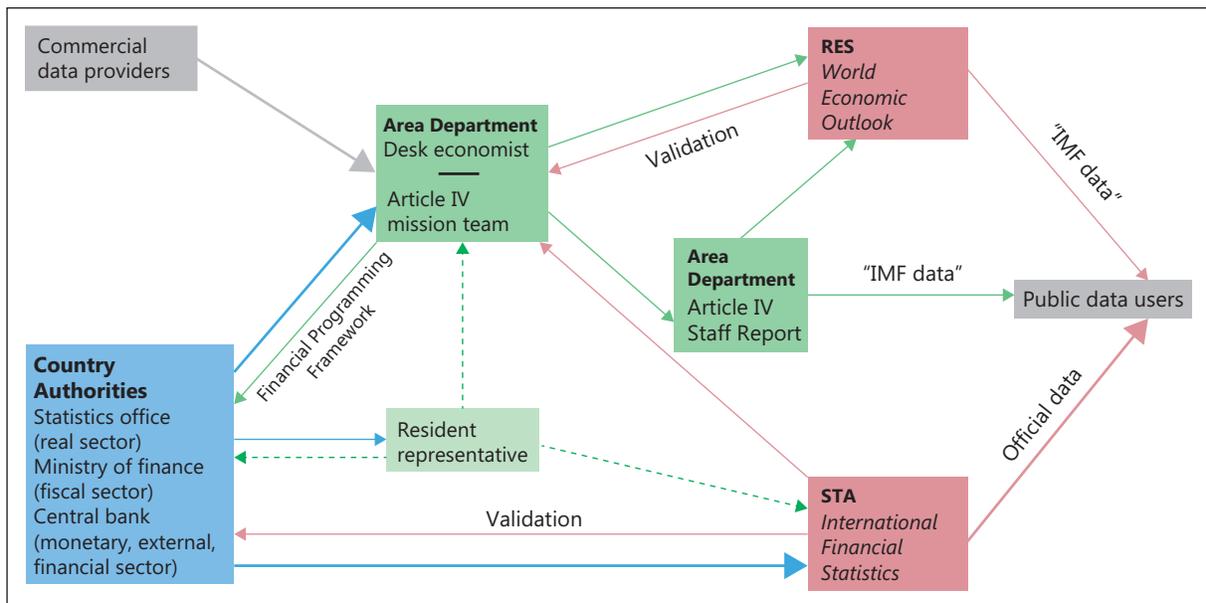


Figure 5. Data Flows at the IMF



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.<sup>68</sup>

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

<sup>68</sup>For 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."

<sup>69</sup>In 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.<sup>70</sup>

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.<sup>71</sup> 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,<sup>72</sup> 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)<sup>73</sup> ([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 ...*

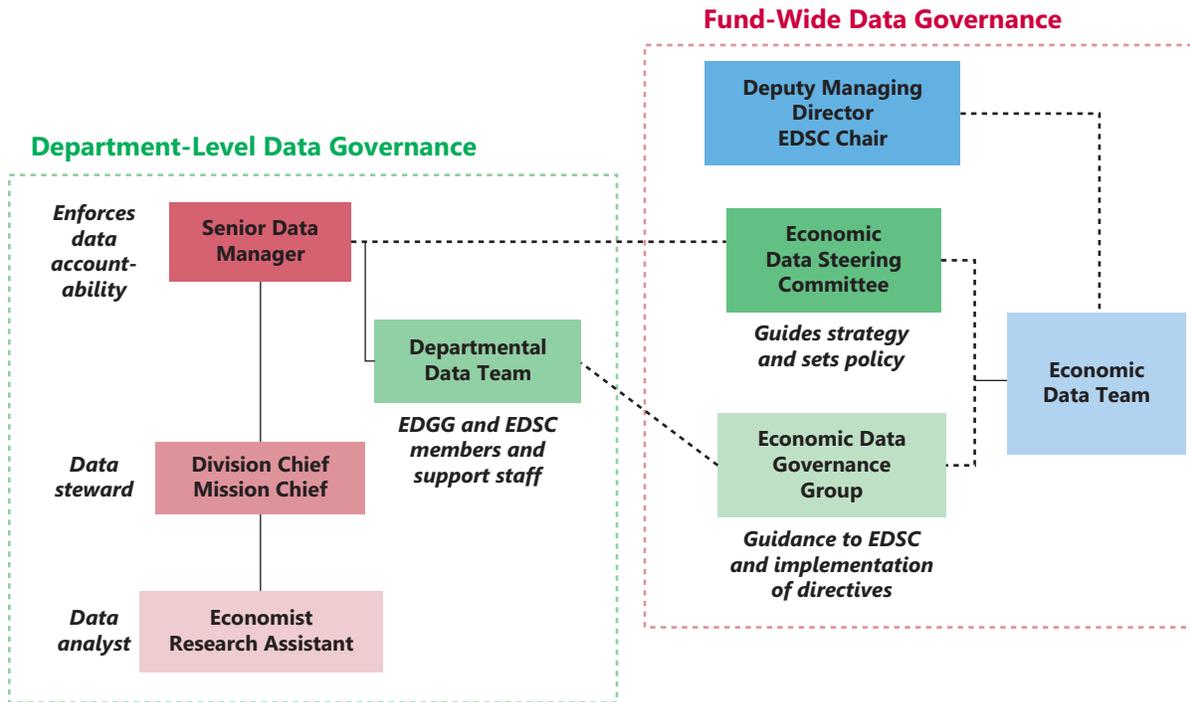
74. Staff’s assessment of the new data governance structure is mixed. Among those surveyed staff familiar with it, a majority believe it has been helpful in improving data management practices, but many think it “just adds another layer of bureaucracy.” In interviews, members of the EDSC and EDGG expressed 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.

<sup>70</sup>The 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.”

<sup>71</sup>See also De Las Casas and Pedraglio (2016).

<sup>72</sup>Gartner 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.

<sup>73</sup>The 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 Fund<sup>1</sup>

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.

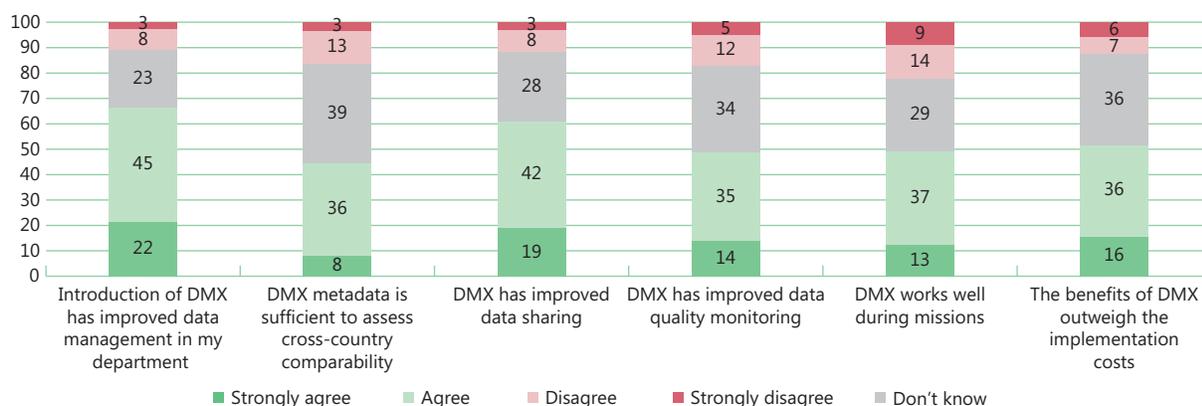
In the post-EDMI context, with the migration to structured databases completed, the EDT has turned to the creation of the Common Surveillance Databases (CSD)—with the explicit aims of facilitating data sharing, integrating data used for bilateral and multilateral surveillance exercises into one common 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.

<sup>1</sup>Based 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,<sup>74</sup> 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,<sup>75</sup> arguing that STA might be better suited for this task, given its core expertise.<sup>76</sup> 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

<sup>74</sup>The 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.

<sup>75</sup>Minutes 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.

<sup>76</sup>The 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.<sup>77</sup> 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),<sup>78</sup> 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

<sup>77</sup>Indeed, 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.

<sup>78</sup>An 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.<sup>79</sup> 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.<sup>80,81</sup>

<sup>79</sup>On 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.

<sup>80</sup>From 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."

<sup>81</sup>Staff 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.<sup>82</sup> 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.<sup>83</sup> 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.

<sup>82</sup>While 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.

<sup>83</sup>In 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.<sup>1</sup> 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.

<sup>1</sup>The 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.

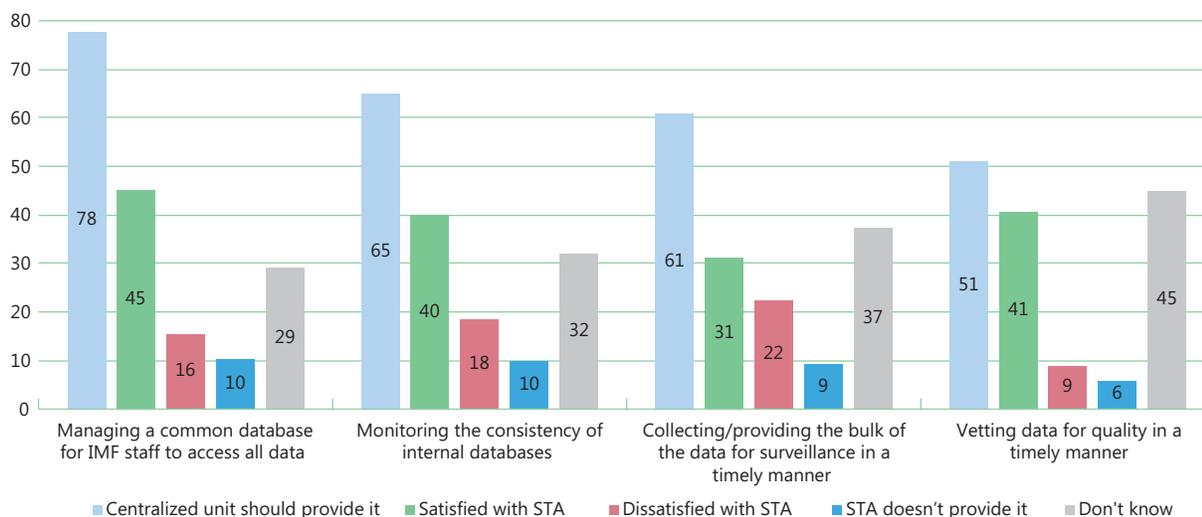
- This unit also ensures that (preliminary) data are available to analysts with minimal delay.
- A common nomenclature is used across all series stored in official databases, and this nomenclature is maintained by the centralized data unit.
- Desk economists use the institution's database because they 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)



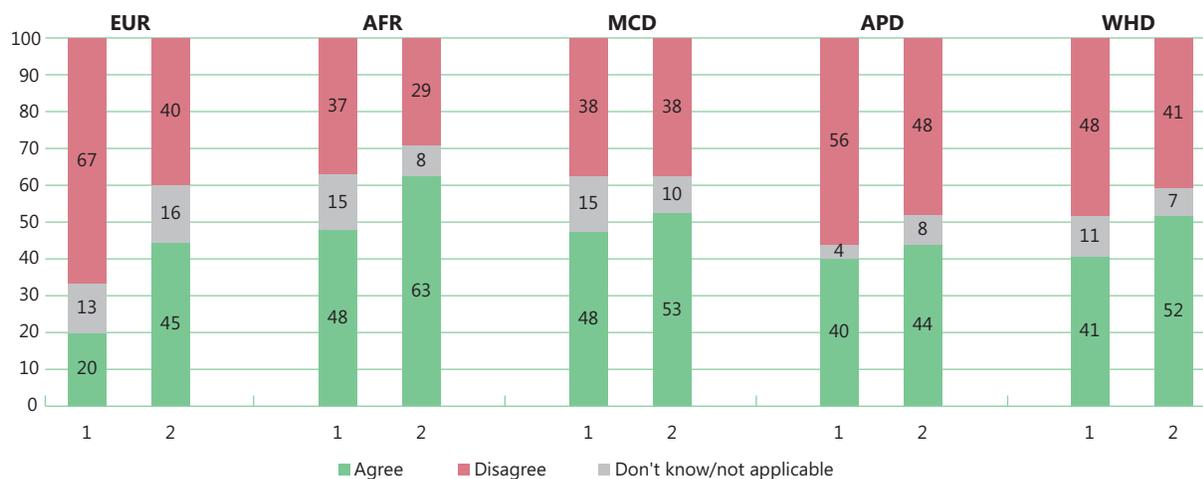
Source: IEO Survey.

### ... 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.<sup>84</sup>

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:

<sup>84</sup>In 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