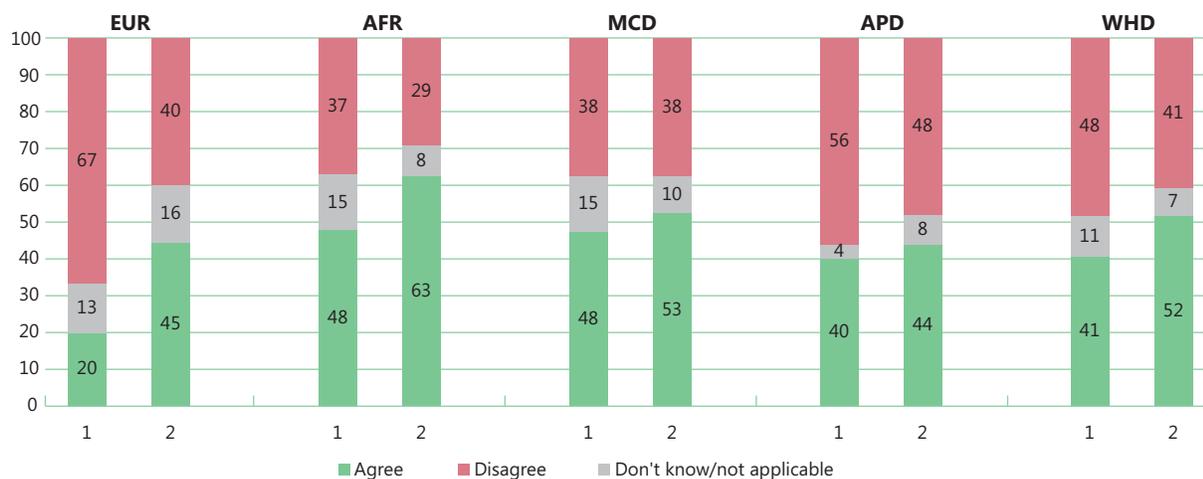
Source: IEO Survey.

... and getting the incentives right.

85. The problem of staff incentives for proper data management remains largely unsolved. During interviews, staff made clear that good data management in the Fund relies mostly on personal interests and attitudes towards data, and that the low visibility of such

work discourages staff from investing time in it. In fact, only one-third of survey respondents perceive data work as being part of their annual performance review discussions, although data management guidelines claim this should be mandatory (Figure 9). The de facto incentive structure is perceived as not rewarding good data management. Indeed, according to staff interviews,

Figure 9. Staff Perceptions of Data Management Practices Across Departments
(In percent)



1. My supervisor typically discusses my work on data management as part of the APR discussion.
2. My department provides incentives for good data management.

Source: IEO Survey.

being too closely associated with managing data was seen as potentially harmful to career prospects.⁸⁴

86. Data management guidelines do not provide adequate incentives for staff. Quality audits—reviewing the work of country desks—for data and metadata in the CSDs do not meet this need, at least in the view of some EDSC/EDGG members, who expressed a rather pessimistic opinion on this issue. Nor do departmental guidelines facilitate proper data management: in practice, their complexity and length (in some cases well over a hundred pages) discourage staff from reading them, let alone applying them on a daily basis. The same guidelines call for periodic assessments on compliance to be conducted annually or semiannually, but such reports are not being prepared in the form and with the frequency mandated—in some departments, none have yet been issued—and are not widely accessible within or across departments, eliminating their presumed positive effect on discipline through peer pressure.

87. Ten years ago, the Towe Report (IMF, 2005) identified eight major recommendations present in the Fund's many previous reports on data management:

⁸⁴In the words of an interviewed senior manager: "Research papers are valued here . . . if the analysis is done right, no one will mark you down for bad data management;" and those of a senior economist: ". . . excellent data management skills? Not on my annual performance review! That would imply I'm not a strategic thinker."

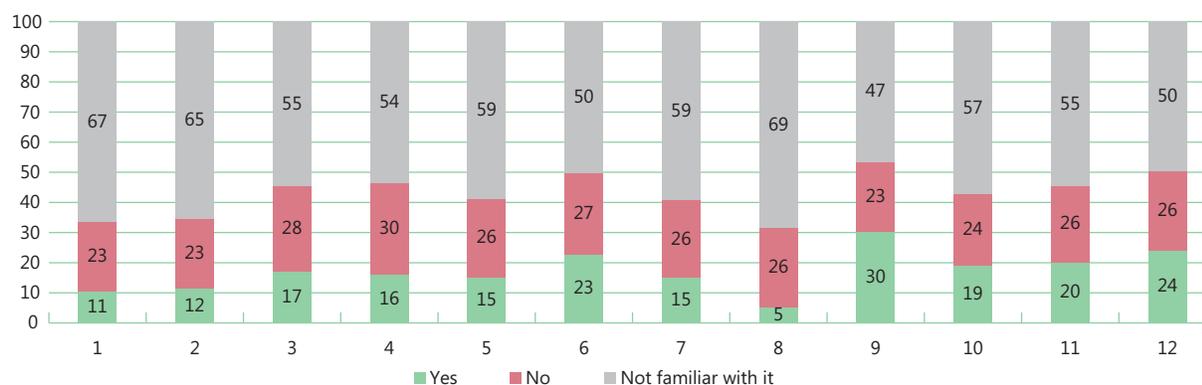
(i) improving the data of member countries; (ii) improving the tools available; (iii) staff training; (iv) establishing data management guidelines; (v) increasing incentives to follow the guidelines; (vi) shifting responsibility to research assistants; (vii) reconciling STA and country data; and (viii) centralizing the data collection process. Its diagnosis, over the previous 15 years, was that little progress had been made, except for the first two recommendations.

88. Today the diagnosis would be largely unchanged: while work on improving members' data continues apace and some improvements have been made regarding available tools, progress with the other recommendations has been limited, at risk of being unraveled, or nonexistent. During interviews, staff repeatedly expressed the view that to address the Fund's data management problems would require from Management a more forceful and mandatory approach than has been the case so far.

E. Data Dissemination and International Cooperation

The IMF disseminates large amounts of data . . .

89. The IMF is not just a collector of information for its own purposes. It also disseminates a vast array of

Figure 10. Use of IMF Specialized Databases*(In percent)*

1. Coordinated Direct Investment Survey (CDIS).
2. Coordinated Portfolio Investment Survey (CPIS).
3. Currency Composition of Official Foreign Exchange Reserves (COFER).
4. Data Template on International Reserves and Foreign Currency Liquidity.
5. Financial Access Survey (FAS).

Source: IEO Survey of Data Users.

6. Financial Soundness Indicators (FSIs).
7. Joint External Debt Hub.
8. Monitoring of Fund Arrangements Database (MONA).
9. Primary Commodity Prices.
10. Principal Global Indicators (PGI).
11. Public Sector Debt Statistics Online Centralized Database.
12. Quarterly External Debt Statistics (QEDS).

data and statistics through a variety of databases, documents, and publications. The IMF's data dissemination has grown exponentially,⁸⁵ propelled not only by the expansion of its membership, but also by technological developments and the relentless growth in the demand for information.

90. In general, the users polled for this evaluation have a positive perception of the data disseminated by the IMF and consider them better than, or at least as good as, those provided by comparable sources⁸⁶ in terms of quality, timeliness, and ease of access. There are only a few exceptions: users consider the ECB/Eurostat superior in terms of timeliness and Haver Analytics superior in timeliness and ease-of-access. IMF-provided data are heavily used by external stakeholders, with the *IFS* and the *WEO*, by a wide margin, the most commonly used resources. At the same time, aside from a few of the well-known databases, interviewees noted that it was difficult to find data on the Fund's website, a finding confirmed by the fact that the

IMF's specialized databases are largely unknown and rarely used (Figure 10).

... but the Fund is wrongly perceived as "endorsing" the data.

91. Users widely misperceive the Fund as ensuring the quality of the data it disseminates. Survey respondents consider, almost unanimously, that Fund-provided data are reliable and accurate, with an overwhelming majority believing these data are endorsed by the institution (Figure 11). But the Fund's data validation capacity is limited, and the perception of the Fund awarding a "seal of approval" could incur reputational risk for the institution.

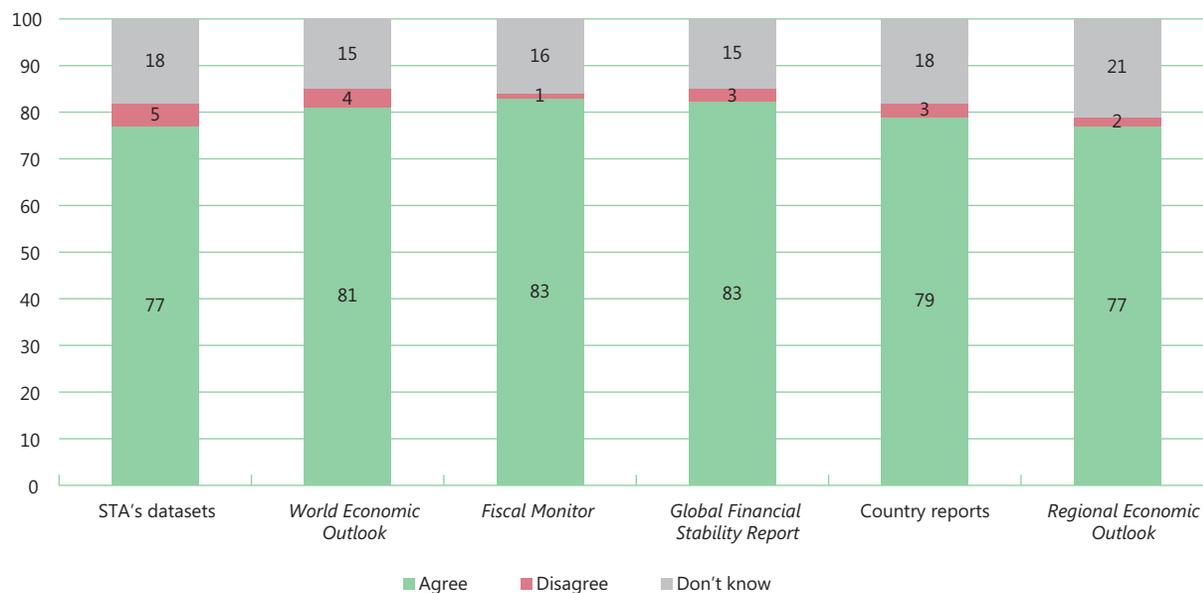
92. Concerns have also been voiced in the Fund for decades about the reputational risk stemming from data discrepancies and lack of comparability across IMF databases and publications.⁸⁷ These discrepancies reflect not only the differences among the inputs provided by countries and the different processes of

⁸⁵As of 2014, for example, the *IFS* disseminated up to 670 times series for each of 194 countries in the print version, but maintained more than 119,000 time series in its electronic database, up from 36 time series for 56 countries in its first print issue.

⁸⁶BIS, ECB, Eurostat, EIU, Haver Analytics, OECD, UN, and World Bank.

⁸⁷Initially, these concerns were expressed in terms of the *IFS* and *WEO*, as the *WEO* was the only IMF flagship document. Today, the challenge of data consistency extends across a much broader array of flagship documents, including the *WEO*, *GFSR*, *Fiscal Monitor*, *Spillover Reports*, *External Sector Reports*, and Article IV reports.

Figure 11. Survey Responses: “Data Quality is Monitored and Endorsed by the IMF”
(In percent)



Source: IEO Survey of Data Users.

Table 1. Discrepancies Among IMF Data Sources
(In percent of number of countries)¹

	Deviation Threshold (In percent)	Article IV vs. World Economic Outlook			International Financial Statistics vs. World Economic Outlook		
		Low-income countries	Advanced and emerging market economies	Total	Low-income countries	Advanced and emerging market economies	Total
Real GDP growth rate	<10	74.6	75.0	74.8	61.3	77.5	70.4
	10–30	12.7	18.8	15.1	19.4	15.0	16.9
	>30	12.7	6.3	10.1	19.4	7.5	12.7
Current account	<10	70.8	83.0	75.6	36.7	83.3	57.4
	10–30	16.7	12.8	15.1	16.7	4.2	11.1
	>30	12.5	4.3	9.2	46.7	12.5	31.5

Source: Jerven (2016).

¹Based on data for 74 low-income countries and 48 advanced and emerging market economies.

internal validation, but also the differing goals and frequencies of IMF outputs; for example, the *IFS* disseminates official data that seek to meet international definitions and standards, while country reports need to work with timely data understood by the authorities.

93. In line with the findings of previous internal IMF reports, the evaluation team found significant discrepancies in the data published by the Fund for the

same country and year in various datasets. [Table 1](#) summarizes a quantification of these discrepancies for disseminated figures of real GDP growth and the current account balance.⁸⁸ While discrepancies are typically

⁸⁸See Jerven (2016) for full results and a complete description of data sources and methodology.

wider for low-income countries, they also appear for advanced and emerging market economies.

The Fund’s recent move to providing data free of charge is an important step, but does it go far enough?

94. The Fund took a major step forward in January 2015, when it began to provide online access to its main databases free of charge.⁸⁹ This decision was praised by country authorities, academia, and other external stakeholders, and almost doubled the average number of users of Fund data during the first three months of operation. But a free data policy is not an open data policy, as the latter, despite its public good nature, could prove controversial at the Fund.

95. While often confused, *free data* are different from—and less ambitious than—*open data*. As indicated above, the Fund manages two broad types of country data: (i) *IFS*-style “official” data, which are intended to be internationally comparable and are basically a pass-through from country authorities; and (ii) operational data collected by country teams from the authorities or generally available sources. The former are the focus of the move to providing data free of charge. But it could be argued that, in the Internet era, when most countries’ official data can be found online, there is little value-added in just passing these on. The latter data, which can be more timely and “unique” to the Fund’s interaction with members, are shared only in as much as they are available in the Fund’s flagship and Article IV reports, but data as presented in the country reports are not “user-friendly.”⁹⁰

96. Other comparable organizations and academia have already adopted open data, which has become best practice. At the IMF, an open data policy—implying easy, universal access to most of the Fund’s operational data and the data underlying its research and other publications—would have positive ramifications. It would boost the Fund’s transparency and credibility, as external data users could more easily replicate and double-check the institution’s work. By the same token, it would contribute to the accountability of the Fund and member countries. A number of IMF staff interviewees believed it could also encourage IMF staff to pay greater attention to data if they knew that these data (and estimates to

“fill in the blanks”) would be subject to public scrutiny. Moreover, it could foster a move toward greater data comparability and quality in member countries if the staff’s operational data differed from the “official” data.

97. But an open data policy at the IMF would require a careful balancing of the institution’s roles as watchdog and trusted advisor. In its latter capacity, the Fund receives from member countries, as part of its operational data, confidential information that also is often market-sensitive. Such confidentiality must not be compromised, as mistrust could severely impact data provision by members, ultimately impairing the quality of the Fund’s work. Both country authorities and Fund staff raised concerns during interviews regarding these implications of open data.

The IMF also actively promotes data dissemination by member countries . . .

98. The IMF’s Data Standards Initiatives (e.g., the SDDS, GDDS, and most recently, the SDDS Plus) have played an important role in advancing data dissemination worldwide. However, after a surge of interest at the outset, these initiatives had languished for some years, with few countries graduating from the GDDS to the SDDS.⁹¹ Lately, though, these initiatives have gained some momentum (Box 9). The GDDS has been enhanced with the introduction of active monitoring of the countries’ dissemination practices—thus becoming the e-GDDS. The enhancement aims to foster dialogue during Article IV consultation missions on constraints and capacity-building needs, thereby providing incentives for countries to graduate to the SDDS and drawing policymakers’ attention to the need for statistical development (IMF, 2015a). On its part, the SDDS Plus should help address data gaps identified during the global financial crisis.

99. The dissemination initiatives, as their name indicates, focus on dissemination practices, not on verification of data quality.⁹² As indicated above, “quality” is difficult to define or assess, particularly as the IMF is not in a position to examine the production process of each specific statistic and gauge errors and events that may have influenced quality. The IMF thus chooses to leave the assessment of quality to users, prescribing the dissemination of

⁸⁹The Fund had lagged behind other international and regional organizations in its move to providing data free of charge.

⁹⁰A common wish of external data users was for the dissemination of country-report data in a downloadable format, for example, allowing the user to click on a table and immediately download the associated data.

⁹¹Some GDDS country authorities explained during interviews that, while they wanted to subscribe to the SDDS, their country was unable to graduate because of the Fund’s rigid approach to subscription and failure to understand national peculiarities.

⁹²There was debate during early Board discussion of the dissemination standards as to the appropriate focus. Indeed, one Executive Director noted that “. . . a set of standards that does not deal with the quality of statistics is empty. . . .”

Box 9. China: Subscribing to the SDDS

China's statistics have attracted unusual attention in recent years, particularly in the area of national accounts. Given China's status as the world's second largest economy, the controversy surrounding the quality of its statistics, particularly in regard to the actual size and rate of growth of the economy, is watched closely by academics, markets, and politicians the world over. The controversy is at its most heated in regard to quarterly real growth figures, where analysts often display a wide range of estimates at variance from the official preliminary figures.

Similarly, in line with the lower growth rates they estimate, some analysts believe China has overestimated the size of its economy. However, a recent study (Rosen and Bao, 2015) delved into the details of the Chinese statistical system, conducted robustness checks, and concluded that "China has made great progress in modernizing GDP statistics" and, if anything, the overall size of China's economy is underestimated. Indeed, they find that, if China were to switch from using the 1993 SNA to the 2008 version, its economic size could be as much as 13–16 percent larger—not a minor discrepancy for an economy of such global import.

The IMF has not stayed on the sidelines of these developments. Staff missions have discussed perceived data weaknesses with the authorities and included their assessment of the adequacy of the data in their reports. These assessments present a picture of slow but steady improvement over time. Whereas in the 1990s, staff raised major concerns across

virtually all sectors of the economy, viz., "... deficiencies in China's economic statistics are seriously complicating economic policy making and hampering effective surveillance . . .,"¹ by 2005, staff was balancing the discussion of weaknesses with recognition of the efforts at improvement made by the authorities. From 2008 onwards, economic statistics were deemed to be broadly adequate for surveillance (a rating of B in the Statistical Issues Appendix), despite some shortcomings, particularly in the areas of national accounts and government finance.

Part of the perceived improvement can be attributed to the technical assistance provided by the IMF and other international organizations. Over the past 25 years, the Fund sent close to 160 technical assistance missions on statistics to China. These missions covered all sectors of the economy, with an initial emphasis on the balance of payments, monetary statistics, and the national accounts. Subsequently, fiscal and financial sector statistics acquired more prominence. As an important step, on September 30, 2015, China for the first time reported the currency composition of its international reserves (COFER) to the IMF on a partial basis, with plans to gradually move to full coverage within two to three years. The progress made in the statistical area enabled China to subscribe to the SDDS on October 7, 2015.

¹ IMF (1996a).

information on methodologies and sources—monitorable proxies—to facilitate this assessment. These metadata are published in an IMF-supported bulletin board, the Dissemination Standards Bulletin Board (DSBB), "as provided to the IMF," leaving the responsibility for their accuracy and reliability with the subscribing country. This setup implies that a country may be in full observance of the standards, and reported as such in the DSBB, while at the same time providing faulty data—potentially in breach of its obligations under Article VIII, Section 5 (see Box 2 above). This could have clear repercussions for the credibility of the dissemination initiatives.

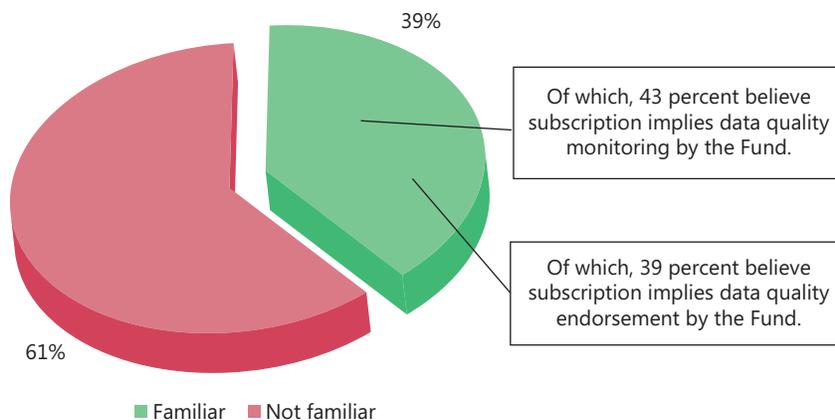
100. Data users' misperception regarding the endorsement of data quality by the Fund spreads to the dissemination initiatives. While the evaluation found a remarkable lack of familiarity with and use of the initiatives (Figure 12), more important was the finding that, among those who are familiar, a significant number believe that a country's participation in the GDDS or SDDS implies that the Fund is monitoring and/or endorsing the data quality.

101. Data producers, especially in low-income and emerging market countries, expressed positive opinions regarding the impact of subscription on dissemination practices, data quality, and third parties' perception of national data (Figure 13). Their opinions were more mixed, however, regarding the effect of these initiatives on access to financial markets. Empirical analysis for this evaluation (De Resende and Loyola, 2016) could not find convincing evidence of the effects of the SDDS on subscribers' gross foreign direct investment inflows, exchange rate volatility, or sovereign borrowing costs, in contrast to the findings in some earlier work by IMF staff.

... and collaborates with international partners in statistics.

102. In addition to its work on standards and methodologies, the IMF has a long history of collaboration with other international organizations in the statistical realm, including on allocation of data responsibilities, sharing of data, reduction of overlapping data requests

Figure 12. Familiarity with the GDDS and SDDS
(In percent)



Source: IEO Survey of Data Users.

Figure 13. Survey Results: “Subscription to the SDDS/GDDS improved my country’s ...”
(In percent)



Source: IEO Survey of Country Authorities and Data Providers.

to countries, donor coordination to address data deficiencies at the country level, and achieving data consistency among the various organizations (IMF, 1995c). This collaboration took on renewed impetus from the increased attention to statistical issues brought by the global financial crisis, and led to the launching of the G20 Data Gaps Initiative in 2009. International partners of the IMF hold, almost unanimously, a high opinion of the IMF’s collaboration. Recent examples of collaboration include:

- The Fund’s joint work with the Financial Stability Board (FSB) on the Data Gaps Initiative (DGI).

While stakeholders view positively its potential contribution to crisis prevention, the ambitious goals and open-ended nature of the DGI are creating a growing sense of fatigue among participants, with the risk of a loss of momentum.

- The Inter-Agency Group on Economic and Financial Statistics (IAG), chaired by IMF staff, was created in 2008 to address the growing need for coordination on statistical matters, including to help limit duplication of efforts at the international level. According to interviewees, the IAG has made limited progress to date in reducing countries’ data

reporting burden arising from duplicative data requests from various international organizations.⁹³ This slow progress is, in part, due to technical challenges with the Statistical Data and Metadata Exchange (SDMX) platform (see below), but also, to a lesser degree, “protecting one’s turf” among institutions.

- The Statistical Data and Metadata Exchange (SDMX)—a joint initiative by the BIS, ECB, Eurostat, IMF, OECD, World Bank, and UN—aims to foster the efficient exchange of data and metadata by adopting common standards and guidelines, together with information technology systems that would facilitate a move from the current “push” system for data reporting (i.e., countries must send their data to each institution) to a “pull” system

⁹³The survey (and interviews) of data providers indicated that 65 percent (and almost three-quarters among advanced economies) still experienced duplication in the data requests from IAG members.

(i.e., countries upload their data to a single web-based repository, and institutions draw on the data as needed). When fully implemented, this could greatly reduce the data reporting burden for member countries and facilitate a much more timely provision of data to analysts.⁹⁴

- The IMF Statistical Forum—created in 2013 and hosted by STA—is intended to become a space where data users, data providers, and policymakers come together to discuss cutting-edge statistical issues. However, so far, these events have been almost exclusively the domain of data providers.⁹⁵

⁹⁴The Open Data Platform for Africa, developed by the IMF in partnership with the African Development Bank is SDMX-based. During interviews, African authorities assessed very positively the impact of this initiative on the standardization and streamlining of data submissions, reducing the reporting burden.

⁹⁵For example, although all Fund staff have been invited to attend, non-STA Fund economists largely have ignored these forums, illustrating their indifference towards statistical issues.