

## Medium-Term Forecasts

*You've got to be very careful if you don't know where you're going, because you might not get there.*

—Attributed to Yogi Berra, American baseball player and amateur philosopher

91. The importance of medium-term forecasts for critical IMF products, and their significant methodological differences relative to short-term forecasts, motivate a separate assessment of medium-term forecasts—that is, those for three, four, and five years ahead.<sup>48</sup> Like the previous chapter, the analysis concentrates on GDP growth forecasts reported in the *WEO*. Forecasts of medium-term growth rates are important to several policy-relevant items in the IMF's surveillance and lending activities. Increasingly, many IMF products—including not only the *WEO*, *GFSR*, and *Fiscal Monitor*, but also spillover reports, pilot external sector reports, and early warning and vulnerability exercises—deal with medium-term issues. These products contain analysis, risk assessments, and scenarios that extend to longer horizons and rely on medium-term IMF forecasts as inputs.<sup>49</sup> The quality of medium-term forecasts significantly affects the conclusions drawn in these products, including assessments of the sustainability of fiscal policy frameworks, estimates of equilibrium real exchange rates, measures of sustainable (and desirable) current account positions, and risks to financial and macroeconomic stability, all of which influence the IMF's policy advice.

92. This chapter first documents the views on medium-term forecasts gathered from a survey of country authorities, the private sector, and IMF staff, and the importance of these forecasts for key products and activities of the Fund. Section C discusses methodological difficulties of estimating a country's potential output, which help explain why, despite their importance, medium-term growth forecasts are inherently uncertain.

Section D describes the medium-term forecasting methods and process in use in the IMF, and section E reviews the quality of the forecasts, applying the methodology used in Timmermann (2006). Like that in Chapter 4, the analysis is based on point or central forecasts. Section F provides an overall assessment.

### A. User and Staff Perspectives on Medium-Term Forecasts

93. The survey conducted by the evaluation team found that more than half the respondents from the private sector think that for policy discussions IMF medium-term forecasts are more important than short-term forecasts. The corresponding percentage among country authorities is about one-third. Very few respondents in either group said that they ignored medium-term forecasts.

94. In interviews, country officials, especially those from less developed economies, stressed the importance of paying attention to medium-term forecasts. Some noted explicitly the potentially damaging effect that misleading medium-term forecasts may have on the IMF's analysis for both surveillance and lending purposes.<sup>50</sup>

95. Yet IMF country desk economists who responded to the survey feel that medium-term forecasts entail too much uncertainty to constitute a reliable guide for policy discussions. In addition, during post-survey interviews with IMF staff a large number of interviewees gave the impression that medium-term forecasts—although integrated with short-term forecasts

<sup>48</sup>This chapter draws on a detailed treatment in de Resende (2014).

<sup>49</sup>See Annex 3 for more details.

<sup>50</sup>For example, a senior official from a major emerging market economy argued that medium-term forecasts are not sufficiently based on countries' structural and demographic characteristics, and that exchange rate assessments and current account projections, including those carried out in the context of external balance assessments, are therefore likely to be inaccurate.

in the Fund’s macro framework—are less carefully made and certainly use fewer resources than short-term forecasts.

96. These findings should not be taken to imply that the IMF as an institution disregards longer-term analysis more broadly defined. As noted, medium-term scenarios and risk analyses are increasingly important in the IMF’s flagship and bilateral surveillance reports. But if the point forecasts for medium-term GDP growth that ultimately drive the risk scenarios are not given appropriate attention, the risk assessments themselves may be compromised.

## B. The Importance of Medium-Term Forecasts in IMF Surveillance Products

97. This section describes the importance of having unbiased central forecasts for three important IMF products: debt sustainability analysis, pilot external balance assessments, and risk assessments based on fan charts. While these products deal with issues and involve techniques that are different from those of medium-term forecasting, they generally use medium-term forecasts as inputs or as base-lines. For this reason the quality of the medium-term forecasts will have an impact on the quality of these other products.

### Debt sustainability analysis

98. Projecting the evolution of a country’s government debt, relative to its underlying ability to generate income—often summarized in the debt-to-GDP ratio—is a key element in debt sustainability analysis, an activity in which medium-term forecasts of GDP growth are obviously important. Erroneous forecasts may produce a distorted view of the future debt level and lead to misguided policy advice today. Variations as small as 1 percentage point, which are not uncommon, in forecasts of medium-term real growth and inflation can make a difference between a sustainable and an explosive path for the debt-to-GDP ratio. Both de Resende (2014) and IMF staff research (see IMF, 2004) suggest that poorly made projections for longer horizons may greatly undermine assessments of debt sustainability.<sup>51</sup>

<sup>51</sup>The Fund’s recently revised framework for debt sustainability analysis in market access countries (IMF, 2013) recognizes explicitly the various sources of uncertainty that can impact debt projections.

### External balance assessment

99. The pilot external balance assessment (EBA) is another important IMF product that uses medium-term, specifically five-year-ahead, forecasts of GDP growth. The analysis in the EBA of appropriate levels of real exchange rates and current account balances is partly based on an empirical estimate of the relationship linking five-year-ahead growth differentials (between the country being examined and a GDP-weighted “world counterpart”) and the observed real exchange rate and current account balance. This relationship is used, together with projections about future growth differentials and normative considerations about current policies, to assess whether the current level of the real exchange rate and the current account balance are appropriate.<sup>52</sup>

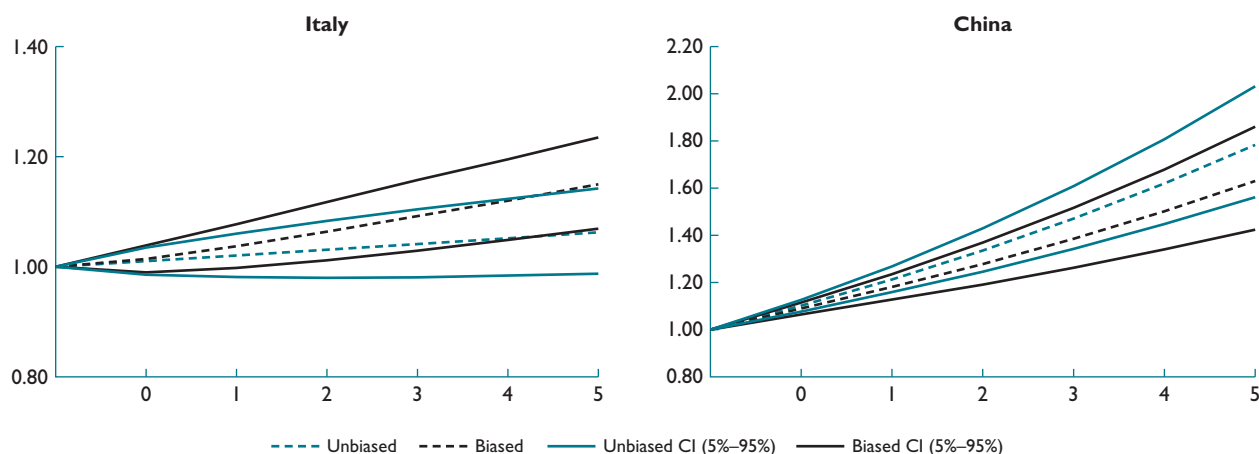
### Risk assessments based on fan charts

100. Assessments of risks to the global economy, to specific regions, and to individual economies are highly valued aspects of IMF analysis. Fan charts have become a popular device in this context, showing the range of possible future values a variable may take given the uncertainty associated with the point forecast.<sup>53</sup> Such fan charts have become regular features of the *WEO* and certain *Regional Economic Outlooks*, and have been used occasionally also in Article IV country reports. Central banks routinely use them to illustrate the uncertainty around their forecasts.

101. A bias in the central forecast can have important consequences for the assessment of risks. This is illustrated in Figure 10 for two G20 countries; Italy, which Timmermann (2006) identifies as a country for which the IMF has regularly produced optimistic forecasts, and China, for which IMF forecasts have tended to be pessimistic on average (Annex 2, Table A2.1). The importance of using a correct central forecast is clear in both cases. For Italy the 90 percent confidence interval using the unbiased forecasts shows that there is a non-negligible possibility that output would decline during

<sup>52</sup>de Resende (2014) shows how biases in the growth forecast for a country or for its trading partners can change the assessment of the norm for the real effective exchange rate, potentially leading to situations in which the rate can be judged under- or overvalued relative to the norm, when it is actually in equilibrium.

<sup>53</sup>The midpoint of a fan chart is usually determined by the central (or point) forecast, and the uncertainty associated with the central forecasts determines the width of the fan. This statement is strictly true only if the distribution of the possible outcomes is symmetric.

**Figure 10. Fan Charts Around Biased and Unbiased GDP Forecasts***(GDP level (year  $-t = 1$ ))*

Source: IEO calculations.

Note: The blue lines in the figures illustrate 90 percent confidence intervals around the unbiased forecasts of the level of real GDP five years into the future, and the black lines show the 90 percent confidence intervals around the biased forecasts. Since the true probability distributions, from the model used in the forecasts, are unknown to the evaluation team, the confidence intervals are based on the standard deviation of forecast errors across different WEO vintages (Spring) for the different horizons, assuming a normal distribution. For further details on the construction of the fan charts, see de Resende (2014).

the forecast period (see the lower blue line in Figure 10), whereas the risk analysis carried out with the biased forecast would put a very low probability on such an outcome (see the black lines in Figure 10). Similarly, for China the upper limit of the fan constructed using the biased (black) forecast would barely exceed the midpoint of the fan using the unbiased forecast. In both cases the consequences of using inaccurate central forecasts for risk assessments are significant.

### C. Measures of Potential Output and the Output Gap

102. A key difference between short- and longer-term forecasts is the relative reliance they place on cyclical versus structural determinants. The longer the forecasting horizon, the greater the importance attached to structural factors. Longer-term forecasts of GDP growth, in particular, try to identify a trend that is often associated with the long-run aggregate supply curve, usually interpreted as the level of potential output, to which the economy reverts when the effects of temporary shocks that cause cyclical fluctuations dissipate. Since medium-term forecasts are more likely than short-term forecasts to abstract from these cyclical factors, forecasting GDP growth over the medium term requires an idea of the level to which the economy will converge and of the speed of convergence to

this level. Because of the importance of the notion of potential output for medium-term forecasts, Box 5 contains a brief description of the methods available to estimate it.

103. During interviews that the evaluation team conducted with staff, it became clear that having an estimate of the level of potential output is a critical step in the process of obtaining medium-term forecasts of GDP growth and other variables. The estimate of potential output is also needed for the calculation of the output gap—the difference between actual and potential GDP—which is a key indicator of the degree of slack in the economy and is typically used in short-term forecasts of inflation and the measurement of cyclically adjusted fiscal and current account balances, factors that are critical in the IMF's policy advice to member countries.

104. A problem associated with virtually all of the methods of estimating potential output is that the results are sensitive to updating the sample on which they are based.<sup>54</sup> Estimates of past potential growth rates will in general change when new data become available, implying considerable uncertainty in the original estimates. This is illustrated in Figure 11, which plots the growth rate of potential GDP in the United States as recorded in different WEO vintages. Relative to the

<sup>54</sup>This is the case particularly at the end of a sample period when such estimates are the most important for forecasts and policy analysis.

### Box 5. Methods to Estimate Potential Output<sup>1</sup>

Methods for estimating potential output fall into three broad classes. At one end of the spectrum are purely statistical univariate approaches that only use the information contained in the GDP series itself to generate estimates of potential output.<sup>2</sup> Their simplicity and more limited data requirements allow these approaches to be applied to a wide range of economies, but their lack of economic content introduces difficulties in identifying the trend and cycle components of the GDP series, which are interpreted as potential output and the output gap, respectively.

<sup>1</sup> For details, including references to research carried out by IMF staff, see de Resende (2014), Section III.

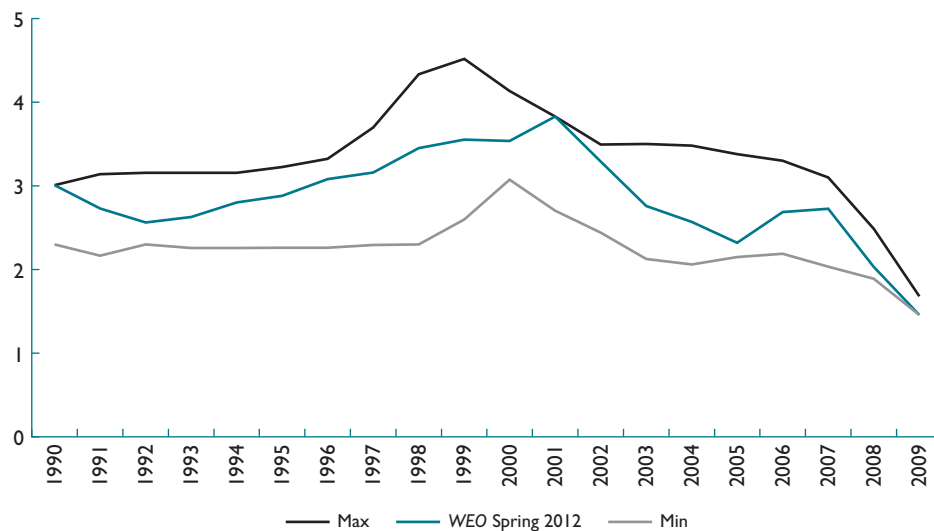
<sup>2</sup> An example of this approach is the popular Hodrick-Prescott filter.

At the other extreme, structural methods rely on a fairly detailed model representation of the economy and typically use data from several variables in the estimation process. These methods are more theoretically coherent, but are also substantially more data-intensive and difficult to implement.

In between these two polar cases are bivariate and multivariate approaches that represent a compromise between the greater theoretical coherence and data intensity of structural methods and the simplicity and lack of economic content of univariate methods.

All methods have shortcomings and there can be no “one-size-fits-all” approach. The proper method to use depends on data availability and on the structure of the economy being analyzed, which may be changing, and should be complemented by the forecaster’s overall knowledge of the country and judgment.

**Figure 11. Growth Rate of Potential GDP in the United States Across Different WEO Vintages**  
(In percent)



Source: de Resende (2014).

growth rates of potential GDP published in the Spring 2012 WEO (blue line), the estimates for almost all years in the sample change substantially across different WEO vintages. For example, the Spring 2012 WEO estimates the growth rate of potential GDP in 1998 to have been 3.5 percent, but the estimates for the same year have been as low as 2.3 percent (Spring 1998) and as high as 4.3 percent (Spring 2010), a difference of 2 percentage points. Such wide variation in estimates of past growth rates of potential GDP growth is also seen

in other economies and may be interpreted as a measure of how uncertain these estimates are.<sup>55</sup> The inherent difficulties in estimating potential output translate into substantial uncertainty and revisions also in measures

<sup>55</sup> The narrower difference between minimal and maximal estimates of the U.S. potential growth rate at the end of the sample in Figure 11 should not be taken to imply more precise estimates. It is rather due to the smaller number of WEO vintages used to compute the mean, minimal, and maximal potential growth rates for more recent years.

of the output gap and may therefore, distort the policy advice in real time that is based on it.

## D. Estimation Methods and Process

105. The process of generating medium-term forecasts by country desk economists is similar to that for short-term forecasts in that both types of forecasts are integrated into the macro framework used in the day-to-day work of desk economists.

106. However, while for short-term growth forecasts the IMF has a process in place to promote global consistency, it does not appear to have a comparable process to construct a view of the growth potential of major regions and economies based on fundamental structural determinants. In particular, no interdepartmental committee is in charge of maintaining analytical consistency among medium-term forecasts at the country, regional, and global levels.

107. The evaluation survey and post-survey follow-up interviews with IMF staff, as well as analysis of Fund documents, indicate that in making medium-term forecasts of GDP growth IMF economists use (or have used) all classes of methods described in Box 5 for estimating potential output, appropriately complemented by a widespread use of judgment. The heterogeneity of methods partly reflects the heterogeneity of the IMF membership in data availability, structural change, development stage, etc., but may also be explained by the lack of top-down guidance of the type provided for short-term forecasts.

108. IMF economists stressed that their choice of method for medium-term forecasting depends on data availability and on the structure of the economy they cover. Systematic assessments of past forecast performance do not appear to influence the choice of methods. More generally, desk economists often indicated that they have little incentive to search for better methods since medium-term forecasts are not viewed as particularly important.

109. Free choice of methods is not a problem if economists are well informed about the methods and the best way to use them in different countries or situations, and if the right incentives to seek information about these methods are in place. Interviews with staff indicated, however, that these conditions are not always met, and that having some guidance on the methods would be desirable.<sup>56</sup>

<sup>56</sup>Other organizations such as the OECD and the EC have processes in place to coordinate medium-term forecasts and to provide a consistent

## E. The Quality of IMF Medium-Term Forecasts

110. Overall, the evaluation finds that IMF medium-term forecasts for GDP growth in the *WEO* meet the basic forecasting efficiency standards in most countries, with little evidence of a built-in organizational bias.<sup>57</sup>

111. However, problems exist in the medium-term forecasts for a notable number of member countries, many of them the same as those Timmerman (2006) identified in the context of current- and next-year forecasts. *WEO* medium-term forecasts have a tendency to overpredict GDP growth. In a universe of 180 countries over the sample period 1990–2011, between two-thirds and three-fourths show predicted growth rates on average higher than actual growth rates. In 20 percent to 30 percent of the countries this bias is statistically significant. Measured on an annual basis, the average bias ranges between 0.14 percentage point and 0.76 percentage point, depending on the forecast horizon, the measure used (median or mean), and the method (descriptive or regression-based statistics from either country or panel regressions). As in short-term forecasts, the existence of bias in medium-term forecasts of GDP growth seems to be largely a reflection of the inability to predict crises and large recessions.

112. The tendency to overpredict medium-term GDP growth is present in economies across all IMF area departments—except for the Middle East and Central Asia Department, where underprediction is the norm—across different levels of development, and regardless of IMF program participation status. Statistically significant biases are more frequent among economies in the African Department and in emerging market and low-income economies. The magnitude of the optimistic biases is however greater in advanced economies, reflecting the fact that large biases are concentrated in a few G7 economies.<sup>58</sup> Medium-term program

view on potential output developments. Both these organizations have a smaller, less heterogeneous membership than does the IMF, so a narrower choice of forecasting methods may be appropriate. In the case of the EC, member countries indeed require homogeneity of forecasting methods because of the preeminent role of cyclically adjusted fiscal balances in the institutional quantitative assessments for these countries. A central unit within the EC coordinates the efforts of the teams producing medium-term forecasts for individual economies, with a view to assuring not only accounting, technical, and statistical consistency, but also analytical and economic consistency.

<sup>57</sup>See de Resende (2014), which contains an analysis of the quality of *WEO* medium-term growth forecasts following the methodology in Timmermann (2006).

<sup>58</sup>The average optimistic bias for G7 economies as a group is always substantially larger than that for the full sample.



forecasts of GDP growth for countries with a history of IMF-supported programs over the sample period (1990–2011) are more optimistic than for nonprogram countries by a sizable and statistically significant margin.<sup>59</sup> Program-related optimistic biases tend to be concentrated in the year the program starts and to be corrected within one year.

113. Though the tendency towards overoptimism in IMF medium-term forecasts of GDP growth exists in the statistical sense, the qualifications with respect to the reliability of statistical tests of accuracy and efficiency emphasized in Faust (2013) and already noted in the discussion of short-term forecasts should be kept in mind.

114. In the IMF's medium-term forecasts serial correlation is less frequent than bias—which limits the scope for using past errors to help improve forecasts. However, for a notable share of countries, the evaluation finds that forecasts of GDP growth in the United States, Germany, or China, as well as of oil prices, can help explain medium-term forecast errors and thus could in principle be used to improve forecasts. This result mirrors the findings for short-term forecasts reported in Chapter 4. This type of inefficiency is more frequent in medium-term than in short-term forecasts—possibly because efforts to account for international spillovers and interrelationships between countries in *WEO* forecasts, as previously recommended by Timmerman (2006), seem to have been implemented more in short-term than in medium-term forecasting,<sup>60</sup> and possibly because less effort is made to coordinate medium- than short-term forecasts.

115. Finally, while IMF medium-term forecasts of GDP growth are clearly more accurate than naïve (no change) and mechanical (e.g., Hodrick-Prescott filter) forecasts, they are somewhat less accurate than private sector forecasts as published by Consensus Economics.

## F. Assessment

116. To assess the quality of IMF medium-term forecasts, this chapter has relied on two metrics—the appropriateness of the methods and procedures used and the quality of IMF medium-term forecasts of GDP growth in terms of informational efficiency and accuracy.

117. The chapter concludes that:

- The methods used in medium-term forecasts of GDP growth for individual economies are broadly appropriate, but sufficient coordination for better analytical consistency at the regional and global levels is lacking.
- The overall quality of IMF medium-term forecasts for GDP growth is acceptable, although there is more room for improvement relative to short-term forecasts.
  - Forecasts are unbiased for most countries. Optimistic biases are found in a nonnegligible share of the membership, including many G20 countries, in all but one IMF area department, in countries at all stages of development, and regardless of IMF program participation status. As with short-term forecasts, however, these biases do not seem systemic and are critically influenced by an entrenched inability to predict recessions, which is not particular to the IMF.
  - IMF medium-term forecasts outperform forecasts made using mechanical and naïve methods (suggesting that staff judgment adds quality to forecasts), but they are somewhat less accurate than private sector forecasts.
  - Serial correlation is not frequent in IMF medium-term forecasts, but there are substantially more signs of informational inefficiency—related to insufficient consideration of spillovers and interdependencies in the global economy—than in short-term forecasts.
- Central or point forecasts are appropriately complemented by a broader set of discussions about medium-term issues—including risk assessments described with fan charts and alternative scenarios to the baseline projections based on the central forecasts. These discussions are increasingly important in IMF flagship documents, notably the *WEO*, and are also included in other multilateral products such as spillover reports and pilot external sector reports. Country authorities highly value analyses using risk assessments and scenarios as well as the point forecasts on which they are based.

<sup>59</sup>Using a different dataset, Luna (2014b) investigates biases in the context of programs and reaches similar conclusions, which are discussed in the next chapter.

<sup>60</sup>Another possibility is that forecasts made subsequent to Timmermann's study are not well reflected in the sample used to assess medium-term forecasts, because the corresponding actual values are not yet known.