# An Evaluation of the *World Economic Outlook* Forecasts

ALLAN TIMMERMANN*

*This paper conducts a series of statistical tests to evaluate the quality of the* World Economic Outlook *(WEO) forecasts for a very large cross section of countries. It assesses whether forecasts were unbiased and informationally efficient, characterizes the process whereby WEO forecasts get revised as the predicted period draws closer, and compares the precision of the WEO forecasts to private sector forecasts known as ''consensus forecasts'' and published by* Consensus Economics *on a monthly basis. The results suggest that the performance of the WEO forecasts is similar to that of the consensus forecasts.* [JEL E17, E37, F17, F47]

The *World Economic Outlook* (WEO) is a significant source of forecasts of global economic activity and is a key vehicle in the International Monetary Fund's multilateral surveillance activities. It is published twice a year, in April and September. Given the central role of the WEO forecasts, it is important that they are periodically evaluated to assess their usefulness and to look for ways to improve the forecasting process. This study is the fourth in a series of such evaluations (following Artis, 1988 and 1997; and Barrionuevo, 1993). There are some notable differences between the current

---

study and the earlier ones:

- First, we analyze forecasts for 178 countries in seven economic regions (Africa, Central and Eastern Europe, the Commonwealth of Independent States (CIS) countries and Mongolia, Developing Asia, the Middle East, the Western Hemisphere, and the Advanced Economies) since 1990. Earlier evaluations had focused on forecasts for only the Group of Seven (G-7) countries and regional aggregates.
- Second, we include an extensive comparison between the accuracy of WEO forecasts and consensus forecasts. The latter is a widely used source that compiles the forecasts of economists working in the private sector. Through this comparison, we assess WEO forecasts not just against absolute benchmarks, but also against a relative benchmark of other forecasters.
- Third, we consider the revisions to the forecasts, both over time and within each forecast round. The latter is important because there is a long gestation lag in the preparation of the forecasts in each round, and it is important to know the gains—in terms of accuracy—of frequent forecast updates.

Our analysis focuses on the current-year and next-year WEO forecasts of real gross domestic product (GDP) growth and inflation. In the case of real GDP growth, we find that the WEO forecasts display a tendency for systematic overprediction—that is, predicted growth, on average, tends to exceed actual growth. From a statistical perspective, these biases are most significant in the next-year forecasts. This tendency for overprediction of growth performance is also persistent over time. Moreover, the evidence suggests that forecasts of U.S. GDP growth are positively and significantly correlated with current-year forecast errors of output growth in a substantial number of advanced economies. (The forecast of German GDP growth also has predictive power over output growth forecast errors in some regions.) Our analysis also finds that, in some cases, accuracy problems appear related to the standing WEO assumption that the output gap is eliminated after 5 years. In particular, the paper notes a predominant negative relationship between the output gap and the forecast error in the GDP growth, notably for Germany, France, and Italy.

Turning to the inflation forecasts, we find a bias toward underprediction of inflation, with these biases significant in the next-year forecasts for many African, Central and Eastern European, and Western Hemisphere countries. The underprediction bias is generally found to be weaker in the current-year forecasts. With regard to their predictability, there is evidence that the next-year inflation forecast errors are often linked to U.S. GDP forecasts.

Prior to the publication of the WEO forecasts in April and September, a first set of predictions is presented to the IMF Executive Board in February and July. Subsequently, the forecasts are revised before they are published. These revisions add considerable informational value. For the February/

April same-year forecasts, the average reduction in the forecast error is about one-fifth for the advanced economies. The reduction is nearly 30 percent for the July/September same-year forecasts, but only 5 percent for the next-year forecasts.

The study compares the WEO projections to consensus forecasts for GDP growth and inflation over the period 1990–2003.[1] The data cover all the G-7 economies (Canada, France, Germany, Italy, Japan, the United Kingdom, and the United States), seven Latin American economies (Argentina, Brazil, Chile, Colombia, Mexico, Peru, and Venezuela), and nine Asian economies (China, Hong Kong SAR, India, Indonesia, Korea, Malaysia, Singapore, Taiwan Province of China, and Thailand). Overall, the comparison suggests that the forecast performance of the WEO is similar to that of the consensus forecast. The paper highlights, however, that the timing of the comparison with the consensus forecast matters. WEO current-year forecasts generally perform quite well against current-year consensus forecasts reported in March and perform considerably better against the February consensus forecasts. Given the relatively long gestation lag in their preparation, they tend to perform considerably worse against the consensus forecasts reported in April.

## I. Description of the WEO Data Set

### Data Coverage

To assess the forecasting performance, we make use of the fact that four sets of short-term forecasts are available for the same variable, because the WEO publishes both April and September current- and next-year forecasts. For example, four forecasts of GDP growth in the year 2000 are reported, namely, the April and September 1999 next-year forecasts and the April and September 2000 current-year forecasts. Access to different forecast vintages allows us to address issues such as whether (and by how much) the error in the forecast gets reduced as the time toward the target dates shrinks. It also allows us to test another efficiency property embedded in an optimal forecast, namely that forecast revisions should themselves be unpredictable. In some cases, we find evidence of significant biases in revisions, suggesting simple ways of improving on the forecasts.

The WEO data set contains information on 178 countries over the period 1990–2003. These countries are collected into seven groups or regions; namely, Africa (50 countries), Central and Eastern Europe (15), CIS and Mongolia (13), Developing Asia (24), Middle East (14), Western Hemisphere (33), and Advanced Economies (29). Data availability and data quality vary significantly across regions and there can be significant differences even

---

[1]The so-called "consensus forecasts" published by *Consensus Economics* on a monthly basis are forecasts for a number of macroeconomic variables. The first forecasts for the major industrial countries were published in October 1989. Since then, the coverage has expanded steadily and now includes many emerging market countries.

within each region. Data quality and the extent to which outliers affect the results also depend on the type of variables being analyzed.

## Timing Conventions

Because the target variables are subject to data revisions, a choice has to be made concerning which data vintage to use to measure realized values or outcomes. To this end, we follow the practice from earlier studies of WEO forecasts, such as Artis (1997), and use the first-available data in the April WEO issue of year $t+1$ to measure the outcome of the predicted variable in period $t$ (labeled $y_t$). Next-year forecasts for period $t+1$ are compared to the realized values for year $t+1$ ($y_{t+1}$) reported in the September WEO issue of year $t+2$. The idea here is that immediacy of an actual value against which the precision of the forecast is measured is particularly important for the short-term forecasts, so the first-available (April) measure is used for these forecasts. This is less of a concern for the longer term (next-year) forecasts, where the more precisely measured September data are consequently used.

In the analysis, we will also make use of the fact that we have both April and September forecasts of same-year and next-year realizations. This means that we have two sets of current-year forecasts generated in April and September, $\hat{y}_{t,t}^{Apr}$, $\hat{y}_{t,t}^{Sep}$, and two sets of next-year forecasts generated during the same months, $\hat{y}_{t+1,t}^{Apr}$, $\hat{y}_{t+1,t}^{Sep}$. In this notation, the first subscript indicates the period being predicted and the second subscript indicates the year when the forecast was generated. The superscript indicates the month of the WEO issue in which the WEO forecast was reported. This convention gives rise to four separate forecast errors:

$$e_{t,t}^{Apr} = y_t - \hat{y}_{t,t}^{Apr}, \qquad \text{April current-year forecast error}$$
$$e_{t,t}^{Sep} = y_t - \hat{y}_{t,t}^{Sep}, \qquad \text{September current-year forecast error}$$
$$e_{t+1,t}^{Apr} = y_{t+1} - \hat{y}_{t+1,t}^{Apr}, \qquad \text{April next-year forecast error}$$
$$e_{t+1,t}^{Sep} = y_{t+1} - \hat{y}_{t+1,t}^{Sep}, \qquad \text{September next-year forecast error.}$$

In addition, we consider current-year and next-year forecast revisions, defined as

$$rev_{t,t} = \hat{y}_{t,t}^{Sep} - \hat{y}_{t,t}^{Apr}, \qquad \text{revision to the current-year forecast}$$
$$rev_{t+1,t} = \hat{y}_{t+1,t}^{Sep} - \hat{y}_{t+1,t}^{Apr}, \qquad \text{revision to the next-year forecast.}$$

The data are trimmed in some regions because of missing observations or extreme observations that would otherwise dominate the regional averages. For example, at least eight September current-year forecasts are available for only 41 out of 50 African countries, and only 11 of the 24 developing Asian economies had more than eight data points for this variable. Fortunately, data on April and September next-year forecasts tend to be more complete,

although again there are some countries with incomplete data. Measured by data coverage, the data set is most complete for the advanced economies and least complete for CIS and Mongolia.

## II. Properties of Optimal Forecasts

To evaluate the quality of the WEO forecasts, it is necessary to establish a set of testable properties that an optimal forecast should have. In this section, we discuss the nature of such properties. In all cases, the properties are established under the assumption that the objective function is of the mean-squared error (MSE) type so that the forecasts minimize a symmetric, quadratic loss function. Different properties hold for other loss functions. To the extent that the costs associated with over- and underpredicting variables, such as GDP growth and inflation, are not symmetric, then it is, in fact, optimal to bias the forecast. Elliott, Komunjer, and Timmermann (2005) find that this has important consequences when evaluating the optimality properties of a forecast. Patton and Timmermann (2006) show how standard optimality properties that a forecast has under MSE loss get violated under asymmetric loss and a nonlinear data-generating process.

### Unbiasedness and Lack of Serial Correlation

Under MSE loss, the optimal forecast, $\hat{y}_{t,\tau}^{*} = \arg\min_{\hat{y}_{t,\tau}} E[(y_t - \hat{y}_{t,\tau})^2 | \Omega_\tau]$, where $\Omega_\tau$ is the forecaster's information set at time $\hat{o} < t$. Under broad conditions, such as the existence of expected loss and covariance stationarity of the forecast error, we have $E[e_t | \Omega_\tau] = 0$, which implies unbiasedness of the optimal forecast and absence of serial correlation in the forecast errors. Define the generic forecast errors for period $t$ or $t+1$ as

$$e_t = y_t - \hat{y}_{t,t}(\tau \le t),$$

$$e_{t+1} = y_{t+1} - \hat{y}_{t+1,t}(t \le t+1).$$

One can now perform the following simple regressions:

$$e_t = \alpha + \varepsilon_t, \tag{1}$$

$$e_{t+1} = \alpha + \beta e_t + \varepsilon_{t+1}. \tag{2}$$

For an efficient forecast, we must have $\alpha = 0$ (unbiasedness) in Equation (1), and and $\alpha = 0$, $\beta = 0$ in Equation (2), implying unbiasedness and absence of serial correlation. The first regression gives rise to a simple Student's $t$-test of $\alpha = 0$, whereas the second leads to an $F$-test. Adding the forecast $\hat{y}_{t+1,t}$ to both sides of Equation (2), this regression is easily seen to be equivalent to the conventional Mincer-Zarnowitz (1969) levels regression

$$y_{t+1} = \alpha + \beta \hat{y}_{t+1,t} + \varepsilon_{t+1}. \tag{3}$$

In this regression, unbiasedness of the forecast translates into a requirement that $\alpha = 0$, $\beta = 1$.

## Efficiency Properties More Generally

Unbiasedness and absence of serial correlation in the forecast errors can be thought of as weak efficiency requirements. A much more general and stricter orthogonality condition holds for optimal forecasts under MSE loss. Because an optimal forecast should be the conditional expectation of the predicted variable of interest, if the forecaster uses all available information efficiently, then no variable in the current information set should be able to predict future forecast errors. To test this, let $z_t$ be any such variable in the forecaster's information set at time $t$, $\Omega_t$. An implication of informational efficiency is that $\alpha = \beta = 0$ in the regression

$$e_{t+1} = \alpha + \beta z_t + \varepsilon_{t+1}, \tag{4}$$

where $\varepsilon_{t+1}$ is a serially uncorrelated, zero-mean-error term. The relationship between unbiasedness and absence of serial correlation on the one hand (equation and informational efficiency according to Equation (4), on the other, more generally is similar to the relationship between the weak and semistrong versions of the market efficiency hypothesis. According to the weakly efficient hypothesis, past values of the variable itself should not help predict future values. The semistrong version tightens this restriction by requiring that no publicly available information helps forecast future values.

## Tests on Forecast Revisions

Forecast revisions are of fundamental interest in a forecast evaluation exercise for one simple reason: If a sequence of forecasts is optimal, then the forecast revisions should themselves be unpredictable (technically a martingale difference sequence). Indeed, if this were not the case and, say, forecast revisions between February and April were themselves predictable, then the original (February) forecast would not be optimal. Suppose, for example, that it is known that on average the April forecast of next-year output growth tends to be $\frac{1}{4}$ of 1 percent higher than the February forecast. Then the February forecast should be revised upward by this amount to reflect the better information available in April of each year.

Another advantage of studying revisions is that predictable patterns in revisions, if detected, automatically tell the forecaster how to improve the original forecast, namely, by amending it by the fitted value of the forecast revision. Hence, if the February forecast of the revision in the forecast between February and April is

$$rev_{t+1,t}^{Apr} = \hat{\alpha} + \hat{\beta} z_t^{Feb},$$

the original February forecast, $\hat{y}_{t+1,t}^{Feb}$, can be replaced by an improved forecast, $\tilde{y}_{t+1,t}^{Feb}$, as follows:

$$\tilde{y}_{t+1,t}^{Feb} = \hat{y}_{t+1,t}^{Feb} + rev_{t+1,t}^{Apr}. \tag{5}$$

More generally, if $\Omega_t^{Apr}$ is the forecaster's information set in April, $\Omega_t^{Feb}$ is the information set in February (which is a subset of the April information

set, $\Omega_t^{Apr} \subseteq \Omega_t^{Feb}$) and if forecasts are formed optimally as conditional expectations—that is, $\hat{y}_{t+1,t}^{Feb} = E[y_{t+1}|\Omega_t^{Feb}]$ and $\hat{y}_{t+1,t}^{Apr} = E[y_{t+1}|\Omega_t^{Apr}]$—then by the law of iterated expectations $E[\hat{y}_{t+1,t}^{Apr}|\Omega_t^{Feb}] = \hat{y}_{t+1,t}^{Feb}$, and so the revision, defined as $rev_{t+1,t} = \hat{y}_{t+1,t}^{Apr} - \hat{y}_{t+1,t}^{Feb}$, must be zero-mean:

$$E[rev_{t+1,t}|\Omega_t^{Feb}] = 0. \tag{6}$$

A similar result holds for the current-year revisions, $rev_{t,t} = \hat{y}_{t,t}^{Apr} - \hat{y}_{t,t}^{Feb}$,

$$E[rev_{t,t}|\Omega_t^{Feb}] = 0. \tag{7}$$

Notice, however, that in general $E[rev_{t+1,t}|\Omega_t^{Apr}] \neq 0$ and $E[rev_{t,t}|\Omega_t^{Apr}] \neq 0$, provided that new information arrives between February and April of year $t$. It is worth emphasizing that we ignore estimation errors, which can induce serial correlation in the forecast errors even if the forecaster knows the true model. This is akin to learning effects—see Timmermann (1993) for a discussion of this point in the context of predictability of financial returns.

An important implication follows from these simple results: forecast optimality can be tested without having data on the target variable $y$. This is important because, given the availability of different vintages of the target variable, it is not clear whether the forecasts should be compared to the first-issue, second (revised), or "final" data revision. This matters considerably in practice as witnessed by the recent literature on real-time macroeconomic data (see Croushore, 2006). By analyzing data revisions, we can effectively construct a test that is not sensitive to how well the underlying data are being measured.

## Nonincreasing Variance of Forecast Errors as the Forecast Horizon is Decreased

A final property of an optimal forecast is declining variance of the forecast error as more information becomes available. This means that the February current-year (next-year) forecast errors should have a greater variance than the April current-year (next-year) forecast errors:

$$Var(e_{t+1,t}^{Apr}) \leq Var(e_{t+1,t}^{Feb}),$$
$$Var(e_{t,t}^{Apr}) \leq Var(e_{t,t}^{Feb}). \tag{8}$$

Intuitively this simply reflects that more information about the outcome in the current or next year is known in April than in February of the same year. This can be formally tested through a variance ratio test or (more appropriately given the small sample size here) by considering patterns in the variance of forecast errors associated with different forecast horizons.

## III. Empirical Results

With the data set and benchmark properties of an optimal forecast in place, we proceed to analyze the empirical evidence. Table 1 reports summary

### Table 1. Descriptive Statistics for Forecast Errors, by Variable and Region
*(Averages across countries in region)*

| | Mean | Median | Standard Deviation | Serial Correlation | Fraction of Positive Errors |
|---|---|---|---|---|---|
| | | | Real GDP (annual change in percent) | | |
| **April current-year forecast errors** | | | | | |
| Africa | −1.17 | −0.81 | 3.19 | 0.21 | 0.34 |
| Central and Eastern Europe | −1.17 | −0.71 | 3.49 | 0.37 | 0.46 |
| Commonwealth of Independent States (CIS) and Mongolia | −1.93 | −1.48 | 8.28 | 0.31 | 0.53 |
| Developing Asia | −0.38 | −0.33 | 2.22 | 0.25 | 0.49 |
| Middle East | −1.66 | 0.20 | 6.38 | 0.37 | 0.53 |
| Western Hemisphere | −0.64 | −0.61 | 2.41 | 0.23 | 0.39 |
| Advanced Economies | −0.04 | −0.14 | 1.36 | 0.21 | 0.48 |
| **September current-year forecast errors** | | | | | |
| Africa | −0.60 | −0.52 | 2.81 | 0.24 | 0.40 |
| Central and Eastern Europe | −0.11 | 0.11 | 2.37 | 0.33 | 0.56 |
| CIS and Mongolia | −1.05 | −0.56 | 6.35 | 0.27 | 0.61 |
| Developing Asia | 0.16 | 0.24 | 1.24 | 0.27 | 0.57 |
| Middle East | 0.67 | 0.22 | 3.67 | 0.35 | 0.58 |
| Western Hemisphere | −0.26 | −0.12 | 2.02 | 0.24 | 0.46 |
| Advanced Economies | 0.09 | −0.02 | 0.81 | 0.22 | 0.55 |
| **April next-year forecast errors** | | | | | |
| Africa | −1.45 | −1.37 | 4.07 | 0.28 | 0.33 |
| Central and Eastern Europe | −1.63 | −0.92 | 3.90 | 0.37 | 0.39 |
| CIS and Mongolia | −2.17 | −1.93 | 8.40 | 0.48 | 0.51 |
| Developing Asia | −0.63 | −0.67 | 2.86 | 0.33 | 0.45 |
| Middle East | −1.06 | 0.11 | 6.63 | 0.32 | 0.49 |
| Western Hemisphere | −1.33 | −1.34 | 3.08 | 0.26 | 0.33 |
| Advanced Economies | −0.55 | −0.74 | 2.06 | 0.29 | 0.42 |
| **September next-year forecast errors** | | | | | |
| Africa | −1.48 | −1.41 | 4.02 | 0.23 | 0.33 |
| Central and Eastern Europe | −1.40 | −0.97 | 3.76 | 0.34 | 0.41 |
| CIS and Mongolia | −2.39 | −2.78 | 9.60 | 0.46 | 0.52 |
| Developing Asia | −0.53 | −0.68 | 2.84 | 0.31 | 0.45 |
| Middle East | −1.34 | 0.06 | 6.15 | 0.31 | 0.53 |
| Western Hemisphere | −1.16 | −1.16 | 2.96 | 0.24 | 0.35 |
| Advanced Economies | −0.36 | −0.48 | 1.97 | 0.24 | 0.44 |
| **Current-year forecast revision** | | | | | |
| Africa | −0.83 | −0.54 | 2.00 | 0.23 | 0.36 |
| Central and Eastern Europe | −0.94 | −0.45 | 2.36 | 0.26 | 0.51 |
| CIS and Mongolia | −1.02 | −0.85 | 5.42 | 0.42 | 0.53 |
| Developing Asia | −0.46 | −0.51 | 1.86 | 0.29 | 0.45 |
| Middle East | −2.33 | −0.62 | 6.17 | 0.21 | 0.44 |
| Western Hemisphere | −0.34 | −0.27 | 1.56 | 0.23 | 0.38 |
| Advanced Economies | −0.11 | −0.07 | 0.99 | 0.26 | 0.49 |

| | Mean | Median | Standard Deviation | Serial Correlation | Fraction of Positive Errors |
|---|---|---|---|---|---|
| **Table 1** *(continued)* | | | | | |
| **Next-year forecast revision** | | | | | |
| Africa | −0.08 | −0.07 | 1.55 | 0.28 | 0.45 |
| Central and Eastern Europe | −0.29 | −0.24 | 0.99 | 0.31 | 0.45 |
| CIS and Mongolia | −0.37 | −0.36 | 2.19 | 0.24 | 0.47 |
| Developing Asia | −0.20 | −0.22 | 1.25 | 0.28 | 0.45 |
| Middle East | 1.10 | 0.27 | 4.69 | 0.46 | 0.47 |
| Western Hemisphere | −0.47 | −0.44 | 1.22 | 0.31 | 0.35 |
| Advanced Economies | −0.20 | −0.22 | 0.71 | 0.21 | 0.39 |
| | | Inflation (in percent per year) | | | |
| **April current-year forecast errors** | | | | | |
| Africa | 57.48 | 0.60 | 162.18 | 0.29 | 0.57 |
| Central and Eastern Europe | 7.36 | 2.30 | 24.37 | 0.39 | 0.53 |
| CIS and Mongolia | 340.63 | 126.95 | 978.65 | 0.69 | 0.49 |
| Developing Asia | 1.72 | 1.06 | 8.20 | 0.33 | 0.53 |
| Middle East | −2.03 | −0.86 | 8.49 | 0.28 | 0.34 |
| Western Hemisphere | 18.69 | 1.78 | 50.40 | 0.38 | 0.56 |
| Advanced economies | −0.08 | −0.03 | 0.94 | 0.20 | 0.44 |
| **September current-year forecast errors** | | | | | |
| Africa | 39.70 | 0.12 | 133.39 | 0.23 | 0.54 |
| Central and Eastern Europe | 2.03 | 0.17 | 8.79 | 0.27 | 0.47 |
| CIS and Mongolia | 163.16 | 64.00 | 554.43 | 0.60 | 0.42 |
| Developing Asia | 1.00 | 0.09 | 5.97 | 0.23 | 0.47 |
| Middle East | −0.59 | −0.41 | 8.61 | 0.19 | 0.43 |
| Western Hemisphere | 7.06 | 0.58 | 23.87 | 0.27 | 0.50 |
| Advanced economies | −0.09 | −0.05 | 0.49 | 0.23 | 0.42 |
| **April next-year forecast errors** | | | | | |
| Africa | 81.72 | 2.50 | 177.60 | 0.26 | 0.66 |
| Central and Eastern Europe | 16.05 | 4.03 | 34.06 | 0.41 | 0.60 |
| CIS and Mongolia | 229.71 | 177.67 | 592.49 | 0.83 | 0.70 |
| Developing Asia | 1.45 | 1.28 | 9.16 | 0.42 | 0.55 |
| Middle East | −0.77 | −0.96 | 11.32 | 0.38 | 0.38 |
| Western Hemisphere | 10.85 | 2.66 | 62.08 | 0.38 | 0.59 |
| Advanced economies | −0.12 | −0.13 | 1.43 | 0.36 | 0.43 |
| **September next-year forecast errors** | | | | | |
| Africa | 74.94 | 1.80 | 164.02 | 0.26 | 0.62 |
| Central and Eastern Europe | 16.04 | 3.27 | 32.43 | 0.39 | 0.58 |
| CIS and Mongolia | 190.99 | 153.65 | 590.62 | 0.74 | 0.65 |
| Developing Asia | 1.63 | 0.62 | 9.15 | 0.36 | 0.52 |
| Middle East | −1.77 | −1.15 | 10.96 | 0.27 | 0.32 |
| Western Hemisphere | 7.81 | 1.60 | 54.87 | 0.29 | 0.58 |
| Advanced economies | −0.19 | −0.13 | 1.20 | 0.33 | 0.40 |
| **Current-year forecast revision** | | | | | |
| Africa | 17.03 | 0.27 | 38.11 | 0.22 | 0.56 |
| Central and Eastern Europe | 5.00 | 1.78 | 19.15 | 0.20 | 0.58 |
| CIS and Mongolia | 182.11 | 120.45 | 551.08 | 0.61 | 0.56 |

| | Mean | Median | Standard Deviation | Serial Correlation | Fraction of Positive Errors |
|---|---|---|---|---|---|
| **Table 1** *(concluded)* | | | | | |
| Developing Asia | 0.31 | 0.15 | 3.21 | 0.29 | 0.52 |
| Middle East | 0.19 | −0.01 | 3.18 | 0.34 | 0.51 |
| Western Hemisphere | 9.67 | 0.29 | 29.17 | 0.18 | 0.51 |
| Advanced economies | 0.03 | 0.03 | 0.76 | 0.19 | 0.51 |
| **Next-year forecast revision** | | | | | |
| Africa | 3.92 | 0.30 | 11.01 | 0.21 | 0.57 |
| Central and Eastern Europe | 1.60 | 0.69 | 7.30 | 0.15 | 0.63 |
| CIS and Mongolia | 36.83 | 11.26 | 112.55 | 0.42 | 0.69 |
| Developing Asia | 0.22 | 0.22 | 2.78 | 0.19 | 0.56 |
| Middle East | 0.74 | 0.40 | 3.84 | 0.34 | 0.57 |
| Western Hemisphere | 3.20 | 0.07 | 12.27 | 0.18 | 0.57 |
| Advanced economies | 0.05 | 0.01 | 0.70 | 0.26 | 0.53 |

statistics for the forecast errors and forecast revisions grouped by variable and region. We show the mean, median, and standard deviation of the forecast error; the average absolute value of the coefficient of first-order serial correlation in the forecast errors; and the percentage of positive values of the forecast error. In all cases, these statistics are computed based on the cross section of countries within a particular region. For example, both the median and standard deviations are computed from the cross section of average values across countries in a given region. We next discuss the main empirical findings.

## GDP Growth

### Current-year forecasts

For the real GDP growth rate variable, the mean of the current-year forecast error (that is, the bias averaged across time and across countries) is very close to zero for the advanced economies. Biases in April current-year forecasts are much larger—exceeding more than 1 percent—and negative for Africa, Central and Eastern Europe, CIS and Mongolia, and the Middle East. As expected, this bias is reduced significantly in the September current-year forecasts. Although the April biases appear to be rather large, it should also be noted that they reflect some very large outliers whose values are predominantly negative and thus represent overpredictions. Indeed, the standard deviations of the April current-year forecast errors tend to be largest for those regions where the greatest biases were found—exceeding 8 percent for CIS and Mongolia and 6 percent for the Middle East.

Such outliers in the data lead us to consider more robust statistics as well, for example, the median forecast error and the proportion of positive forecast errors (underpredictions). Provided that the underlying shocks are not drawn from asymmetric distributions, one would expect the median to be close to

zero and the proportion of positive forecast errors to be close to 50 percent on average if the underlying forecasting model is not misspecified. Again the data reveal systematic problems for some of the regions: between 34 and 40 percent of the same-year forecasts for the African region are overpredictions of subsequent GDP growth (negative mean forecast errors). Consistent with this, the median forecast error remains large and negative ($-0.81$ for this region), as it does for Central and Eastern Europe and CIS and Mongolia.

Forecasts in all regions pass the test that the variance of the September forecast errors should be smaller than the variance of the April forecast errors of the same variable. Furthermore, in many regions the reduction in uncertainty between the April and September forecast appears to be quite large. For example, the average standard deviation of the current-year forecast error in the advanced economies is reduced from 1.36 percent in April to 0.81 percent in September, representing a 40 percent reduction.

### Next-Year Forecasts

Biases in the next-year forecast errors generally exceed those observed in the current-year forecasts. Interestingly, in every single region the mean April or September biases are negative, and this also holds for the median bias in all regions, with exception of the Middle East. This suggests that the WEO in general overpredicts next-year GDP growth. Furthermore, whereas the average bias in the current-year predictions for the advanced economies is very small, it is quite sizable in the next-year forecast, where it takes values of $-0.36$ and $-0.55$ percent, depending on the reporting date of the forecast. Estimates of the standard deviations of the forecast errors associated with the April and September next-year forecasts are much more similar than their current-year counterparts. This suggests that far less is learned between April and September about next-year growth than is learned between these months about growth in the current year.

The proportion of positive next-year forecast errors is again very low for Africa (0.33) and the Western Hemisphere (0.35). The predominance of regions with proportions of positive signs below 0.5 is consistent with the tendency of the WEO forecasts to overpredict next-year GDP growth.

Serial correlation in the forecast errors also appears to be a problem in some regions. The fourth column of Table 1, which reports the average of the absolute value of the first-order autocorrelation in the forecast error, is quite high in CIS and Mongolia in particular.

Turning to the forecast revisions between the April and September WEO publications, which should have a mean of zero, there is systematic evidence of negative biases. This is consistent with the April and September forecasts both overpredicting GDP growth on average, but the April forecast being more optimistic than the September value (so the mean change is negative). Hence, on average, the September forecast is being revised downward when compared with the April value. This finding is corroborated in the median values as well as in the proportion of positive forecast revisions (which

consistently lies below one-half) and is information that could easily be used to improve on the WEO growth forecasts.

Another feature worth noting in the forecast revisions is that the standard deviation of the revision is generally quite a bit larger for the current-year values than for the next-year values. Again, this reinforces the earlier observation that information arriving between April and September more strongly affects current-year than next-year forecasts.

### Inflation

Very high inflation rates characterized a number of countries during the sample period, so it is not surprising that outliers tend to be very large for this variable and certainly larger than for real GDP growth. As a consequence, we focus our analysis on the relatively robust measures of forecasting performance, such as the proportion of positive forecast errors. For the current-year forecasts, this does not deviate too strongly from 50 percent in any of the regions, except for the Middle East, where only between 34 and 43 percent of the April and September current-year forecast errors are positive, and to a lesser extent for the Advanced Economies, where 43 percent of the signs are positive.

A rather different picture emerges for the next-year forecast errors. Between 60 and 70 percent of the April forecast errors are positive for Africa, Central and Eastern Europe, and CIS and Mongolia. These proportions are closer to 60 percent for the September forecasts, but remain somewhat higher than 50 percent, indicating a tendency toward underprediction of inflation in these countries. Furthermore, all forecast revisions have positive means and more than 50 percent of the forecast revisions are positive. A particularly high percentage is observed among the next-year revisions for CIS and Mongolia and Central and Eastern Europe, which generally see the average forecast revised upward. Hence there is a tendency for both the WEO's current-year and next-year inflation forecasts to be raised between April and September. Since the September forecasts are generally more accurate than their April counterparts, this suggests that the April WEO inflation forecasts can be improved by increasing their value.

We also consider whether the standard deviation of the April forecast errors is greater than that of the September forecast errors. Although outliers make it difficult to interpret some of the values, this appears generally to be the case.

### IV. Analysis of Statistical Significance

Whether the biases documented in the previous table should be of concern depends on how systematic they are. This issue can best be addressed by undertaking a more in-depth statistical analysis. Such an analysis is of course tempered by the short data sample, which potentially invalidates inference relying on asymptotic distributions and also lowers the power of a statistical analysis to detect misspecification in the forecasting models, even when this is present. Again, countries with fewer than eight observations will be excluded from the statistical analysis. Considerable caution should be exercised when

interpreting the statistical inference results, because the sample size used here is very small, and finite-sample distortions of standard test statistics that correct for heteroskedasticity and autocorrelation in the regression residuals are well known (see Den Haan and Levin, 1997; Kiefer, Vogelsang, and Bunzel, 2000; and Kiefer and Vogelsang, 2002).

To deal with the problem that the small-sample properties of the simple $t$- and $F$-statistics are such that standard critical levels may not provide a reliable guide to inference, we designed a bootstrap experiment. This procedure repeatedly draws values of the forecast errors $(e_1, \ldots, e_T)$ with replacement from the empirical distribution function to construct a sample whose length $(T)$ is identical to that of the original data sample. Having constructed an artificial sample in this way $(e^b_{1(b)}, \ldots, e^b_{T(b)})$, where $b$ is an indicator for the $b$th bootstrap and $1(b)$, $T(b)$ are randomly drawn integer values between 1 and $T$, we recalculate the test statistics of interest, for example, $t$- and $F$-statistics associated with the efficiency regressions. We repeat this in 5,000 bootstrap experiments to construct a histogram for the distribution of the test statistic. The value of the test statistic found for the actual data is then compared with this bootstrapped distribution to get bootstrapped $p$-values. We shall report the proportion of countries for which the actual test statistic exceeds the 95th percentile of the bootstrapped distribution (using a two-sided test for the $t$-statistic).

Using Equation (1), the first two columns of Table 2 report the proportion of included countries in the various regions for which the $t$-statistic associated with the mean forecast error is less than $-2$ or greater than 2.[2] The third column reports the proportion of bootstrapped $p$-values for $\alpha = 0$ that fall below 0.05 using a two-sided test. It is instructive to compare the proportion of $t$-statistics that exceed 2 in absolute value against the bootstrapped $p$-values. In almost all cases the latter lead to far fewer rejections, indicating the small-sample size distortions that affect conventional test statistics.

The fourth column reports the percentage of regressions for which the absolute value of the $t$-statistic of $\beta$ in the weak efficiency regression, Equation (2), is greater than 2. The fifth column reports the percentage of cases where the $F$-test for the joint hypothesis $\alpha = 0$, $\beta = 0$ in Equation (2) exceeds its 5 percent critical level, and the final column reports the percentage of significant values of a sign test for whether the proportion of positive forecast errors differs from one-half, again using a 5 percent critical level. The purpose of reporting so many test statistics is to get a broader picture of possible forecast inefficiencies and to account for the fact that the individual test statistics are surrounded by more than the usual uncertainty, owing to the very small samples entertained here. Caution should therefore be exercised when interpreting the results.

---

[2]These $t$-statistics should be viewed only as broad indicators of statistical significance and are reported only because these measures are conventionally used to test statistical significance.

## Table 2. Tests for Biasedness and Serial Correlation of Forecast Errors
### (Share of countries in region with significant test statistics)

| | Forecast Error Bias ($\hat{\alpha}$) | | | Serial Correlation ($\hat{\beta}$) | | |
| --- | --- | --- | --- | --- | --- | --- |
| | *T*-value for $\hat{\alpha}$ | | Fraction of bootstrap *P*-value < 0.05 | $\left\|t_{\hat{\beta}}\right\| > 2$ | Fraction of bootstrap *P*-value < 0.05 | Fraction of significant sign tests (*P*-value < 0.05) |
| | $t_{\hat{\alpha}} < -2$ | $t_{\hat{\alpha}} > 2$ | | | | |
| *Real GDP* | | | | | | |
| **April current-year forecast errors** | | | | | | |
| Africa | 0.40 | 0.00 | 0.25 | 0.06 | 0.15 | 0.17 |
| Central and Eastern Europe | 0.13 | 0.00 | 0.07 | 0.13 | 0.13 | 0.20 |
| Commonwealth of Independent States (CIS) and Mongolia | 0.08 | 0.08 | 0.15 | 0.08 | 0.00 | 0.15 |
| Developing Asia | 0.10 | 0.05 | 0.05 | 0.00 | 0.00 | 0.15 |
| Middle East | 0.00 | 0.08 | 0.08 | 0.31 | 0.15 | 0.08 |
| Western Hemisphere | 0.19 | 0.00 | 0.13 | 0.03 | 0.06 | 0.03 |
| Advanced Economies | 0.03 | 0.03 | 0.07 | 0.00 | 0.03 | 0.10 |
| **September current-year forecast errors** | | | | | | |
| Africa | 0.20 | 0.00 | 0.15 | 0.12 | 0.15 | 0.15 |
| Central and Eastern Europe | 0.00 | 0.08 | 0.00 | 0.15 | 0.00 | 0.08 |
| CIS and Mongolia | 0.00 | 0.17 | 0.08 | 0.17 | 0.08 | 0.17 |
| Developing Asia | 0.00 | 0.09 | 0.00 | 0.00 | 0.00 | 0.18 |
| Middle East | 0.00 | 0.20 | 0.20 | 0.30 | 0.20 | 0.10 |
| Western Hemisphere | 0.10 | 0.05 | 0.05 | 0.05 | 0.00 | 0.05 |
| Advanced Economies | 0.00 | 0.10 | 0.07 | 0.07 | 0.10 | 0.17 |
| **April next-year forecast errors** | | | | | | |
| Africa | 0.38 | 0.04 | 0.33 | 0.15 | 0.33 | 0.35 |
| Central and Eastern Europe | 0.27 | 0.00 | 0.07 | 0.27 | 0.27 | 0.07 |
| CIS and Mongolia | 0.15 | 0.00 | 0.00 | 0.46 | 0.08 | 0.08 |
| Developing Asia | 0.17 | 0.04 | 0.09 | 0.30 | 0.13 | 0.09 |
| Middle East | 0.00 | 0.07 | 0.00 | 0.14 | 0.14 | 0.07 |
| Western Hemisphere | 0.36 | 0.00 | 0.27 | 0.09 | 0.15 | 0.18 |
| Advanced Economies | 0.28 | 0.03 | 0.24 | 0.14 | 0.21 | 0.14 |
| **September next-year forecast errors** | | | | | | |
| Africa | 0.40 | 0.02 | 0.33 | 0.13 | 0.29 | 0.33 |
| Central and Eastern Europe | 0.20 | 0.00 | 0.07 | 0.20 | 0.20 | 0.13 |
| CIS and Mongolia | 0.08 | 0.00 | 0.00 | 0.38 | 0.15 | 0.08 |
| Developing Asia | 0.04 | 0.04 | 0.04 | 0.17 | 0.13 | 0.09 |
| Middle East | 0.00 | 0.07 | 0.07 | 0.14 | 0.21 | 0.07 |
| Western Hemisphere | 0.33 | 0.00 | 0.24 | 0.06 | 0.21 | 0.09 |
| Advanced Economies | 0.24 | 0.07 | 0.21 | 0.00 | 0.14 | 0.10 |

| Table 2 *(continued)* | | | | | | |
|---|---|---|---|---|---|---|
| | Forecast Error Bias ($\hat{\alpha}$) | | | Serial Correlation ($\hat{\beta}$) | | |
| | *T*-value for $\hat{\alpha}$ | | Fraction of bootstrap *P*-value < 0.05 | $\left\| t_{\hat{\beta}} \right\| > 2$ | Fraction of bootstrap *P*-value < 0.05 | Fraction of significant sign tests (*P*-value < 0.05) |
| | $t_{\hat{\alpha}} < -2$ | $t_{\hat{\alpha}} > 2$ | | | | |
| **Current-year forecast revision** | | | | | | |
| Africa | 0.29 | 0.00 | 0.17 | 0.10 | 0.10 | 0.12 |
| Central and Eastern Europe | 0.18 | 0.00 | 0.00 | 0.00 | 0.00 | 0.09 |
| CIS and Mongolia | 0.00 | 0.00 | 0.00 | 0.20 | 0.00 | 0.00 |
| Developing Asia | 0.08 | 0.00 | 0.08 | 0.08 | 0.08 | 0.08 |
| Middle East | 0.00 | 0.00 | 0.00 | 0.10 | 0.00 | 0.00 |
| Western Hemisphere | 0.20 | 0.00 | 0.13 | 0.00 | 0.00 | 0.00 |
| Advanced Economies | 0.04 | 0.04 | 0.04 | 0.07 | 0.04 | 0.11 |
| **Next-year forecast revision** | | | | | | |
| Africa | 0.07 | 0.00 | 0.02 | 0.10 | 0.05 | 0.07 |
| Central and Eastern Europe | 0.22 | 0.00 | 0.00 | 0.22 | 0.11 | 0.00 |
| CIS and Mongolia | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 |
| Developing Asia | 0.08 | 0.00 | 0.08 | 0.00 | 0.00 | 0.00 |
| Middle East | 0.00 | 0.00 | 0.00 | 0.38 | 0.00 | 0.00 |
| Western Hemisphere | 0.20 | 0.00 | 0.13 | 0.07 | 0.13 | 0.07 |
| Advanced Economies | 0.19 | 0.00 | 0.07 | 0.07 | 0.07 | 0.11 |
| Inflation | | | | | | |
| **April current-year forecast errors** | | | | | | |
| Africa | 0.00 | 0.19 | 0.13 | 0.06 | 0.13 | 0.21 |
| Central and Eastern Europe | 0.00 | 0.07 | 0.00 | 0.33 | 0.07 | 0.13 |
| CIS and Mongolia | 0.00 | 0.08 | 0.00 | 0.69 | 0.31 | 0.00 |
| Developing Asia | 0.00 | 0.05 | 0.00 | 0.15 | 0.00 | 0.15 |
| Middle East | 0.08 | 0.00 | 0.00 | 0.15 | 0.00 | 0.15 |
| Western Hemisphere | 0.03 | 0.19 | 0.13 | 0.22 | 0.09 | 0.16 |
| Advanced Economies | 0.07 | 0.03 | 0.00 | 0.07 | 0.07 | 0.07 |
| **September current-year forecast errors** | | | | | | |
| Africa | 0.04 | 0.08 | 0.00 | 0.06 | 0.04 | 0.16 |
| Central and Eastern Europe | 0.07 | 0.00 | 0.00 | 0.20 | 0.00 | 0.00 |
| CIS and Mongolia | 0.00 | 0.08 | 0.00 | 0.62 | 0.15 | 0.00 |
| Developing Asia | 0.08 | 0.00 | 0.08 | 0.04 | 0.08 | 0.13 |
| Middle East | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.07 |
| Western Hemisphere | 0.03 | 0.06 | 0.03 | 0.12 | 0.06 | 0.06 |
| Advanced Economies | 0.17 | 0.00 | 0.03 | 0.07 | 0.03 | 0.00 |

## Table 2 *(concluded)*

| | Forecast Error Bias ($\hat{\alpha}$) | | | Serial Correlation ($\hat{\beta}$) | | |
|---|---|---|---|---|---|---|
| | *T*-value for $\hat{\alpha}$ | | Fraction of bootstrap *P*-value < 0.05 | | Fraction of bootstrap *P*-value < 0.05 | Fraction of significant sign tests (*P*-value < 0.05) |
| | $t_{\hat{\alpha}} < -2$ | $t_{\hat{\alpha}} > 2$ | | $\left| t_{\hat{\beta}} \right| > 2$ | | |
| **April next-year forecast errors** | | | | | | |
| Africa | 0.00 | 0.25 | 0.19 | 0.19 | 0.17 | 0.31 |
| Central and Eastern Europe | 0.00 | 0.20 | 0.13 | 0.33 | 0.20 | 0.33 |
| CIS and Mongolia | 0.00 | 0.31 | 0.00 | 0.92 | 0.31 | 0.31 |
| Developing Asia | 0.00 | 0.17 | 0.13 | 0.35 | 0.26 | 0.26 |
| Middle East | 0.29 | 0.14 | 0.14 | 0.29 | 0.21 | 0.36 |
| Western Hemisphere | 0.15 | 0.27 | 0.24 | 0.33 | 0.39 | 0.36 |
| Advanced Economies | 0.14 | 0.03 | 0.14 | 0.10 | 0.17 | 0.10 |
| **September next-year forecast errors** | | | | | | |
| Africa | 0.02 | 0.22 | 0.16 | 0.14 | 0.14 | 0.34 |
| Central and Eastern Europe | 0.00 | 0.27 | 0.20 | 0.33 | 0.20 | 0.40 |
| CIS and Mongolia | 0.00 | 0.23 | 0.08 | 0.77 | 0.15 | 0.23 |
| Developing Asia | 0.13 | 0.13 | 0.08 | 0.21 | 0.17 | 0.25 |
| Middle East | 0.21 | 0.14 | 0.14 | 0.00 | 0.14 | 0.21 |
| Western Hemisphere | 0.12 | 0.24 | 0.27 | 0.15 | 0.36 | 0.42 |
| Advanced Economies | 0.07 | 0.00 | 0.07 | 0.17 | 0.14 | 0.07 |
| **Current-year forecast revision** | | | | | | |
| Africa | 0.00 | 0.14 | 0.02 | 0.08 | 0.02 | 0.18 |
| Central and Eastern Europe | 0.00 | 0.07 | 0.00 | 0.20 | 0.00 | 0.07 |
| CIS and Mongolia | 0.00 | 0.08 | 0.00 | 0.62 | 0.31 | 0.15 |
| Developing Asia | 0.00 | 0.08 | 0.04 | 0.08 | 0.04 | 0.17 |
| Middle East | 0.00 | 0.00 | 0.00 | 0.29 | 0.00 | 0.21 |
| Western Hemisphere | 0.00 | 0.12 | 0.00 | 0.06 | 0.00 | 0.12 |
| Advanced Economies | 0.03 | 0.03 | 0.07 | 0.07 | 0.03 | 0.07 |
| **Next-year forecast revision** | | | | | | |
| Africa | 0.02 | 0.08 | 0.04 | 0.06 | 0.06 | 0.18 |
| Central and Eastern Europe | 0.00 | 0.13 | 0.07 | 0.00 | 0.00 | 0.20 |
| CIS and Mongolia | 0.00 | 0.00 | 0.00 | 0.38 | 0.08 | 0.38 |
| Developing Asia | 0.00 | 0.17 | 0.04 | 0.09 | 0.09 | 0.13 |
| Middle East | 0.00 | 0.00 | 0.00 | 0.21 | 0.07 | 0.07 |
| Western Hemisphere | 0.00 | 0.06 | 0.03 | 0.06 | 0.03 | 0.24 |
| Advanced Economies | 0.03 | 0.10 | 0.03 | 0.14 | 0.10 | 0.10 |

Source: Author's calculations.

## GDP Forecasts

First consider the April current-year forecasts. For close to 40 percent of the countries in the African region, the GDP growth forecasts were systematically too large.[3] The bootstrapped test statistics confirm a significant bias for a much larger proportion of African countries—close to 25 percent—than should be expected if the forecasts were genuinely unbiased. This proportion is reduced to 15 percent when bias and serial correlation are jointly tested, most likely because of the weaker power of the joint test, which requires estimation of an additional parameter. In fact, we can identify significant serial correlation for only about 6 percent of the African countries (column 4 of Table 2). Similarly, for about 15 percent of the African countries, the proportion of positive signs in the current-year forecast errors is significantly different from one-half at the 5 percent critical level (column 5 of Table 2).

Between 10 and 20 percent of the countries in CIS and Mongolia and the Western Hemisphere also show evidence of a significant bias in the forecasts. Serial correlation in the forecast errors appears to be most important in the Middle East, where 15 percent of the countries generate significant bootstrapped test statistics. These findings mostly carry over to the September current-year forecasts. Forecast errors continue to be biased and serially correlated for about 15 percent of the countries in Africa and there is strong evidence of serial correlation for the Middle East. In contrast, there is very little evidence that the current-year forecasts are biased or serially correlated in developing Asia or the advanced economies. Overall, the proportion of cases with a significant bias is lower in the September current-year forecasts compared with the April current-year forecasts.

Turning to the next-year forecast errors, there is evidence of a significant upward bias in the forecasts for about 35 percent of the countries in Africa and almost 25 percent of the countries in the Western Hemisphere (column 3 of Table 2). Significant biases also affect more than 20 percent of the countries among the advanced economies. Serial correlation in next-year forecast errors plagues all regions, particularly Africa. All told, the bootstrapped *p*-values show a pattern of biased or serially correlated next-year forecast errors in all regions.

Current-year forecast revisions are biased for Africa and the Western Hemisphere but there is little evidence of serial correlation. Next-year forecast revisions are biased and serially correlated for more than 10 percent of the countries in the Western Hemisphere, but otherwise the evidence against (weak) efficiency tends to be relatively mild.

## Inflation Forecasts

As mentioned previously, the inflation data are affected by numerous outliers, so we will not rely on standard test statistics and instead will move

---

[3]Because the forecast error is defined as realization minus prediction, $e = y - \hat{y}$, a negative mean forecast error shows that the prediction on average exceeds the realization and thus negative *t*-values represent overpredictions.

directly to consider the bootstrap results. These reveal mild evidence of inefficiency in the current-year inflation forecasts. There appears to be some positive bias (underprediction of inflation) in the case of Africa and the Western Hemisphere. By far the strongest evidence against efficiency is found in the next-year forecast errors, which reveal forecasts that are systematically downward-biased in most regions except for the Advanced Economies. However, forecast errors are serially correlated in the latter region so the null of no bias or serial correlation is rejected for about 15 percent of all countries (more than double the level expected under the null).

For the next-year forecasts, with the exception of CIS and Mongolia, a greater-than-expected proportion of countries in the various regions generates a significant test statistic associated with the bias. The strongest evidence against efficiency comes from the serial correlation tests in column 5 of Table 2, which show that p-values below 5 percent were generated for between 15 percent and 40 percent of the countries in the various regions. In particular, more than 30 percent of the countries in the Western Hemisphere show evidence of significant serial correlation in the forecast errors. With few exceptions, forecast revisions reveal little systematic evidence of biases or serial correlation.

## V. Can the WEO Forecast Errors Be Predicted?

The process whereby the WEO forecasts are generated puts considerable emphasis on integrating predictions across countries, regions, and variables in order to produce a coherent and internally consistent projection of current and future economic activity. One way to analyze whether the procedures that are currently in place have their intended effect is to test for informational efficiency using a range of indicators of global economic activity. Such tests build on the moment condition $E[e_{t+1}|\Omega_t] = 0$—where $\Omega_t$ is the forecaster's information set at the time of the forecast ($t$)—and are hence versions of the efficiency tests in Equation (4).

In our empirical application we focus on four such predictor variables. First, we consider the WEO prediction of U.S. GDP growth. This is an obvious choice given the size of the U.S. economy and the leading role it plays in shaping global economic activity. The second instrument is the WEO prediction of German output growth—again motivated by the significance of this economy to regional and global growth.[4] Finally, we also use the WEO forecast of oil prices and a global current account discrepancy instrument as

---

[4]For both U.S. and German growth, we use the April and September current-year and next-year WEO forecasts as instruments in predicting the corresponding April and September current-year and next-year forecast errors. These data are more up to date than the corresponding realized values (which are available only with a lag) and have the further advantage that they are the data used to forecast growth in other economies. Hence, if the predicted value of U.S. or German output growth helps explain forecast errors in other economies, it must be that the internal WEO projections were not fully utilized in producing a forecast for those other economies.

predictors. Oil prices are an obvious choice because they are an important determinant of economic growth and inflation in a number of economies. The global current account discrepancy is constructed as the sum total of current accounts across all countries scaled by 15 global exports. This figure should be equal to zero but may differ from this value owing to measurement errors.

Table 3 shows the outcome of this exercise. Within each region and for each of the predictor variables the table reports the proportion of $t$-values below $-2$ and above 2, respectively. Results indicative of a failure to fully account for the predicted U.S. GDP growth should show up in the form of a proportion of significant $t$-values somewhat higher than 5 percent. There is also information in the sign of the $t$-statistic. For GDP growth, a higher proportion of positive and significant values than negative and significant $t$-statistics would reveal a failure to fully account for the spillover of U.S. GDP growth to other countries.

There are only a few cases where the WEO prediction of U.S. GDP growth appears to be correlated with the forecast errors. However, the ones that we find are of considerable interest. Indeed, the evidence suggests that, for the advanced economies, 31 percent of the April current-year forecasts and 24 percent of the September current-year U.S. GDP forecasts generate a $t$-value above 2 and hence predict the forecast errors. This leads to a significantly positive $t$-statistic for 29 percent of the current-year forecast revisions in this region. In contrast, there is no evidence that the U.S. GDP forecast has predictive power over the next-year forecast errors. The only other instance registering a greater-than-expected proportion of significant $t$-values is the current-year forecasts for Central and Eastern Europe, where 33 percent of the $t$-values exceed a value of 2. For many of the countries in this region, the revision to the current-year forecast that takes place between April and September is predicted by the U.S. GDP forecast.

Turning to the WEO forecast of German output growth, interestingly this is positively correlated and significant in explaining forecast errors in a high proportion of countries in CIS and Mongolia (particularly for the next-year forecast errors) but not to nearly the same extent in other regions.

With a few interesting exceptions—namely, CIS and Mongolia, for which predicted oil prices are positively correlated with forecast errors in GDP growth, and Western Hemisphere and advanced economies, for which a negative correlation emerges—the WEO forecasts of oil prices do not appear to be overly important in explaining forecast errors in output growth.

Interestingly, the global current account discrepancy is significant for close to 40 percent and 25 percent of the advanced economies in explaining the April current-year and next-year forecast errors, respectively.

There is evidence that the next-year inflation forecast errors are linked to U.S. GDP forecasts, particularly for countries in Central and Eastern Europe, CIS and Mongolia, Developing Asia, Western Hemisphere, and the advanced economies. Once again the WEO forecast of German output growth is significant in explaining the inflation forecast error for a very large proportion of the countries in the CIS and Mongolia region.

**Table 3. Predictability of Forecast Errors in Relation to Current Information Variables**

*(Fraction of all countries in region with t-values for additional variables above or below indicated threshold)*

| | U.S. GDP Growth | | German GDP Growth | | Oil Prices | | Global Current Account Discrepancy | |
|---|---|---|---|---|---|---|---|---|
| | $t_{\hat{\beta}} < -2$ | $t_{\hat{\beta}} > 2$ | $t_{\hat{\beta}} < -2$ | $t_{\hat{\beta}} > 2$ | $t_{\hat{\beta}} < -2$ | $t_{\hat{\beta}} > 2$ | $t_{\hat{\beta}} < -2$ | $t_{\hat{\beta}} > 2$ |
| Real GDP | | | | | | | | |
| **April current-year forecast errors** | | | | | | | | |
| Africa | 0.02 | 0.04 | 0.00 | 0.06 | 0.04 | 0.04 | 0.06 | 0.06 |
| Central and Eastern Europe | 0.00 | 0.33 | 0.00 | 0.07 | 0.00 | 0.13 | 0.00 | 0.20 |
| Commonwealth of Independent States (CIS) and Mongolia | 0.00 | 0.00 | 0.08 | 0.23 | 0.08 | 0.08 | 0.23 | 0.00 |
| Developing Asia | 0.00 | 0.05 | 0.00 | 0.00 | 0.00 | 0.00 | 0.05 | 0.10 |
| Middle East | 0.08 | 0.08 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 |
| Western Hemisphere | 0.00 | 0.09 | 0.06 | 0.00 | 0.06 | 0.06 | 0.03 | 0.09 |
| Advanced economies | 0.00 | 0.31 | 0.07 | 0.03 | 0.07 | 0.00 | 0.00 | 0.38 |
| **September current-year forecast errors** | | | | | | | | |
| Africa | 0.00 | 0.05 | 0.05 | 0.07 | 0.00 | 0.12 | 0.07 | 0.02 |
| Central and Eastern Europe | 0.00 | 0.08 | 0.00 | 0.00 | 0.00 | 0.15 | 0.08 | 0.15 |
| CIS and Mongolia | 0.00 | 0.00 | 0.00 | 0.08 | 0.00 | 0.25 | 0.33 | 0.08 |
| Developing Asia | 0.00 | 0.09 | 0.18 | 0.00 | 0.00 | 0.00 | 0.09 | 0.00 |
| Middle East | 0.00 | 0.00 | 0.10 | 0.00 | 0.10 | 0.00 | 0.00 | 0.00 |
| Western Hemisphere | 0.00 | 0.00 | 0.05 | 0.00 | 0.00 | 0.00 | 0.05 | 0.00 |
| Advanced economies | 0.03 | 0.24 | 0.07 | 0.07 | 0.00 | 0.03 | 0.00 | 0.14 |
| **April next-year forecast errors** | | | | | | | | |
| Africa | 0.00 | 0.00 | 0.02 | 0.02 | 0.02 | 0.13 | 0.06 | 0.02 |
| Central and Eastern Europe | 0.00 | 0.00 | 0.07 | 0.00 | 0.13 | 0.07 | 0.00 | 0.07 |
| CIS and Mongolia | 0.00 | 0.08 | 0.08 | 0.54 | 0.00 | 0.08 | 0.00 | 0.00 |
| Developing Asia | 0.00 | 0.09 | 0.04 | 0.04 | 0.00 | 0.04 | 0.04 | 0.04 |

| | | | | | | | | |
|---|---|---|---|---|---|---|---|---|
| Middle East | 0.00 | 0.07 | 0.00 | 0.07 | 0.07 | 0.00 | 0.07 | 0.00 |
| Western Hemisphere | 0.03 | 0.06 | 0.09 | 0.06 | 0.06 | 0.00 | 0.00 | 0.09 |
| Advanced Economies | 0.21 | 0.03 | 0.03 | 0.00 | 0.21 | 0.00 | 0.00 | 0.24 |
| **September next-year forecast errors** | | | | | | | | |
| Africa | 0.02 | 0.00 | 0.02 | 0.04 | 0.02 | 0.06 | 0.06 | 0.06 |
| Central and Eastern Europe | 0.00 | 0.00 | 0.07 | 0.00 | 0.07 | 0.00 | 0.07 | 0.00 |
| CIS and Mongolia | 0.00 | 0.00 | 0.00 | 0.38 | 0.00 | 0.00 | 0.00 | 0.00 |
| Developing Asia | 0.09 | 0.04 | 0.04 | 0.04 | 0.00 | 0.00 | 0.00 | 0.04 |
| Middle East | 0.07 | 0.07 | 0.00 | 0.07 | 0.07 | 0.00 | 0.07 | 0.00 |
| Western Hemisphere | 0.06 | 0.00 | 0.06 | 0.15 | 0.15 | 0.00 | 0.03 | 0.12 |
| Advanced economies | 0.07 | 0.00 | 0.10 | 0.00 | 0.21 | 0.00 | 0.00 | 0.03 |
| **Current-year forecast revision** | | | | | | | | |
| Africa | 0.05 | 0.05 | 0.05 | 0.02 | 0.07 | 0.00 | 0.05 | 0.05 |
| Central and Eastern Europe | 0.00 | 0.36 | 0.00 | 0.09 | 0.00 | 0.00 | 0.00 | 0.00 |
| CIS and Mongolia | 0.00 | 0.00 | 0.10 | 0.50 | 0.00 | 0.20 | 0.00 | 0.00 |
| Developing Asia | 0.00 | 0.15 | 0.08 | 0.00 | 0.00 | 0.00 | 0.08 | 0.00 |
| Middle East | 0.10 | 0.10 | 0.00 | 0.00 | 0.00 | 0.00 | 0.20 | 0.00 |
| Western Hemisphere | 0.00 | 0.07 | 0.07 | 0.00 | 0.00 | 0.07 | 0.13 | 0.13 |
| Advanced economies | 0.00 | 0.29 | 0.04 | 0.04 | 0.04 | 0.00 | 0.00 | 0.32 |
| **Next-year forecast revision** | | | | | | | | |
| Africa | 0.07 | 0.02 | 0.02 | 0.07 | 0.02 | 0.02 | 0.02 | 0.05 |
| Central and Eastern Europe | 0.00 | 0.00 | 0.11 | 0.00 | 0.00 | 0.11 | 0.22 | 0.00 |
| CIS and Mongolia | 0.00 | 0.00 | 0.00 | 0.14 | 0.00 | 0.14 | 0.00 | 0.00 |
| Developing Asia | 0.00 | 0.00 | 0.00 | 0.08 | 0.00 | 0.00 | 0.17 | 0.00 |
| Middle East | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.50 | 0.00 |
| Western Hemisphere | 0.07 | 0.00 | 0.07 | 0.07 | 0.20 | 0.00 | 0.00 | 0.00 |
| Advanced economies | 0.04 | 0.00 | 0.00 | 0.04 | 0.04 | 0.00 | 0.00 | 0.11 |
| | | | | Inflation | | | | |
| **April current-year forecast errors** | | | | | | | | |
| Africa | 0.10 | 0.08 | 0.00 | 0.15 | 0.02 | 0.02 | 0.06 | 0.17 |
| Central and Eastern Europe | 0.00 | 0.20 | 0.07 | 0.40 | 0.00 | 0.00 | 0.13 | 0.00 |

| Table 3 (*concluded*) | | | | | | | | |
|---|---|---|---|---|---|---|---|---|
| CIS and Mongolia | 0.00 | 0.08 | 0.00 | 0.69 | 0.00 | 0.00 | 0.15 | 0.15 |
| Developing Asia | 0.00 | 0.05 | 0.05 | 0.30 | 0.00 | 0.05 | 0.00 | 0.00 |
| Middle East | 0.15 | 0.00 | 0.15 | 0.00 | 0.00 | 0.00 | 0.00 | 0.15 |
| Western Hemisphere | 0.03 | 0.25 | 0.00 | 0.22 | 0.06 | 0.03 | 0.03 | 0.03 |
| Advanced Economies | 0.00 | 0.14 | 0.07 | 0.03 | 0.00 | 0.07 | 0.21 | 0.03 |
| | | | | | | | | |
| **September current-year forecast errors** | | | | | | | | |
| Africa | 0.02 | 0.00 | 0.02 | 0.12 | 0.00 | 0.08 | 0.02 | 0.14 |
| Central and Eastern Europe | 0.00 | 0.07 | 0.07 | 0.27 | 0.07 | 0.07 | 0.07 | 0.13 |
| CIS and Mongolia | 0.00 | 0.00 | 0.08 | 0.62 | 0.00 | 0.00 | 0.54 | 0.08 |
| Developing Asia | 0.00 | 0.08 | 0.08 | 0.08 | 0.08 | 0.04 | 0.04 | 0.04 |
| Middle East | 0.07 | 0.07 | 0.21 | 0.00 | 0.07 | 0.00 | 0.00 | 0.07 |
| Western Hemisphere | 0.00 | 0.18 | 0.00 | 0.18 | 0.00 | 0.00 | 0.06 | 0.00 |
| Advanced economies | 0.03 | 0.03 | 0.00 | 0.10 | 0.00 | 0.14 | 0.07 | 0.00 |
| | | | | | | | | |
| **April next-year forecast errors** | | | | | | | | |
| Africa | 0.00 | 0.08 | 0.00 | 0.15 | 0.06 | 0.06 | 0.04 | 0.06 |
| Central and Eastern Europe | 0.00 | 0.33 | 0.00 | 0.33 | 0.00 | 0.00 | 0.00 | 0.00 |
| CIS and Mongolia | 0.00 | 0.23 | 0.00 | 0.92 | 0.00 | 0.00 | 0.00 | 0.00 |
| Developing Asia | 0.13 | 0.30 | 0.17 | 0.26 | 0.00 | 0.04 | 0.09 | 0.09 |
| Middle East | 0.14 | 0.14 | 0.07 | 0.00 | 0.00 | 0.00 | 0.00 | 0.07 |
| Western Hemisphere | 0.00 | 0.30 | 0.00 | 0.27 | 0.06 | 0.12 | 0.09 | 0.06 |
| Advanced economies | 0.07 | 0.17 | 0.24 | 0.14 | 0.03 | 0.00 | 0.00 | 0.00 |
| | | | | | | | | |
| | | | | | | | | |
| **September next-year forecast errors** | | | | | | | | |
| Africa | 0.00 | 0.12 | 0.02 | 0.12 | 0.06 | 0.06 | 0.02 | 0.06 |
| Central and Eastern Europe | 0.00 | 0.33 | 0.00 | 0.47 | 0.07 | 0.00 | 0.00 | 0.00 |

| | | | | | | | | |
|---|---|---|---|---|---|---|---|---|
| CIS and Mongolia | 0.00 | 0.23 | 0.00 | 0.77 | 0.00 | 0.00 | 0.08 | 0.08 |
| Developing Asia | 0.08 | 0.25 | 0.17 | 0.25 | 0.00 | 0.04 | 0.13 | 0.00 |
| Middle East | 0.07 | 0.07 | 0.14 | 0.00 | 0.00 | 0.00 | 0.00 | 0.07 |
| Western Hemisphere | 0.00 | 0.27 | 0.00 | 0.27 | 0.06 | 0.06 | 0.15 | 0.06 |
| Advanced economies | 0.07 | 0.17 | 0.21 | 0.10 | 0.00 | 0.03 | 0.00 | 0.00 |
| **Current-year forecast revision** | | | | | | | | |
| Africa | 0.04 | 0.08 | 0.00 | 0.08 | 0.04 | 0.08 | 0.04 | 0.02 |
| Central and Eastern Europe | 0.00 | 0.13 | 0.13 | 0.27 | 0.00 | 0.00 | 0.13 | 0.07 |
| CIS and Mongolia | 0.00 | 0.08 | 0.00 | 0.54 | 0.00 | 0.00 | 0.00 | 0.31 |
| Developing Asia | 0.00 | 0.08 | 0.00 | 0.04 | 0.13 | 0.08 | 0.00 | 0.04 |
| Middle East | 0.07 | 0.07 | 0.07 | 0.07 | 0.00 | 0.07 | 0.00 | 0.00 |
| Western Hemisphere | 0.00 | 0.06 | 0.00 | 0.06 | 0.03 | 0.09 | 0.00 | 0.03 |
| Advanced economies | 0.00 | 0.10 | 0.10 | 0.07 | 0.00 | 0.03 | 0.07 | 0.03 |
| **Next-year forecast revision** | | | | | | | | |
| Africa | 0.00 | 0.04 | 0.06 | 0.04 | 0.02 | 0.12 | 0.02 | 0.02 |
| Central and Eastern Europe | 0.00 | 0.13 | 0.00 | 0.00 | 0.00 | 0.00 | 0.07 | 0.00 |
| CIS and Mongolia | 0.00 | 0.15 | 0.00 | 0.23 | 0.00 | 0.00 | 0.00 | 0.23 |
| Developing Asia | 0.00 | 0.04 | 0.09 | 0.04 | 0.04 | 0.04 | 0.00 | 0.04 |
| Middle East | 0.00 | 0.00 | 0.07 | 0.14 | 0.00 | 0.21 | 0.07 | 0.00 |
| Western Hemisphere | 0.00 | 0.00 | 0.15 | 0.03 | 0.06 | 0.12 | 0.00 | 0.03 |
| Advanced economies | 0.00 | 0.07 | 0.00 | 0.10 | 0.00 | 0.03 | 0.00 | 0.00 |

Source: Author's calculations.

## Output Gap

The output gap—measured as the difference between actual and potential GDP—plays an important role in the WEO forecasts. Implicit in these is an assumption that the output gap is eliminated after 5 years. If this assumption is unrealistic and leads to biased forecasts, then one would expect that the predicted value of the output gap itself would be accountable for forecast errors. For example, if it takes longer to eliminate the output gap than assumed in the WEO, then the WEO will tend to overpredict forecasts for countries with large output gaps.

We have data on output gaps for the 29 advanced economies. For each of these, we regress the forecast error on an intercept and on the predicted output gap whose timing corresponds to the forecast with which it gets matched.

Table 4 presents the results in the form of $t$-statistics for current- and next-year forecast errors and forecast revisions. A pattern that stands out for the GDP forecasts is that the signs of the estimated $t$-values predominantly are negative. About 15 percent of the $t$-statistics exceed 2 in absolute value. The large negative $t$-statistics for Germany, France, and Italy are particularly interesting because, as we shall see subsequently, these were also economies for which the WEO output growth forecasts were systematically biased upward during the period. This finding suggests that the reduction in the output gap assumed in computing the WEO forecasts could lead to overpredictions: All three economies had large output gaps during the 1990s, as did Japan—the output gap averaged $-1.63$, $-1.99$, $-2.30$, and $-4.16$ for France, Germany, Italy, and Japan, respectively. These were among the highest output gaps in the 29 countries. An assumption in the WEO forecasts that these output gaps would be reduced too fast might lead to a greater prediction of output growth and hence to an upward bias in the forecast.

The sign of the regression coefficient of the output gap is predominantly positive in the case of the inflation forecast errors, that is, the opposite of the sign of what was found for the GDP forecasts. Hence, the larger the output gap—that is, the greater an economy's unused capacity—the more the WEO tends to underpredict inflation. This effect can be quite large and is borderline significant for countries such as France, Germany, and Korea.

Finally, turning to the regression results for the current account, there are many instances with large and significant predictability from the output gap over subsequent forecast errors, although the sign of the regression coefficient varies quite a bit. Countries for which a significant degree of predictability is found include Hong Kong SAR, Japan, the Netherlands, Singapore, and Sweden.

## VI. Revisions from Board to Published Forecasts

WEO forecasts are published twice a year, in April and September. Several rounds of forecast revisions precede the published version. A first set of

## Table 4. Output Gaps and the Predictability of Forecast Errors in Advanced Economies

*(Value of t-statistics for the coefficient of the output gap in forecast efficiency regression)*

| | Current-Year | | Next-Year | | Forecast Revisions | |
|---|---|---|---|---|---|---|
| | April | September | April | September | Current-year | Next-year |
| | | | Real GDP | | | |
| Australia | −1.64 | 0.91 | −0.03 | 0.91 | −2.69 | −1.91 |
| Austria | −1.56 | −1.41 | −1.09 | −1.34 | −1.10 | −0.24 |
| Belgium | −2.01 | −1.51 | −1.35 | −1.54 | −0.73 | 0.13 |
| Canada | −0.58 | −0.53 | 0.53 | −0.22 | −0.28 | −0.04 |
| Cyprus | 0.02 | 0.02 | 0.00 | 0.00 | −0.44 | 0.76 |
| Denmark | −1.67 | −0.25 | −1.20 | −0.85 | −3.02 | −2.11 |
| Finland | −0.17 | −0.04 | −0.30 | −0.09 | −0.42 | 0.25 |
| France | −2.32 | −1.77 | −1.73 | −2.42 | −1.53 | −0.87 |
| Germany | −2.59 | 0.11 | −4.44 | −2.19 | −2.43 | −1.34 |
| Greece | 1.79 | 0.62 | 0.17 | −0.29 | −0.25 | −0.20 |
| Hong Kong SAR | −1.75 | 0.01 | −2.30 | −4.08 | −1.23 | −0.99 |
| Iceland | −1.72 | −1.13 | −0.96 | −0.62 | −0.75 | 0.18 |
| Ireland | −0.31 | 0.84 | −0.36 | −0.93 | −0.69 | −0.68 |
| Israel | −1.46 | −0.98 | −1.43 | −1.57 | −0.42 | −0.26 |
| Italy | −1.97 | −4.05 | −3.14 | −2.30 | −1.38 | −1.32 |
| Japan | −0.31 | 1.40 | −2.63 | −0.62 | −0.17 | −0.47 |
| Korea | −1.19 | −1.74 | −1.98 | −2.32 | −0.37 | 0.69 |
| Luxembourg | 1.14 | 0.70 | 1.11 | 0.93 | −0.08 | 0.69 |
| Netherlands | −0.14 | −0.05 | −1.26 | −0.36 | −0.32 | 0.63 |
| New Zealand | 1.21 | −0.67 | −0.55 | −0.47 | 1.21 | 0.53 |
| Norway | −1.45 | −2.99 | 0.53 | 0.21 | −0.37 | 0.32 |
| Portugal | 0.61 | −0.24 | 0.68 | 0.11 | −0.42 | 0.96 |
| Singapore | −0.75 | −0.69 | −3.08 | −2.67 | −0.33 | 0.48 |
| Spain | −0.79 | −1.02 | 0.19 | 0.14 | −0.54 | −0.14 |
| Sweden | −2.03 | −0.38 | −1.33 | −2.63 | −1.65 | −1.73 |
| Switzerland | 0.54 | 2.38 | −0.69 | −0.02 | 0.28 | 0.54 |
| Taiwan Province of China | −0.73 | 0.55 | −0.83 | −1.26 | −0.85 | −0.42 |
| United Kingdom | −1.57 | −0.67 | −0.21 | −0.82 | −1.76 | −0.42 |
| United States | −0.32 | −1.02 | 0.62 | −0.07 | 0.95 | 1.21 |
| | | | Inflation | | | |
| Australia | −0.31 | −0.27 | 0.32 | −0.06 | −0.09 | 0.23 |
| Austria | −0.35 | −0.28 | 0.06 | 0.86 | −0.28 | −0.32 |
| Belgium | 1.13 | 1.84 | 1.34 | 0.71 | 0.93 | 0.25 |
| Canada | 0.68 | 0.74 | 0.99 | 2.34 | 1.22 | 0.82 |
| Cyprus | −0.90 | −4.18 | 0.00 | 0.00 | −0.71 | −0.21 |
| Denmark | 0.29 | −1.08 | 0.59 | −0.39 | 0.42 | 0.00 |
| Finland | 2.26 | 2.17 | 0.77 | 1.81 | 1.60 | −0.37 |
| France | 1.80 | 1.34 | 1.69 | 2.91 | 1.14 | 0.44 |
| Germany | 1.07 | 1.71 | 0.65 | 1.42 | 1.11 | −0.12 |
| Greece | 0.07 | 0.08 | 1.42 | 0.42 | −0.16 | −0.30 |
| Hong Kong SAR | −0.07 | 0.01 | −0.93 | −1.00 | 0.07 | −0.81 |
| Iceland | 0.72 | 1.60 | 1.55 | 0.97 | 1.21 | 0.51 |
| Ireland | 1.23 | 0.67 | 2.31 | 1.48 | 1.52 | 1.38 |

| Table 4 (concluded) | | | | | | |
| --- | --- | --- | --- | --- | --- | --- |
| | Current-Year | | Next-Year | | Forecast Revisions | |
| | April | September | April | September | Current-Year | Next-Year |
| Israel | 1.77 | 0.41 | 1.53 | 0.40 | 1.63 | −1.23 |
| Italy | 0.11 | −1.31 | −1.29 | −0.52 | 0.59 | −0.99 |
| Japan | 1.54 | −1.09 | −0.13 | −0.52 | 2.64 | −0.06 |
| Korea | 1.58 | 3.15 | 1.33 | 2.08 | 2.07 | −0.78 |
| Luxembourg | −0.19 | 0.26 | −0.28 | 0.15 | −0.06 | −0.55 |
| Netherlands | 1.86 | 1.21 | 3.16 | 3.39 | 1.01 | −0.27 |
| New Zealand | 1.14 | 0.22 | 0.53 | 0.19 | 0.81 | −0.54 |
| Norway | 0.05 | −0.64 | −1.18 | −2.09 | 0.93 | 0.51 |
| Portugal | 1.89 | 1.24 | 2.16 | 1.11 | 0.57 | 0.85 |
| Singapore | −0.37 | −0.74 | −0.57 | 0.08 | −0.20 | 1.30 |
| Spain | −0.12 | −0.45 | −0.30 | −0.18 | 0.70 | −0.12 |
| Sweden | 0.64 | −0.25 | 1.30 | 0.34 | 0.47 | 0.54 |
| Switzerland | 1.19 | −0.57 | 0.65 | 1.78 | 1.40 | −0.56 |
| Taiwan Province of China | −2.19 | −1.45 | −0.92 | −0.53 | −1.47 | −0.09 |
| United Kingdom | 1.24 | 0.15 | −0.40 | −0.21 | 2.27 | −0.09 |
| United States | −0.12 | −0.18 | 0.99 | 0.11 | −0.26 | 0.99 |

Source: Author's calculations.

predictions is presented to the IMF board in February and July each year, preceding the April and September WEO publications. To assess the informational value of forecast revisions that occur between the Board version and the published version, we obtained data on Board forecasts of current-year GDP growth in February and next-year board forecasts of GDP growth reported in July. We refer to these forecasts as $\hat{y}_{t,t}^{Feb}$ and $\hat{y}_{t+1,t}^{July}$, respectively. Further, let the forecast revisions from the board to the published WEO forecasts be given by $rev_{t,t}^{pub-Board}$ and $rev_{t+1,t}^{pub-Board}$. If the revisions occurring between the board and published forecasts contain useful information, we should expect that they help predict the errors in the original board forecasts, defined as $e_{t,t}^{Board} = y_t - \hat{y}_{t,t}^{Feb}$ and $e_{t+1,t}^{Board} = y_{t+1} - \hat{y}_{t+1,t}^{July}$. We test this proposition through the regressions

$$
e_{t,t}^{Board} = \alpha + \beta rev_{t,t}^{pub-Board} + \varepsilon_t,
$$
$$
e_{t+1,t}^{Board} = \alpha + \beta rev_{t+1,t}^{pub-Board} + \varepsilon_{t+1}.
$$
(9)

If the revisions incorporated in the published WEO forecasts do not add any value to the original board forecast, then we should expect to find β-coefficients near zero. Conversely, we would expect to find significant and positive values of β and nonzero $R^2$-values in case the revisions contain valuable information. Estimation results based on Equation (9) are reported in Table 5. The current-year forecast errors for the advanced economies reveal strong evidence that the board-to-publication revision contains

Table 5. Real GDP: Significance of Forecast Revisions After Executive Board Meeting

*(Average across regions except for fractions)*

| | Fractions of | | | | Mean Squared Error Ratio |
| | *T*-values | | $\widehat{\beta}$ | $R^2$ | |
| | $t_{\hat{\beta}} < -1$ | $t_{\hat{\beta}} > 2$ | Coefficients $> 0$ | | |
|---|---|---|---|---|---|
| **April current-year forecast errors** | | | | | |
| Africa | 0.00 | 0.13 | 0.81 | 0.11 | 0.83 |
| Central and Eastern Europe | 0.00 | 0.07 | 0.60 | 0.09 | 0.85 |
| Commonwealth of Independent States (CIS) and Mongolia | 0.00 | 0.15 | 0.54 | 0.07 | 0.48 |
| Developing Asia | 0.00 | 0.15 | 0.50 | 0.16 | 0.72 |
| Middle East | 0.00 | 0.00 | 0.62 | 0.03 | 0.77 |
| Western Hemisphere | 0.00 | 0.13 | 0.53 | 0.08 | 0.64 |
| Advanced economies | 0.00 | 0.52 | 0.90 | 0.23 | 0.81 |
| | | | | | |
| **September current-year forecast errors** | | | | | |
| Africa | 0.00 | 0.07 | 0.41 | 0.07 | 0.55 |
| Central and Eastern Europe | 0.00 | 0.23 | 0.69 | 0.14 | 0.69 |
| CIS and Mongolia | 0.00 | 0.33 | 0.67 | 0.18 | 0.67 |
| Developing Asia | 0.00 | 0.27 | 0.55 | 0.15 | 0.43 |
| Middle East | 0.00 | 0.60 | 0.80 | 0.36 | 0.75 |
| Western Hemisphere | 0.00 | 0.38 | 0.67 | 0.24 | 0.74 |
| Advanced economies | 0.00 | 0.41 | 0.83 | 0.27 | 0.72 |
| | | | | | |
| **April next-year forecast errors** | | | | | |
| Africa | 0.00 | 0.10 | 0.63 | 0.08 | 0.92 |
| Central and Eastern Europe | 0.13 | 0.00 | 0.27 | 0.14 | 0.77 |
| CIS and Mongolia | 0.00 | 0.00 | 0.38 | 0.04 | 0.36 |
| Developing Asia | 0.00 | 0.04 | 0.48 | 0.07 | 0.63 |
| Middle East | 0.07 | 0.00 | 0.43 | 0.07 | 1.64 |
| Western Hemisphere | 0.03 | 0.12 | 0.52 | 0.08 | 0.66 |
| Advanced economies | 0.03 | 0.14 | 0.62 | 0.12 | 0.95 |
| | | | | | |
| **September next-year forecast errors** | | | | | |
| Africa | 0.04 | 0.04 | 0.38 | 0.06 | 0.81 |
| Central and Eastern Europe | 0.00 | 0.07 | 0.60 | 0.08 | 0.63 |
| CIS and Mongolia | 0.00 | 0.00 | 0.38 | 0.01 | 0.55 |
| Developing Asia | 0.04 | 0.17 | 0.48 | 0.15 | 0.74 |
| Middle East | 0.00 | 0.14 | 0.57 | 0.09 | 0.73 |
| Western Hemisphere | 0.03 | 0.06 | 0.45 | 0.09 | 0.80 |
| Advanced economies | 0.00 | 0.07 | 0.59 | 0.07 | 0.90 |

Source: Author's calculations.

valuable information that not only is significantly correlated with the forecast error for about 50 percent of the countries but has the required positive sign for between 80 and 90 percent of the countries. The large $R^2$-value of about 0.25 is further testimony to this effect and suggests that 25 percent of the current-year February or July forecast error can be explained by the revision between the board and published versions.

Much lower levels of significance are obtained for the next-year forecasts, for which, in the case of the advanced economies, close to 60 percent of the β-estimates are positive and only 14 and 7 percent (for April and September forecast respectively) of the coefficients exceed 2. Furthermore, the average $R^2$-value now declines to a level near 0.10.

The $R^2$ values do not in themselves quantify the degree of improvement in the WEO forecast from the board to the published version. A better measure of this is the ratio of MSE-values based on $\sum_{t=1}^{T}\left(e_{t,t}^{Apr}\right)^2 \Big/ \sum_{t=1}^{T}\left(e_{t,t}^{Board}\right)^2$. $\sum_{t=1}^{T}\left(e_{t,t}^{Board}\right)^2$ and $\sum_{t=1}^{T}\left(e_{t+1,t}^{Sep}\right)^2 \Big/ \sum_{t=1}^{T}\left(e_{t+1,t}^{Board}\right)^2$, where $T$ is the sample size. The final column in Table 5 shows these ratios. Values below unity indicate that the WEO forecast gets more precise from the board to the published version and the extent to which the ratio is below unity is a measure of the improvement. For the February/April same-year forecasts, the average ratio is about 0.80 for the advanced economies. This declines to a value near 0.70 for the July/September same-year forecasts, but is closer to 0.95 for the next-year forecasts. These values suggest that much valuable information is learned about current-year economic growth between the time the board forecast is reported and the time of the official publication. Far less information is learned about next-year economic growth between these dates, as witnessed by the $R^2$-values near 0.95.

Turning to the countries outside the advanced economies region, in general the percentage of positive coefficients in the board-to-publication revision regressions in Equation (9) is somewhat lower, as is the fraction of estimates that is statistically significant. In fact, only about 10–15 percent and 5–10 percent of the current-year and next-year coefficient estimates generate positive $t$-values that exceed 2. Interestingly, the MSE ratio tends to be somewhat lower than was found for the advanced economies, especially for the next-year forecasts, suggesting a significant improvement in the next-year forecasts between the board and the published forecasts for the other regions.

## VII. Comparison of WEO and Consensus Forecasts

A comparison of forecasts to subsequent outcomes—which we have done thus far—is an important exercise, or reality check, that allows us to test whether basic efficiency properties are satisfied by the forecasts. This exercise clearly has its limitations, however. For example, it is not evident what constitutes a good forecast in absolute terms. Some series may be intrinsically very difficult to predict (inflation comes to mind) because they are affected by

large exogenous shocks and/or shifts in economic policy whose effects are difficult to predict in advance. Conversely, a forecast can be very uninformative but lead to errors that do not appear to violate efficiency properties, such as unbiasedness and absence of serial correlation.

To address this issue, as Juhn and Loungani (2002) emphasize, it is very informative to compare the WEO forecasts to alternative forecasts such as those produced by a highly reputed source, such as the consensus forecasts. Forecasters included in the consensus survey faced similar difficulties as the WEO forecasters—for example, the higher-than-expected productivity growth for the U.S. economy or the absence of large, global inflationary shocks during most of the 1990s—and therefore serve as a yardstick against which the WEO forecasts can be measured.

## Consensus Data

To investigate the relative performance of the WEO and consensus forecasts, we obtained consensus forecast data on GDP growth, inflation, and the current account balance over the period 1990–2003. The data cover all the G-7 economies, seven Latin American economies (Argentina, Brazil, Chile, Colombia, Mexico, Peru, and Venezuela), and nine Asian economies (China, Hong Kong SAR, India, Indonesia, Korea, Malaysia, Singapore, Taiwan Province of China, and Thailand).

In the baseline scenario, consensus forecasts are measured in March (in the case of the current-year forecasts) and September (in the case of next-year forecasts) except for the Latin American economies, for which data coverage is limited for these months. For this reason, the February and August consensus forecasts were used for these economies. The consensus forecast is computed as the mean forecast across participants in a given monthly survey.

We shall refer to the March current-year consensus forecast as $\hat{y}_{t,t}^{cons}$; the September next-year consensus forecast is denoted as $\hat{y}_{t+1,t}^{cons}$. Although consensus forecasts are now available on a monthly basis, the March and September consensus forecasts are the forecasts that are based on information whose timing is most similar to the WEO April current-year (denoted $\hat{y}_{t,t}^{WEO}$) and September next-year ($\hat{y}_{t+1,t}^{WEO}$) forecasts, so this comparison was deemed most appropriate for measuring the information content of the two sets of forecasts. For completeness we shall later report the outcome of a sensitivity analysis that changes the timing of the consensus forecasts for each G-7 country.

## Statistical Tests of Forecasting Performance

To evaluate the relative performance of the two sets of forecasts, the left panel of Table 6 shows the ratio of consensus over WEO root-mean-squared forecast errors (RMSFE). Values lower than 1 suggest that the consensus forecast performed best over the sample, whereas values greater than 1 suggest that the WEO forecasts were better.

Allan Timmermann

## Table 6. Comparison of WEO and Consensus Forecasts: Ratios of Root-Mean-Squared Forecast Errors[1]
### (Consensus over World Economic Outlook)

| | Consensus Measured in March/September | | | | Consensus Measured in February/August | | | | Consensus Measured in April/October | | | |
| | GDP | | Inflation | | GDP | | Inflation | | GDP | | Inflation | |
| | Current-Year | Next-Year | Current-Year | Next-Year | Current-Year | Next-Year | Current-Year | Next-Year | Current-Year | Next-Year | Current-Year | Next-Year |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| **Group of seven countries** | | | | | | | | | | | | |
| Canada | 0.992 | 0.977 | 0.945 | 1.267 | 1.167 | 0.965 | 1.273 | 1.342 | 0.909 | 0.967 | 0.864 | 1.218 |
| France | 1.150 | 0.948 | 1.088 | 1.074 | 1.349 | 1.024 | 1.274 | 1.103 | 0.986 | 0.849 | 1.003 | 0.973 |
| Germany | 1.092 | 0.906 | 1.155 | 0.986 | 1.202 | 0.957 | 1.321 | 1.013 | 1.071 | 0.796 | 1.060 | 0.917 |
| Italy | 0.944 | 1.014 | 0.995 | 0.783 | 0.987 | 1.066 | 1.202 | 0.891 | 0.899 | 0.832 | 0.854 | 0.891 |
| Japan | 1.053 | 0.926 | 1.081 | 0.785 | 1.128 | 0.973 | 1.097 | 0.858 | 0.972 | 0.873 | 0.886 | 0.656 |
| United Kingdom | 1.184 | 1.024 | 1.194 | 0.924 | 1.291 | 1.069 | 1.419 | 1.018 | 1.064 | 0.940 | 0.989 | 0.996 |
| United States | 1.026 | 0.965 | 0.962 | 1.017 | 1.158 | 1.031 | 1.056 | 1.045 | 0.937 | 0.958 | 0.789 | 0.922 |
| **Latin America** | | | | | | | | | | | | |
| Argentina | 1.036 | 0.998 | 3.333 | 0.909 | 1.036 | 0.998 | 3.333 | 0.909 | 0.838 | 0.910 | 4.198 | 0.795 |
| Brazil | 1.121 | 1.085 | 1.599 | 0.754 | 1.121 | 1.085 | 1.599 | 0.754 | 1.001 | 0.937 | 0.762 | 0.689 |
| Chile | 1.098 | 1.014 | 1.289 | 0.787 | 1.098 | 1.014 | 1.289 | 0.787 | 1.061 | 0.835 | 1.280 | 0.743 |
| Colombia | 1.127 | 1.186 | 1.482 | 0.912 | 1.127 | 1.186 | 1.482 | 0.912 | 1.063 | 1.013 | 1.280 | 0.867 |
| Mexico | 1.107 | 0.988 | 1.207 | 0.995 | 1.107 | 0.988 | 1.207 | 0.995 | 0.888 | 0.969 | 1.333 | 0.999 |
| Peru | 1.272 | 0.990 | 1.366 | 1.922 | 1.272 | 0.990 | 1.366 | 1.922 | 1.218 | 0.893 | 1.345 | 1.656 |
| Venezuela | 1.100 | 0.963 | 1.617 | 0.928 | 1.100 | 0.963 | 1.617 | 0.928 | 0.951 | 0.885 | 1.153 | 0.953 |
| **Asia** | | | | | | | | | | | | |
| China | 0.915 | 0.853 | 1.522 | 1.246 | 0.916 | 0.889 | 1.439 | 1.138 | 0.908 | 0.863 | 1.618 | 1.299 |
| Hong Kong SAR | 0.988 | 1.096 | 0.910 | 0.820 | 1.089 | 1.061 | 0.802 | 0.783 | 0.970 | 1.061 | 0.977 | 0.914 |
| India | 1.153 | 1.066 | 1.378 | 0.935 | 1.087 | 1.036 | 1.045 | 1.031 | 1.071 | 1.049 | 1.282 | 0.974 |
| Indonesia | 0.951 | 1.011 | 1.610 | 1.001 | 1.109 | 1.068 | 0.855 | 0.977 | 0.852 | 0.961 | 1.997 | 1.010 |
| Korea | 0.958 | 0.963 | 1.106 | 0.786 | 1.027 | 0.965 | 0.884 | 0.855 | 0.876 | 0.978 | 0.979 | 0.950 |
| Malaysia | 0.941 | 1.005 | 0.767 | 0.716 | 1.040 | 1.062 | 0.918 | 0.715 | 0.883 | 0.967 | 0.757 | 0.683 |
| Singapore | 0.982 | 1.042 | 1.071 | 1.042 | 1.026 | 1.033 | 0.973 | 1.007 | 0.930 | 1.034 | 1.115 | 1.052 |
| Taiwan Province of China | 1.105 | 0.979 | 0.943 | 1.005 | 1.109 | 0.969 | 0.863 | 0.977 | 1.067 | 0.982 | 0.900 | 1.082 |
| Thailand | 0.954 | 0.902 | 0.932 | 1.110 | 1.034 | 0.927 | 0.900 | 0.998 | 0.896 | 0.859 | 0.922 | 1.198 |

Source: Author's calculations.

[1]Values greater than one indicate a better performance by the WEO forecasts, while values less than one indicate better performance by the consensus forecasts.

Current-year GDP forecasts produced by the WEO are on average better than the consensus forecasts for the G-7 and Latin American economies, because the RMSFE ratio exceeds unity for five of seven G-7 countries and for all Latin American countries—although it should be borne in mind that the current- and next-year forecasts for the latter countries are measured in February and August, respectively. In contrast, the consensus forecasts are better for the Asian economies, because only two of nine current-year RMSFE ratios exceed unity. Turning to the next-year GDP values, the performance of the two sets of forecasts is very similar, with RMSFE ratios between 0.90 and 1.10 in all but two cases. One notable exception is China, for which the WEO forecast is notably worse than the consensus forecast, as witnessed by the RMSFE ratio of 0.85.

The WEO current-year inflation forecasts perform quite well relative to the consensus values in all three regions, particularly in Latin America. Conversely, next-year inflation forecasts produced by the consensus survey are generally better than the WEO next-year forecasts for Latin America (with the exception of Peru). The two sets of forecasts are of similar quality for the G-7 and Asian economies.

Unsurprisingly, given the small samples, few cases produce significant test statistics when they are evaluated using a Diebold-Mariano (1995) test. It is interesting to note, however, that the WEO performs better than the consensus in a greater number of cases than vice versa, particularly when it comes to current-year forecasts. The WEO current-year GDP forecast is best for the United Kingdom, Colombia, Mexico, and India, whereas the WEO next-year forecast is more accurate for Thailand. Current-year WEO inflation forecasts for Argentina, Venezuela, and China surpass their consensus counterparts, but the opposite holds true for next-year inflation forecasts for Italy, Japan, Brazil, Malaysia, and Korea.

## Timing of Consensus Forecasts

The information sets underlying the consensus and WEO forecasts are not perfectly aligned, so it is worthwhile to investigate the sensitivity of the (relative) performance of the two sets of forecasts to changes in the dating. We do so in two ways. First, we compare the published (April/September) WEO forecasts to the consensus forecasts reported in February and August, respectively. This timing clearly benefits the WEO forecasts, which can embody more up-to-date information than is available in February or August. We also reverse the informational advantage by comparing the WEO forecasts to the April/October consensus forecasts, which embody more recent information than the WEO forecasts.

If the consensus and WEO forecasters update their predictions reasonably efficiently, we would expect that the consensus/WEO RMSFE ratios should be higher than when the consensus forecasts are based on the February/August information. Conversely, we would expect to see lower values when using the April/October consensus forecasts.

Table 6 also presents RMSFE ratios when the current- and next-year consensus forecasts are the ones published in February/August or in April/October. The February consensus GDP current-year forecasts generate higher RMSFE values than the WEO forecasts for six, seven, and eight of the G-7, Latin American, and Asian economies, respectively. On the other hand, the WEO next-year GDP forecasts surpass the consensus forecasts only in roughly half of the cases, despite the latter's use of outdated information relative to the WEO forecasts.

The WEO current-year inflation forecasts are most precise when measured against the February consensus forecasts for the G-7 countries and Latin America. Surprisingly, however, for seven out of nine Asian economies, the current-year February consensus inflation forecasts are better than the WEO forecasts despite their informational disadvantage. Furthermore, the next-year WEO inflation forecasts do not measure up well against the February consensus forecasts. Finally, the WEO forecasts of the current account generally perform well compared with the February consensus forecasts of this variable.

Turning to the performance of the April/October consensus forecasts compared with the WEO forecasts, it is clear that the consensus forecasts excel in the majority of cases, the only exception being the current-year forecasts of inflation in Latin America.

## VIII. Conclusion

This paper has undertaken a wide-ranging set of tests to assess several issues in relation to the performance of the WEO forecasts since 1990. In particular, it has addressed (1) how precise the WEO forecasts were when measured against actual outcomes; (2) whether there were simple ways to improve on these forecasts—in particular, whether spillover effects from major economies, such as the United States and Germany, are accounted for in all forecasts; and (3) how well the WEO forecasts performed relative to the consensus forecasts.

## REFERENCES

Artis, Michael J., 1988, ''How Accurate Is the World Economic Outlook? A Post Mortem on Short-Term Forecasting at the International Monetary Fund,'' *Staff Studies for the World Economic Outlook* (Washington, International Monetary Fund), pp. 1–49.

———, 1997, ''How Accurate Are the WEO's Short-Term Forecasts? An Examination of the World Economic Outlook,'' *Staff Studies for the World Economic Outlook* (Washington, International Monetary Fund).

Barrionuevo, J.M., 1993, ''How Accurate Are the World Economic Outlook Projections? *Staff Studies for the World Economic Outlook* ( Washington, International Monetary Fund).

Croushore, Dean, 2006, "Forecasting with Real-Time Macroeconomic Data," in *Handbook of Economic Forecasting*, ed. by Graham Elliott, Clive W.J. Granger, and Allan Timmermann (Amsterdam, North-Holland), pp. 961–982.

Den Haan, Wouter J., and Andrew T. Levin, 1997, "A Practitioner's Guide to Robust Covariance Matrix Estimation," in *Robust Inference*, Handbook of Statistics, Vol. 15, ed. by G.S. Maddala and C.R. Rao (Amsterdam, North-Holland), Chapter 12, pp. 299–342.

Diebold, Francis X., and Roberto Mariano, 1995, "Comparing Predictive Accuracy," *Journal of Business and Economic Statistics*, Vol. 13 (July), pp. 253–263.

Elliott, Graham, Ivana Komunjer, and Allan Timmermann, 2005, "Estimation and Testing of Forecast Rationality under Flexible Loss," *Review of Economic Studies*, Vol. 72 (October), pp. 1107–1125.

Juhn, Grace, and Prakash Loungani, 2002, "Further Cross-Country Evidence on the Accuracy of the Private Sector's Output Forecasts," *IMF Staff Papers*, Vol. 49 (April), pp. 49–64.

———, 2002, "Heteroskedasticity-Autocorrelation Robust Standard Errors Using the Bartlett Kernel Without Truncation," *Econometrica*, Vol. 70 (September), pp. 2093–2096.

Kiefer, Nicholas M., J.Vogelsang Timothy, and Helle Bunzel, 2000, "Simple Robust Testing of Regression Hypotheses," *Econometrica*, Vol. 68 (May), pp. 695–714.

Mincer, Jacob, and Victor Zarnowitz, 1969, "The Evaluation of Economic Forecasts," in *Economic Forecasts and Expectations: Analyses of Forecasting Behavior and Performance*, NBER Studies in Business Cycles, Vol. 19, ed. by Jacob Mincer (New York, National Bureau of Economic Research).

Patton, Andrew J., and Allan Timmermann, 2006, "Properties of Optimal Forecasts under Asymmetric Loss and Nonlinearity," *Journal of Econometrics*, doi:10.1061/j.jeconom.2006.07.018.

Timmermann, Allan G., 1993, "How Learning in Financial Markets Generates Excess Volatility and Predictability in Stock Prices," *Quarterly Journal of Economics*, Vol. 108 (November), pp. 1135–1145.