

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.