

## The Quality of IMF Forecasts

*Prediction is very difficult, especially about the future.*

—Attributed to Niels Bohr, Danish physicist and philosopher

46. This chapter assesses the quality of IMF forecasts during 1990 to 2011, a period that included episodes of relatively sustained global economic growth as well as global, regional, and country-specific crises or recessions.<sup>22</sup> Like virtually all studies that have evaluated IMF forecasts, it focuses on short-term forecasts, that is, those made for the current year and one year ahead.<sup>23</sup> The analysis covers the IMF membership as a whole, in order to investigate whether the forecast quality varied systematically by region or level of economic development. For reasons already explained, the focus is primarily on forecasts of GDP growth.

47. The quantitative analysis is restricted to *WEO* forecasts rather than forecasts published in Article IV consultation reports. First, *WEO* forecasts are more frequent and are issued at regular intervals (twice a year, at roughly the same dates for virtually the whole membership)—which facilitates their comparison with those of other agencies that release forecasts on a regular basis and for many economies. Second, *WEO* forecasts have been analyzed in commissioned studies of IMF forecasting performance since the 1980s (see Section E below), allowing comparisons to be made with those studies and an assessment of how the IMF learns from its past forecasting performance. Third, the *WEO* data are more readily available, being organized in a comprehensive dataset, than data on Article IV forecasts.<sup>24</sup> Finally, as explained in Chapter 3, except for reasons

related to timing there should be no substantial differences between *WEO* and Article IV forecasts, since their preparation follows the same general process.

48. The analysis concentrates on point forecasts. Clearly, informed views about the future require more than just point forecasts: risk scenarios and the analysis of driving forces behind the path of the variables forecast are also important; they are highly valued by country officials according to the survey conducted for this evaluation, and are being increasingly incorporated in IMF flagship documents and Article IV consultation reports in response to the recommendations from the commissioned external evaluation of IMF forecasts by Timmerman (2006). But point forecasts can nonetheless be viewed as the basis, or starting point, for such broader sets of considerations about future economic developments.

49. With these considerations in mind, the assessment of the quality of IMF forecasts is based on three separate metrics—informational efficiency (Sections A and B), accuracy (Section C), and perceptions by country authorities and the private sector (Section D). Section E considers the importance of learning from past forecast performance for the quality of forecasts, and the IMF's current practices in this respect, and section F provides an overall assessment.

### A. Are Forecasts Biased?

50. An issue frequently raised about IMF forecasts in the academic literature and in interviews with country authorities is whether they are systematically biased. The most recently published study of the quality of IMF forecasts commissioned by the IMF (Timmermann, 2006), covering the period 1990 to 2003, concluded that

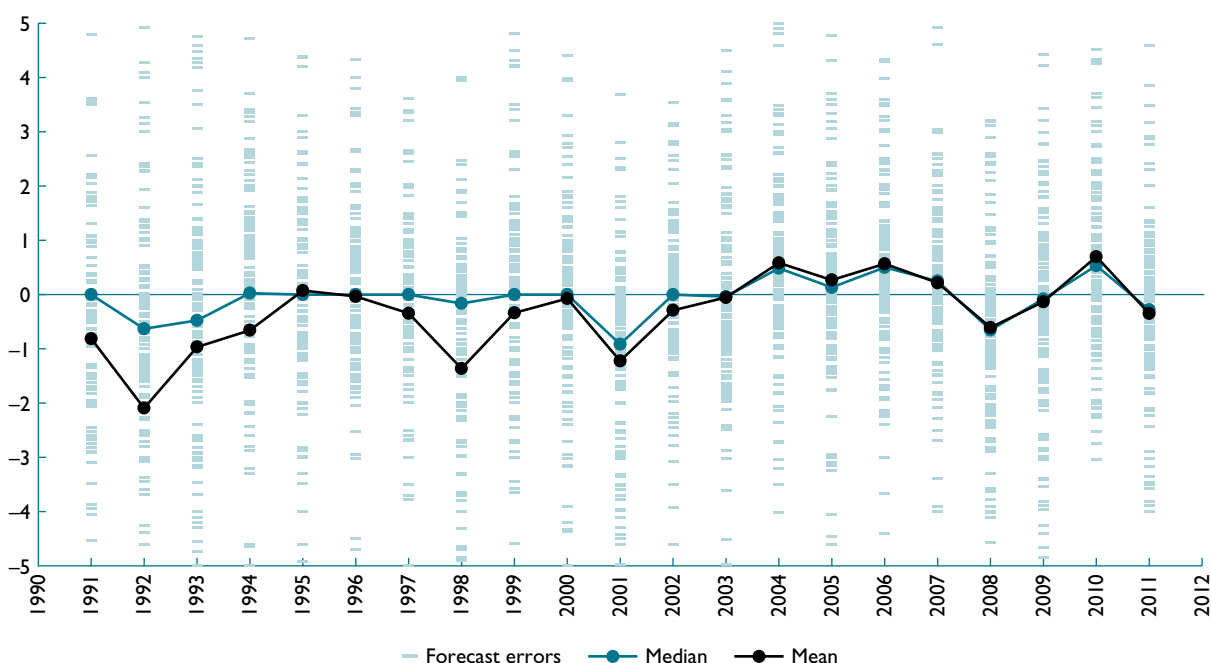
<sup>22</sup>This chapter draws on a more detailed treatment in Genberg and Martinez (2014b).

<sup>23</sup>Medium-term forecasts are discussed in Chapter 5.

<sup>24</sup>To our knowledge, there is no unified dataset that contains data on Article IV forecasts. Recent *WEO* forecasts are easily accessible through the IMF website. Each individual release of the *WEO*'s forecasts is available for more than a decade dating back through 1998, either in the statistical appendixes of the relevant *WEO* publication or in its corresponding database. It is much harder to access historical

forecasts from the *WEO*. Despite the fact that the *WEO* has produced forecasts since 1971, and has published them since 1980, the IMF website provides no information on the forecasts prior to the late 1990s. For additional detail, see Genberg and Martinez (2014b, Annex 1).

**Figure 7. Forecast Errors of GDP Growth**  
(In percentage points)



Source: IEO calculations using the IMF's Spring *World Economic Outlook* current-year forecasts.

“forecasts of real GDP growth display a tendency for systematic overprediction” (p. 6). While several other studies concur with Timmerman’s assessment (e.g., Artis, 1988; and Faust, 2013), a number of authors draw the opposite conclusion or find no evidence of bias.<sup>25</sup>

51. Among the reasons for the different conclusions are differences in the choice of sample period, the countries included in the analysis, and whether or not program countries are included in the sample. Some examples of the implications of these choices follow.

52. Figure 7 illustrates how conclusions can vary depending on the choice of sample period. It shows the errors in GDP growth forecasts for each of 144 member countries as well as the cross-country averages and medians, calculated year by year. The figure makes clear that studies based on cross-country averages and samples that are heavily weighted by the 1990s and early 2000s will tend to find negative cross-country average forecast errors, that is, an optimistic bias.<sup>26</sup>

<sup>25</sup>See Genberg and Martinez (2014b), Section II.B(i) for detailed references.

<sup>26</sup>See Genberg and Martinez (2014b) for details about the sample and the calculations. Forecast errors are calculated as the actual outcome minus the forecast. A negative forecast error for economic growth can thus be labeled an optimistic forecast.

Extending the analysis into the 2000s will include underestimations of GDP growth observed in the middle of the decade—so much so that the overall bias for the whole period since the early 1990s becomes quite small.

53. As shown in Annex 1, for advanced, emerging market, and low-income economies the general message is the same: average forecast errors vary over time, tend to be negative (optimistic) in the 1990s, and larger than zero (pessimistic) for a number of years in the mid-2000s. For low-income economies the average forecast errors also vary over time, but they are more consistently negative (optimistic) than for the other two types of economies (see also Table 1 below).

54. Figure 7 also illustrates other features of the forecast errors that should be kept in mind when inferences are drawn about the nature of IMF forecasts. In particular, even though a majority of the errors cluster around zero in a range of plus-minus 2 percentage points to 3 percentage points, there are significant numbers of errors of a much larger magnitude. As will be discussed in more detail below, these are often associated with economic crises or recoveries therefrom. As a consequence, the cross-country mean can be heavily influenced by outliers. The cross-country median (the blue circles in the figure) is less affected by outliers and

is therefore typically significantly closer to zero than the mean.<sup>27</sup>

55. For individual G20 economies, as for the membership as a whole, overpredictions of GDP growth are the most frequent outcome (Figure A2.1 and Table A2.1 in Annex 2), although there are considerable variations over time and across countries also in this group.<sup>28</sup> Underestimations of inflation are much less frequent among G20 economies than in the membership as a whole. Among these economies, especially emerging market economies, inflation tends to be mostly overpredicted.

56. A recurring feature of forecast errors is the particularly large negative values during regional and global recessions such as the crisis in the European Union in 1992, the Asian crisis in 1997–98, the end of the dot-com bubble in March 2000, and the financial crisis of 2007–09.<sup>29</sup> Table 1 shows how recessions decisively affect the measure of biases in short-term GDP forecasts. For instance, consider the spring vintage of next-year forecasts, which shows the largest optimistic biases. The bias, measured by the median forecast error, ranges from about –0.3 percentage point to –0.5 percentage point, depending on countries’ level of development and their IMF program participation status.<sup>30</sup> However, when the highly optimistic biases observed for recession years (ranging from –4 percentage points to –7 percentage points) are excluded from the sample, optimistic biases are eliminated, reversed, or substantially reduced.<sup>31</sup>

57. Juhn and Loungani (2002) showed that the onsets of recessions are difficult to forecast, as judged by the spectacular failure of private sector forecasters to do so. The IEO evaluation team carried out calculations using these authors’ methodology, focusing on the forecast record of the IMF. The results are equally telling: neither the IMF nor the private sector has been able to forecast the onset of recessions very well.

<sup>27</sup>A feature of the data not visible in the figure but which can be verified by statistical analysis is that sequences of individual country forecast errors are typically not serially correlated even though the cross-country averages appear to be. (See de Resende, 2014; and Genberg and Martinez, 2014a.)

<sup>28</sup>Statistically significant average underpredictions are only observed for China and a few isolated cases in other emerging market economies.

<sup>29</sup>The large average errors in the 1992, 1998, and 2001 forecasts are visible in Figure 7 for the entire membership. Figure A2.1 in Annex 2 shows that significant forecast errors were made in the 2007–09 crisis for a number of G20 countries.

<sup>30</sup>As already noted, overpredictions are particularly prevalent in low-income countries.

<sup>31</sup>Similar results are obtained if outliers (either positive or negative) are eliminated from the sample, a practice that has been suggested in the literature on forecast evaluation. See Genberg and Martinez (2014b), Section II.C(iii), for details.

**Table 1. Median Forecast Errors in GDP Growth, 1991–2011**  
(In percentage points)

	Year-Ahead Forecasts		Current-Year Forecasts	
	Spring	Fall	Spring	Fall
Full sample	–0.29	–0.20	0.00	0.00
For all recessions	–6.27	–5.73	–3.69	–1.52
For nonrecessions	0.00	0.00	0.09	0.07
Advanced countries	–0.30	–0.19	0.07	0.10
For all recessions	–4.14	–3.64	–0.90	–0.16
For nonrecessions	0.00	0.12	0.15	0.11
Emerging and developing countries	–0.04	0.00	0.06	0.10
For all recessions	–6.50	–5.92	–3.38	–1.44
For nonrecessions	0.30	0.30	0.30	0.23
Low-income countries	–0.50	–0.42	–0.20	0.00
For all recessions	–7.03	–6.89	–5.00	–3.59
For nonrecessions	–0.11	–0.08	0.00	0.00
IMF program countries	–0.43	–0.30	–0.05	0.00
For all recessions	–7.03	–6.93	–4.94	–1.94
For nonrecessions	–0.07	0.00	0.00	0.06

Source: IEO calculations using IMF, *World Economic Outlook*.

58. Is it possible to identify institutional factors that explain why large forecast errors tend to be particularly clustered around regional or global recessions? While it is clear that some events may be unpredictable, Juhn and Loungani (2002) argue that private sector forecasters’ inability to predict recessions could arise from a lack of incentives to do so. Within the IMF, whose internal forecasting process may discourage forecasts that “rock the boat,” as noted in Chapter 3, there is little incentive to forecast a recession when neither the private sector nor previous forecast rounds have done so. As part of the Fund’s review process, staff forecasts are checked against those of other forecasters and need to be justified if they are different. Although asking for such justification is perfectly legitimate, desk economists can minimize the amount of scrutiny their forecasts will receive by not differing significantly from the consensus forecast.<sup>32</sup> While this scrutiny operates symmetrically, the cost of forecasting a recession that does not materialize may be perceived as higher than the cost of having wrongly predicted a boom.<sup>33</sup> And efforts

<sup>32</sup>Similar arguments were made in the evaluation of *IMF Performance in the Run-Up to the Financial and Economic Crisis* (IEO, 2011): “The evaluation found that incentives were not well aligned to foster the candid exchange of ideas that is needed for good surveillance” (para. 55), “. . . expressing strong contrarian views could ‘ruin one’s career’” (para. 56).

<sup>33</sup>As noted in de Resende (2014), a complementary explanation would rely on the empirical observation that recessions tend to occur more abruptly and be associated with temporary shocks, while booms are more gradual and frequently related to permanent shocks.

to convince colleagues and supervisors may not seem to promise a large enough pay-off, even if the forecast is ultimately proven right. It should also be noted that forecasting a recession may entail high costs if doing so would in fact precipitate a recession.

59. Optimistic biases are reduced as more information becomes available (Fall vintages) and are typically smaller for shorter forecast horizons (current year) (Table 1). An implication of these findings is that revisions of forecasts, for example from the Spring *WEO* to the Fall *WEO*, typically reduce biases. Timmermann (2006) and Faust (2013) found similar results with respect to forecast accuracy: revisions made in *WEO* forecasts as more information became available regularly led to a reduction in the size of forecast errors.<sup>34</sup>

60. The fact that biases are critically affected by recessions and vary both over time and across regions makes it difficult to argue that there is a consistent institutional bias in IMF forecasts, either optimistic or pessimistic. In addition, as argued by Faust (2013), statistical tests of unbiasedness, accuracy, and overall efficiency may be a poor assessment of the quality of forecasts when there are relatively frequent structural changes in the economies for which forecasts are produced.

61. Finding that biases in *WEO* forecasts of GDP growth are not systemic at the institutional level should not be a reason for complacency, however. Lack of bias only means that positive and negative forecast errors tend to cancel each other out over time. It does not mean that forecast errors are small<sup>35</sup> or that there are no possibilities for improvement in individual countries.

## B. Are Forecasts Efficient?

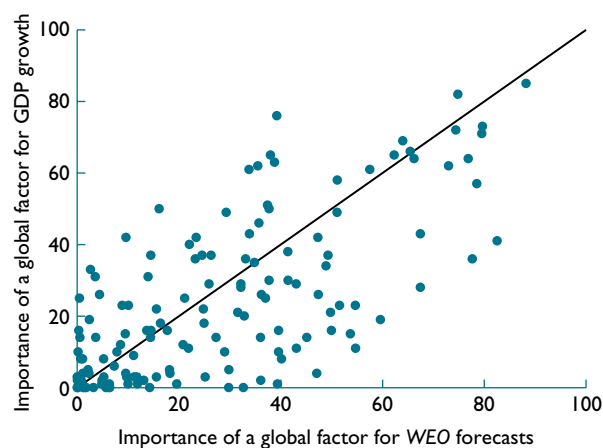
62. Efficiency of forecasts is a wider concept than bias and refers to whether or not the forecasts take into account “all available information.” In the context of multi-country forecasts a particularly interesting question relating to efficiency is whether the forecasters in each individual country take proper account of interdependencies between member countries. Timmermann (2006) showed that forecast errors in the *WEO* are in part explainable by the forecasts of U.S. and German

GDP growth that were available when the forecasts for other countries were made. This result indicates that some interdependencies may not be fully incorporated in all *WEO* forecasts. Timmermann’s (2006) results still hold true for *WEO* forecasts when the sample period is extended to 2011. In addition, information in forecasts for China’s GDP growth also does not appear to have been adequately incorporated in forecasts for some other countries in this extended sample.<sup>36</sup> See Genberg and Martinez (2014b).

63. These results should not be taken to mean that *WEO* forecasts systematically ignore interlinkages between countries. Indeed, the evaluation finds strong evidence that interlinkages are taken into account, albeit perhaps not fully. Consider Figure 8. On the vertical axis it shows a measure of how important global developments are for GDP growth in an economy. One hundred percent would indicate that all of the fluctuations in the economy can be accounted for by a global factor that is common to all countries. Zero percent would mean that fluctuations are completely country (or region) specific. This measure has been calculated for a large number of IMF member countries and reported in Matheson (2013).

64. The horizontal axis measures how important global developments are for *WEO* forecasts of GDP growth for the same countries and time period as those

**Figure 8. Interdependence in the Data and in Forecasts**



Sources: Genberg and Martinez (2014b) and Matheson (2013).

<sup>34</sup>de Resende (2014) contains similar findings with respect to medium-term forecasts: the shorter the forecast horizon the greater the accuracy.

<sup>35</sup>Faust (2013) reports that median and mean year-ahead Fall forecast errors of GDP growth in 2009 in advanced countries were more than 4 percentage points, roughly the same as in our sample (3.6 percentage points).

<sup>36</sup>Chapter 5 below shows that this type of informational inefficiency is also present in medium-term forecasts.

**Table 2. Average Difference in Publication Dates with WEO, 2000–10<sup>1</sup>**  
(In number of days)

	Organization for Economic Cooperation and Development	Consensus Economics <sup>2</sup>	European Commission	World Bank	Economist Intelligence Unit <sup>2</sup>
Spring/summer	62	–1	2	57	11
Fall/winter	87	–2	42	75	24

Source: Genberg and Martinez (2014b).

<sup>1</sup>A positive number indicates that the WEO forecasts are published first.

<sup>2</sup>Publication dates for these institutions were chosen to minimize difference from WEO publications given that their forecasts are released on a more frequent basis.

studied by Matheson. For reasons having to do with the frequency of forecasts relative to the frequency of actual data, the two measures are not identical, but Genberg and Martinez (2014b) show that the two should be positively correlated if WEO forecasts incorporate the global forces identified by Matheson. Inspection of Figure 8 shows this to be the case, and results reported by Genberg and Martinez (2014b) show that the visual impression holds up to statistical scrutiny.

65. We conclude that while WEO forecasts do incorporate linkages among economies to a significant degree, these linkages may still not be fully accounted for in all forecasts. The global economy evolves over time as economies become more linked to each other through trade in goods, services, and financial instruments. Forecasters aiming to incorporate interdependencies among economies are thus shooting at a moving target. IMF desk economists are no exception in this respect, and they need to keep adapting their models and judgment to incorporate new realities. The WEO forecasting process contains elements that are designed to increase individual desk economists' awareness of relevant international developments. In view of the potential inefficiencies mentioned by Timmermann (2006) and confirmed in this evaluation, these elements may need to be strengthened.

### C. Are Forecasts Accurate?

66. The IMF's WEO forecasts are often viewed as a benchmark to use in comparisons with other national and international forecasters. A survey conducted for IEO (2006) found that almost 88 percent of country authorities either agreed or strongly agreed that they "consider the WEO's projections to be the benchmark for assessing economic prospects." More recently, the survey conducted for the present evaluation found that about two-thirds of country authorities who responded either agreed or strongly agreed with the statement that they "use WEO forecasts to check the

accuracy of [their] own forecasts" (Genberg and Martinez, 2014a).<sup>37</sup>

67. Differences in release dates among forecasters can influence the determination of relative forecast performance, especially when a later forecast can incorporate an earlier forecast's information.<sup>38</sup> As shown in Table 2, relative to its main forecast comparators the WEO is released relatively early in each forecasting cycle. This means that the IMF's Fall forecast may be published up to three months before the OECD's forecast—which would give the OECD and other forecasters time to incorporate the IMF's forecast as well any new information that may emerge in the interim. While these timing differences could markedly affect relative forecast performance, only a few past studies of IMF forecasts make more than a passing note of differences in production dates.

68. There is less of a publication timing issue when comparing WEO forecasts with private forecasts such as those issued by Consensus Economics. This is largely because private forecasters produce their forecasts monthly and thus the publication date can be selected so as to minimize the timing differences.

69. For this evaluation Genberg and Martinez (2014b) compared the accuracy of WEO and Consensus Economics forecasts of GDP growth using the most recent data available. Looking across all countries in the comparisons for each category of forecasts, the results show that there is little to differentiate between WEO and Consensus in the Spring forecasts, whether these are for the current year or the year ahead. For the Fall forecasts the results are very sensitive to the vintage chosen for Consensus forecasts. If the September

<sup>37</sup>The survey as well as follow-up interviews with country officials also revealed that IMF forecasts are typically judged to be more accurate than forecasts made by other international organizations. For the domestic economy, country authorities typically view their own forecasts as being more accurate than those of the IMF.

<sup>38</sup>A further complication results from potential differences between the date on which the final forecast was established and the ultimate release/publication date.



forecast is used the *WEO* has a slight edge, whereas if the Consensus October forecast is used, the opposite is true.

70. Focusing more narrowly on G20 countries, IMF forecasts of GDP growth are very similar to Consensus forecasts (Figure A2.3 in Annex 2).<sup>39</sup> For almost all G20 economies, forecast errors for any given year have the same overall pattern and size, and display the same turning points in both cases. This goes against the notion of an organizational bias in IMF forecasts.<sup>40</sup>

## D. User Perspectives on the Quality of IMF Forecasts

71. When asked about IMF forecasts in general, a majority of country authorities responded that they believed they were unbiased. Only a small minority expressed the opposite view. To a more specific question about the accuracy of *WEO* growth forecasts for their own country, three-quarters of country officials responded that they believed these forecasts were “about right.” Six percent believed they were “consistently too high” and 18 percent said they were “consistently too low.” Respondents working in global financial institutions had less sanguine views about the accuracy of *WEO* forecasts: 50 percent believed that they were “about right,” 27 percent “consistently too high,” and 23 percent “consistently too low.”

72. These survey results are interesting because they suggest that country authorities by and large do not question the quality of IMF forecasts. Of course one can argue that when 24 percent of officials feel that *WEO* growth forecasts are consistently either too high or too low, something is amiss. It is also noteworthy that, regarding GDP growth, three-quarters of the officials who feel that IMF forecasts are biased think that growth forecasts are too pessimistic rather than too optimistic.

<sup>39</sup>This raises the question whether IMF forecasts are so close to Consensus forecasts that they do not contain any independent information. Timmermann (2006) investigates this possibility and concludes that it is not the case; both forecasts carry useful information about future growth of GDP and inflation. Luna (2014b) presents results with a similar interpretation for program cases. Hence, the correspondence between IMF and Consensus forecasts is likely to be the result of both using similar (but not exactly the same) information and forecasting methods.

<sup>40</sup>It does not hide the occurrence of occasionally large forecast errors for most individual G20 economies, many of which are associated with recessions or crises and the subsequent recovery.

## E. How Does the IMF Learn from Past Forecasting Performance and Experience?

73. Learning from experience takes place at many levels, individually and institutionally, formally and informally, through introspection and in response to external review, routinely, and in response to significant failures. This section assesses initiatives taken at the institutional level and at the level of individual desk economists to learn from past forecasting performance.

### Commissioned studies: objectives and impact<sup>41</sup>

74. Since the 1980s the IMF Research Department has commissioned four studies by outside experts to evaluate the quality of *WEO* forecasts: Artis (1988, 1996), Timmermann (2006), and Faust (2013).<sup>42</sup> Barriounevo (1993) has been treated, in all the subsequent studies, as part of this series of assessments even though this study was produced inside the IMF. The first study (Artis, 1988) responded to concerns by Executive Directors about bias in IMF forecasts. Each of the subsequent studies was intended as an update of the preceding ones using the most recent data available and, particularly for the last two studies, to provide recommendations for improving the forecasts.

75. The Fund did not put in place a formal process defining what is expected from each successive study; how the results of the study are to be communicated to staff, Management, and the Board; how staff and Management should respond to the recommendations in the study; or how the follow-up should be implemented and documented.

76. The lack of such a process makes it difficult to judge whether practices at the IMF have changed as a result of these external reviews. Freedman (2014) concludes that though some specific changes could be attributed directly to one of them (Timmermann, 2006), it was difficult to pinpoint more generally the effects of the various evaluations on the behavior of forecasters and the way they go about their business.

77. In response to questions by Freedman, senior IMF officials who had been involved in the *WEO* process at the time of the various commissioned studies suggested

<sup>41</sup>This section draws extensively on Freedman (2014).

<sup>42</sup>The Faust (2013) study has not yet been published. All references to that study in this evaluation refer to a draft dated February 5, 2013.

that the studies had helped build an internal consensus about the need to update the Fund's forecasts more often. They quoted the introduction of mini- or mid-term *WEOs* and the increased use of alternative scenarios as examples of how the *WEO* process had become more responsive to changes in global economic conditions.

78. Freedman (2014) identifies several issues related to commissioned studies that have not received sufficient attention. He points to the absence of a structured process to facilitate learning from these commissioned studies and monitor the implementation of their recommendation at the institutional level or at the level of individual desk economists. He also asks whether the forecasting process achieves the right balance between top-down and bottom-up elements.

### Experience of desk economists

79. Country forecasts by the IMF are ultimately the product of country desk economists. From interviews and a survey of IMF country desk staff, it is clear that one common aspect of how forecasts are produced is the reliance on the judgment of the desk economist.

80. Judgment relevant for forecasting can be sharpened by on-the-job learning; by assimilating the knowledge of a predecessor; by investigating and learning from past forecast errors; and by attending specialized formal training. The evaluation team gathered information on each of these elements to assess the nature and effectiveness of forecast-related learning by country desk economists.

### *The relationship between experience of desk economists and forecast accuracy*

81. The experience of desk economists has a significant effect on forecast accuracy. Numerous studies of security analysts in the private sector have found such a relationship, and this evaluation finds similar results for IMF staff: both country-specific and general experience is associated with improved forecasts (see Genberg and Martinez (2014b) for details and Box 4 for a summary). Survey results and interviews with desk economists corroborated the statistical findings. As one desk economist said: “[a]t the beginning [it is] very useful to rely on what is there, while you learn [about] the economy, only [over time] can you think about improving [the forecasts].”

### *Transfer of knowledge from incumbent to successor*

82. Given how dependent desk economists are on the methods used by their predecessor when they first join a country desk (see Chapter 3, Section E), it is important that the transition between desk economists function smoothly.

83. This is not always the case. In interviews, most staff indicated that transitions between country desks were ad hoc and varied substantially from person to person (Figure 9, panel a). While some thought the process worked satisfactorily, and several thought that the standardization of spreadsheets through DMX (Data Management for Excel)<sup>43</sup> had led to improvements, many expressed

<sup>43</sup>DMX is an extension of Excel that provides tools and services to help with macroeconomic data management where data are stored in the form of time series, formulas, and tables.

### **Box 4. The Relationship Between Staff Experience and Forecast Accuracy<sup>1</sup>**

A unique internal IMF dataset was used to compare *WEO* short-term GDP forecast errors for a large set of countries over the period 2007–11 against the experience levels of the desk economists who produced those forecasts. The analysis distinguished among different types and levels of experience (previous country desk assignments, tenure at the IMF, and the attendance of IMF training courses related to forecasting), different groups of countries, and different forecast horizons.

The regression results indicate that greater staff experience is associated with lower absolute forecast errors. They suggest that both country-specific and general experience help improve the forecasts.

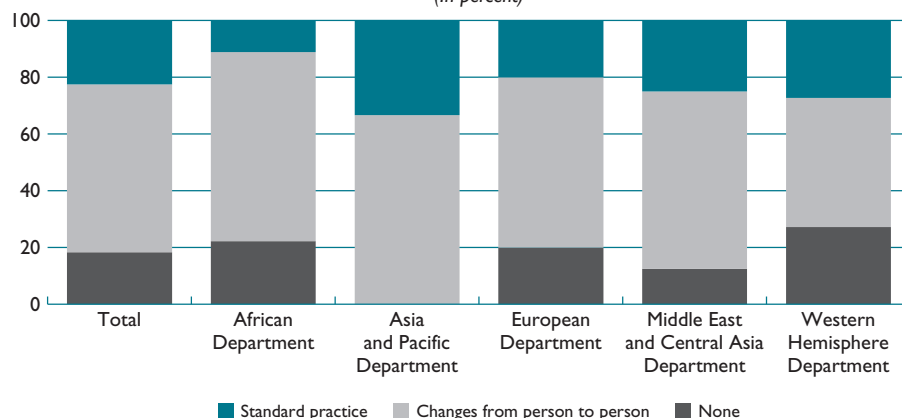
<sup>1</sup> See Genberg and Martinez (2014b), Section III.B.

However, the results are not uniform across all types of countries. While country-specific experience is associated with an improvement in forecast performance for low-income countries, it appears to have little relation with forecast performance in advanced and emerging economies. A possible explanation is that the use of judgment is much more prevalent in IMF forecasts of low-income countries, which tend to have a limited amount of data available and few (if any) other external forecasters.

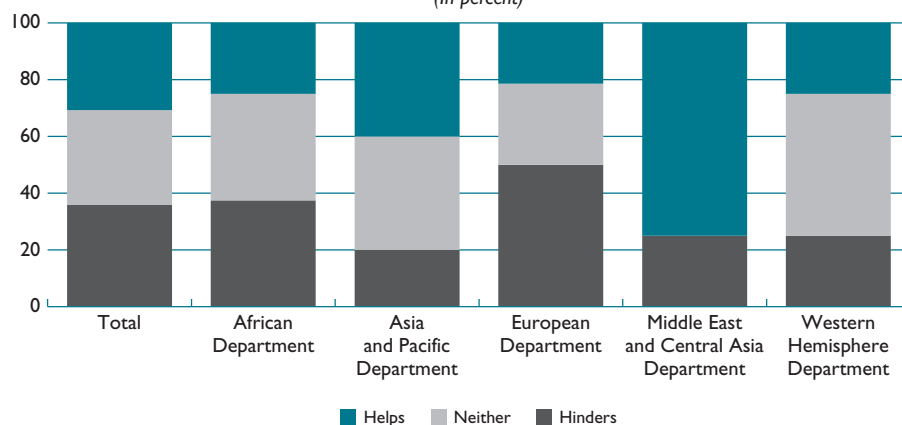
The results also suggest that increases in a desk economist's general work experience and training are related to improvements in forecast accuracy. Mission-chief tenure does not appear to have a significant effect on forecasting performance.

**Figure 9. Preserving Historical Memory**

**a. What Mechanisms Exist for Transmitting Accumulated Knowledge of the Country Between the Old and the New Desk Economist?**  
(In percent)



**b. Does the Way in Which Data and Models Are Passed from the Incumbent to a New Desk Economist Help or Hinder the Process of Generating Forecasts?**  
(In percent)



Source: Genberg and Martinez (2014b).

frustration with how much variation there was. Several desk economists expressed the view that the only thing facilitating transitions between country desks was “good will” on the part of the incumbent desk economist.

84. Around 40 percent of the staff interviewed thought that the ad hoc nature of the transfer of information from outgoing to new staff on country desks hindered the forecasting process (Figure 9, panel b).<sup>44</sup> Some argued that the lack of a standard transition mechanism

helped perpetuate the status quo and led to inertia in making changes. Others said that “a tremendous amount of information gets lost” because there is no standard way to convey this information. A common theme, however, was that the efficiency of the process of passing information is highly dependent on personalities, and that some more formal system would be desirable.<sup>45</sup>

<sup>44</sup>Figure 9 may in fact underrepresent the concerns expressed by staff. Several staff said that the way information was transferred between desks was helpful to the production of forecasts only when the process functioned well—which is not always guaranteed.

<sup>45</sup>This has been a recurrent issue in IEO evaluations. See IEO (2009, 2011b, and 2013). In 2013 the Strategy, Policy, and Review Department issued an internal checklist/guidance note for country assignment handover within the department, to ameliorate the handover process ([www-intranet.imf.org/departments/SPR/OGR/Pages/default.aspx](http://www-intranet.imf.org/departments/SPR/OGR/Pages/default.aspx)). It is too early to tell whether this will have the hoped-for effects, and whether it will be implemented also in area departments.



### Learning from past performance

85. Learning about how an economy functions and evolves can also be achieved through a careful examination of past forecast performance, which can be informative about the appropriateness of a chosen forecast method.<sup>46</sup> Because a large majority of desk economists indicated that forecast accuracy is an important consideration in their choice of a forecast method, it might be expected that assessments of past performance would be conducted regularly. But only 50 percent of the desk economists responding to the survey said that such an analysis had been conducted during their tenure on the desk. About a quarter indicated that they analyzed forecast errors once a year or after each forecast round, 15 percent had analyzed forecast errors at least once, and 10 percent responded that they did not know whether a forecast performance assessment had been carried out.

### Learning from formal training<sup>47</sup>

86. A final aspect of learning relates to participation in formal training courses on forecast methodology. The IMF appears to be the only international organization to provide training on forecasting to its staff. The Fund's Institute for Capacity Development (ICD) provides in-house courses on topics that range from the basic needs of IMF staff to specialized topics presented by renowned external experts.

87. Has the formal training affected forecast performance? The evaluation survey asked desk economists about their own perception of the usefulness of forecast-related courses. Of those who had attended such courses, about 20 percent felt that the training had not influenced their ability to produce better forecasts. An equal percentage responded that it had led to a great improvement, while the remainder perceived the courses as having led to some improvement.

88. In follow-up interviews it emerged that staff saw the value of attending specialized courses on forecasting as limited because (i) the courses are “too academic” and not immediately relevant for the desk work; and (ii) desk economists are too busy with operational work to attend such courses (in particular the longer ones), especially because the institution does not give the right incentives to participate in such events.

89. A number of interviewees suggested that specific tools should be developed to tackle the forecasting needs of desks exposed to different situations dictated by data quality and availability. Training events should then be organized to teach the use of such tools.

## F. Assessment

90. Are IMF forecasts accurate and efficient? Is the IMF learning from past forecast performance? The evaluation finds that:

- Though optimistic biases in forecasts occur in all country groupings—and tend to be larger in low-income countries and in certain program countries—these biases are highly sample-sensitive and do not seem to be systemic or associated with the way the institution conducts its forecasts. In particular, an entrenched inability to predict recessions, which is not particular to the IMF but also plagues competitor forecasters, is critical in explaining the source of measured optimistic biases.
- IMF forecasts take account of interdependencies among economies, but not fully. Forecasts of GDP growth in China, Germany, and the United States, for example, have explanatory power for forecast errors.
- The accuracy of IMF short-term forecasts compares well with that of other institutions providing multi-country forecasts. As for perceptions, the majority of country officials and private sector analysts surveyed for this evaluation seem to trust the integrity of forecasts and generally do not feel that IMF forecasts are biased.
- Learning is the area where the evaluation found more room for improvement. First, while the experience with regular commissioned studies has been positive, the process for disseminating and implementing their recommendations is not fully developed. Second, IMF economists do not frequently and systematically check the past forecast performance for their countries, though this could be a valuable source of learning. Third, experience matters for better forecasts, especially when these are heavily based on judgment, but the relevant experience is not always transmitted effectively between successive country desk economists. Finally, staff comments in the survey and interviews suggest that in-house training is not sufficiently practical to be directly applicable in the economists' daily work.

<sup>46</sup>Faust (2013) stresses the particular importance of examining and learning from forecast errors in periods of significant structural change possibly brought about by events such as the recession of 2007–09.

<sup>47</sup>This section is based on Luna (2014a).