

## 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 financial<sup>12</sup>—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).<sup>13</sup> 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 surveillance<sup>14</sup> 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.<sup>15</sup> 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).

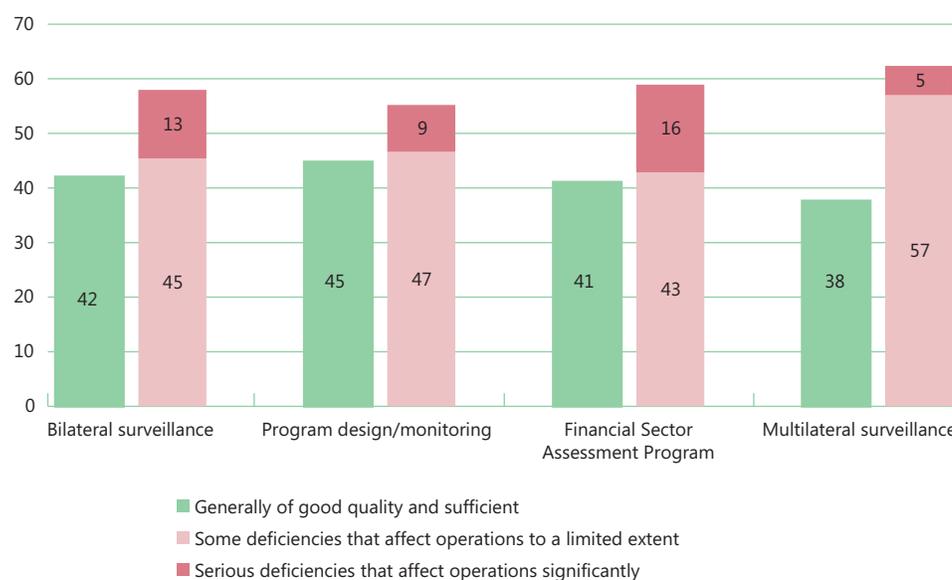
<sup>13</sup>Although this report focuses on surveillance and lending, data deficiencies also can have a bearing on other important areas of Fund work, such as calculating quota shares to guide decisions regarding relative size and distribution of members’ actual quotas.

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

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

<sup>12</sup>The 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.<sup>16</sup> 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.<sup>17</sup>

<sup>16</sup>Most 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.

<sup>17</sup>*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.<sup>18</sup>

### **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,

<sup>18</sup>Intentional 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.<sup>19</sup> 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):<sup>20</sup> (i) lack of candor in staff's discussion of data deficiencies, with an "upward bias in the

<sup>19</sup>Until 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.

<sup>20</sup>The 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.<sup>1</sup> 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

<sup>1</sup>Note 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.<sup>21</sup> 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.

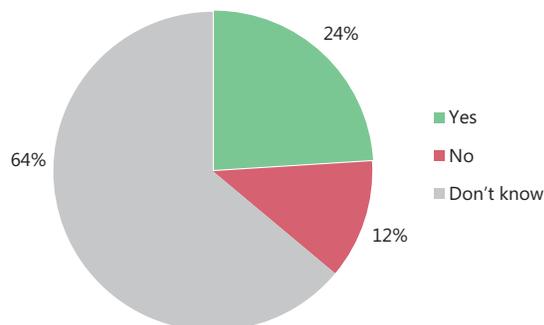
<sup>21</sup>As senior IMF staff members pointed out to the evaluation team, the Board's "lack of attention" to data quality issues at times reflected peer protection and political considerations.

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,<sup>22</sup> 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.<sup>23</sup>

<sup>22</sup>As 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.

<sup>23</sup>Interviews 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).<sup>24</sup>

<sup>24</sup>The 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 countries<sup>25</sup>—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

<sup>25</sup>In 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.<sup>26</sup> 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,<sup>27</sup> 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.”<sup>28</sup>

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

<sup>26</sup>The 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.

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

<sup>28</sup>The 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.