

VI. STATISTICAL TESTS OF WEO FORECASTS ON GROWTH AND INFLATION

This background document provides additional statistical tests of WEO forecasts. The results confirm the optimistic tendency of WEO forecasts for Africa and Latin America, but find a pessimistic tendency in forecasts for industrial countries and the Middle East. Forecasts for emerging Asia and transition economies are by and large unbiased, partly because of the averaging of positive and negative biases over the sample period. The magnitude and direction of the bias in WEO forecasts depends both on the time period and on the country coverage.

A number of previous studies have addressed the statistical properties of WEO forecasts (see Table 6.1 for an overview of these studies). Earlier studies tended to find that the WEO forecasts were optimistic for developing countries, particularly those in Africa, but failed to find such bias for industrial countries. These broad conclusions are supported by more recent works. For example, the latest work of Timmermann (2006), using data from the 1990s, noted that the WEO's growth forecasts appeared biased for individual industrial countries, but the forecast errors were not statistically significant. For the developing countries, however, the same study found optimism, particularly in the WEO's forecasts for Africa and, to a lesser extent, the Commonwealth of Independent States (CIS) and Latin America. Timmermann (2006) further noted that the optimistic bias was even stronger for countries under IMF-supported programs (see also GAO, 2003).

This document adds to these previous studies by using a more up-to-date set of data (1991–2003), employing panel data regression (as opposed to the pooled regressions of most previous studies and the bootstrapping technique of Timmermann 2006), and comparing the WEO forecasts to the Consensus forecasts as well as the forecasts produced by other international or regional public institutions, including the World Bank, the OECD, the Asian Development Bank (AsDB), the African Development Bank (AfDB), and the United Nations Economic Commission for Latin America and the Caribbean (ECLAC).

Table 6.1 Main Findings of Previous Studies on WEO Forecasts

	Sample Period	Country Coverage	Methodology	Unbiasedness	Efficiency	Comparison with Consensus
Barrionuevo (1993)	1971–91	G-7	Individual and pooled regressions on bias and efficiency	Optimisms in the 1970s; pessimism in the 1980s	Growth forecasts inefficient; inflation forecasts efficient	N.A.
Artis (1997)	1971–94	G-7, five area departments, nonfuel exporters	Individual and pooled regressions on bias and efficiency	No bias for G-7 individually; optimism in Africa, Asia, and Western Hemisphere	Efficient for both G-7 and developing countries	WEO and Consensus mean are similar
Loungani (2000)	1989–98	Consensus countries	Pooled regression on bias and efficiency	N.A.	N.A.	WEO and Consensus mean are identical
Blix et al (2001)	1991–2000	France, Germany, Italy, Japan, Sweden, the United States.	Average Mean Forecast Error; comparison with other institutions' forecasts	N.A.	N.A.	70 percent of private forecasters better than WEO
GAO (2003)	1990–2001	G-7 and 87 emerging market economies	Individual regressions on bias and efficiency	Mostly unbiased; program targets biased but accurate	Efficient	N.A.
Timmermann (2006)	1990–2003	178 IMF member countries	Individual and bootstrapped regressions on bias and efficiency; Diebold and Mariano test	Unbiased for industrial countries; signs of optimism for developing countries	Serial correlation, especially for Central and Eastern Europe and CIS countries	WEO and Consensus mean are statistically similar

A. Data and Methodology

In this document, we restrict our attention to forecasts for real GDP growth and inflation in IMF member countries over 1991–2003. For the WEO, we use the April publications for current year projections and the September publications for year-ahead projections, with actual numbers taken from the April 2004 issue.⁹ Forecasts made by other institutions are selected from the closest relevant publication date so as to minimize any difference in timing.¹⁰ In practice, the current-year forecasts (corresponding to the April WEO) come from the March–June period, while the year-ahead forecasts (corresponding to the September WEO) span the August–November period. The regression results for year-ahead forecasts, however, are not formally reported here because they are nearly identical to those obtained for current-year forecasts.¹¹

Given our focus on the IMF’s overall forecast performance, we are mainly concerned here with averages for country groupings or regions, although we also look at forecasts for individual countries when relevant. Regions are as defined by the WEO, except in the case of transition countries for which we combine Central and Eastern Europe (CEE) and CIS. When comparison is made with Consensus forecasts, we follow the definition of regions given by Consensus Economics, Inc.; when comparing WEO forecasts with those produced by other public institutions, we follow their definition of regions or groups (subject, in the cases of emerging Asia and industrial countries, to the availability of Consensus forecasts).

In view of the structural changes or hyperinflation experienced by some of these countries in the 1990s, we eliminated from the sample those years when extreme values were observed for some regions (e.g., the early 1990s for CIS countries). We also eliminated from the sample two countries experiencing conflict¹² as well as small economies (with an annual GDP of less than \$5 billion), in order to secure a balanced set of panel data.

⁹ Some previous studies have used the realized figures from the next available publication for each year. Our preliminary analysis shows that this particular choice of actuals does not materially change the results of our analysis.

¹⁰ For inflation forecasts by the World Bank and the OECD, we use their indices for private consumption prices. In order to make a meaningful comparison, we compute the respective forecast errors in these series by using their own realized numbers.

¹¹ The only difference is that the WEO’s growth forecasts for industrial countries become unbiased when year-ahead forecasts are used.

¹² Angola and the Democratic Republic of Congo.

The regression analysis consists of estimating the following two models, one on bias and the other on efficiency:

$$\text{Test of biasedness: } F_{it} - R_{it} = e_{it} = \beta_0 + \epsilon_{it}, \text{ with } H_0 : \beta_0 = 0 \quad (6-1)$$

$$\text{Test of efficiency: } R_{it} = \beta_0 + \beta_1 F_{it} + \epsilon_{it}, \text{ with } H_0 : \beta_0 = 0 \text{ and } \beta_1 = 1 \quad (6-2)$$

These regressions are widely used in the literature (see Barrionuevo, 1993; Artis, 1997). Panel data regression is used to address the small sample problem. In particular, we use a panel-data GLS estimator, which yields a weighted average of within-group and between-group estimators, and divides forecast errors by standard deviations to take account of the volatility of the underlying series. We also control for serial correlation (panel-specific where possible) and allows for a heteroskedastic error structure for cross-country differences (but not cross-country correlation).

We report the following estimators of the relevant coefficients: (i) pooled OLS estimators; (ii) GLS panel-data estimators; and (iii) OLS Prais-Winsten (p-w) estimators. Given the fixed sample period and variable sample size (the number of countries could change between 12 and 27), individual effects may not be consistently estimated by the panel-data GLS estimator for some regions. It is for this reason that we supplement the GLS panel estimator with two additional estimators, as noted above.¹³ In comparing two sets of forecasts, we use the mean error (ME), mean absolute error (MAE), and root mean squared error (RMSE) statistics. As the means for the respective regions, we report the means of the ME, MAE and RMSE statistics calculated from the individual countries.

B. The Statistical Properties of WEO Forecasts

Biasedness and efficiency of WEO forecasts

The biasedness and efficiency of WEO forecasts were tested by estimating equations (6-1) and (6-2), respectively. The results are reported in Table 6.2 for biasedness and Table 6.3 for efficiency. These results support the view that WEO forecasts were pessimistic for industrial and Middle Eastern countries (as indicated by positive forecast errors for growth and negative forecast errors for inflation) and were inefficient (as indicated by large p-values of the Wald statistic) during the 1990s.¹⁴ For emerging Asia, however, WEO forecasts were unbiased and efficient for both growth and inflation. It is possible that the lack of bias in

¹³ We also used the Generalized Estimating Equations estimator and obtained substantially the same results.

¹⁴ A closer look at individual countries (not reported here) would show that growth was particularly underpredicted for the United States and the United Kingdom, while it was overpredicted for Germany, Italy and France.

growth forecasts was the result of missing downturns in the late 1990s and the early 2000s, which offset the negative errors committed in the other years (Figure 6.1).¹⁵ For transition economies, growth and inflation forecasts were unbiased, but the inflation forecasts were inefficient. The time-series profiles of forecast errors indicate that the absence of bias for both growth and inflation likely reflect the averaging of positive errors in the early 1990s with negative ones in a later period (Figures 6.1 and 6.2). The inefficiency of inflation forecasts could reflect serial correlation rather than bias, as noted by Timmermann (2006).

Table 6.2. Testing the Biasedness of WEO Current-Year Forecasts, 1994–2003 1/

A. Growth		OLS 2/	GLS 3/	OLS P-W 3/
Industrial	Constant (β_0)	-0.41 (0.10)	-0.36 (0.07)	-0.45 (0.11)
Emerging Asia	Constant (β_0)	0.02 (0.23)	0.00 (0.17)	0.22 (0.27)
Latin America	Constant (β_0)	0.39 (0.19)	0.37 (0.13)	0.33 (0.21)
Transition	Constant (β_0)	0.05 (0.34)	-0.24 (0.26)	-0.03 (0.40)
Middle East	Constant (β_0)	-0.67 (0.30)	-0.75 (0.19)	-0.89 (0.31)
Africa	Constant (β_0)	0.66 (0.19)	0.65 (0.14)	0.82 (0.24)

B. Inflation		OLS 2/	GLS 3/	OLS P-W 3/
Industrial	Constant (β_0)	0.17 (0.06)	0.13 (0.04)	0.17 (0.08)
Emerging Asia	Constant (β_0)	-1.07 (0.59)	-0.74 (0.28)	-1.26 (0.55)
Latin America	Constant (β_0)	0.54 (0.74)	-0.05 (0.20)	-3.23 (2.12)
Transition	Constant (β_0)	-8.91 (6.60)	-3.18 (3.07)	-25.60 (20.01)
Middle East	Constant (β_0)	0.80 (0.63)	0.59 (0.21)	0.96 (0.72)
Africa	Constant (β_0)	-2.11 (0.71)	-0.59 (0.52)	-3.46 (1.14)

1/ Standard deviations are in parentheses.

2/ OLS estimation with robust standard errors.

3/ GLS and OLS P-W (Prais-Winsten) estimations with heteroskedastic and panel-specific correlation.

¹⁵ A closer look at individual countries (not reported here) would show that the lack of bias in inflation forecasts likely resulted from overpredicting inflation in systemically important economies, which counterbalanced the underprediction in other countries in the region.

Table 6.3. Testing the Efficiency of WEO Current-Year Forecasts, 1994–2003 1/

		A. Growth			B. Inflation			
		OLS	GLS	OLS P-W	OLS	GLS	OLS P-W	
Industrial	Constant (β_0)	0.11	0.38	0.18	Constant (β_0)	-0.01	0.04	0.05
	β_1	1.11	0.98	1.09	β_1	0.94	0.92	0.92
	Wald Test	8.38	23.81	17.38	Wald Test	5.39	13.99	4.41
Emerging Asia	Constant (β_0)	-0.61	-0.02	-0.70	Constant (β_0)	-0.55	-0.01	0.43
	β_1	1.11	1.01	1.10	β_1	1.22	1.07	1.10
	Wald Test	0.22	0.03	1.15	Wald Test	2.14	4.16	5.78
Latin America 2/	Constant (β_0)	0.16	0.28	0.17	Constant (β_0)	2.54	1.76	3.90
	β_1	0.82	0.80	0.81	β_1	0.71	0.87	0.66
	Wald Test	5.36	21.57	10.18	Wald Test	1.02	11.26	9.94
Transition 3/	Constant (β_0)	0.62	0.91	0.39	Constant (β_0)	-1.87	-2.44	2.72
	β_1	0.96	0.80	0.89	β_1	1.37	1.31	1.40
	Wald Test	0.10	8.66	1.42	Wald Test	19.74	82.16	26.30
Middle East	Constant (β_0)	2.37	2.35	2.41	Constant (β_0)	-0.50	0.35	0.54
	β_1	0.51	0.52	0.53	β_1	0.95	0.72	0.71
	Wald Test	20.92	59.65	30.63	Wald Test	1.66	14.24	5.18
Africa	Constant (β_0)	-0.76	0.41	-0.37	Constant (β_0)	2.00	0.50	3.15
	β_1	1.02	0.77	0.89	β_1	1.01	1.01	1.02
	Wald Test	5.85	35.05	13.45	Wald Test	5.72	1.44	9.66

1/ Wald statistics represents a test of joint hypothesis $H_0: \beta_1 = \beta_2 = \dots = \beta_n = 0$ and $H_1: \beta_1 \neq 0$ or $\beta_2 \neq 0$ or $\dots$ or $\beta_n \neq 0$.

2/ 1995–2003 for inflation in Latin America.

3/ 1996–2003 for inflation in transition countries.

Figure 6.1. Economic Growth and WEO Current-Year Forecast Errors in Selected Regions, 1991–2003 (In percent per year)

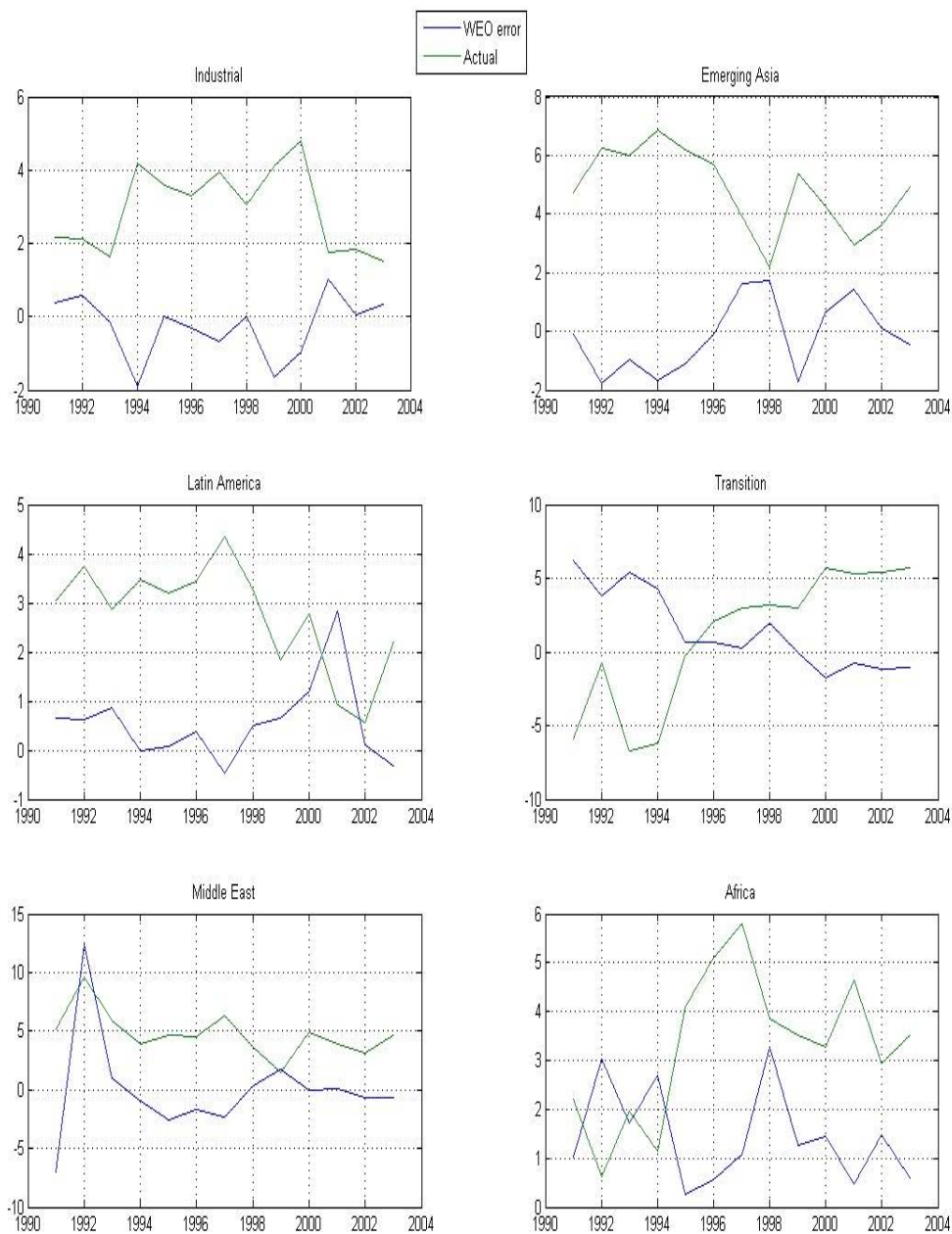
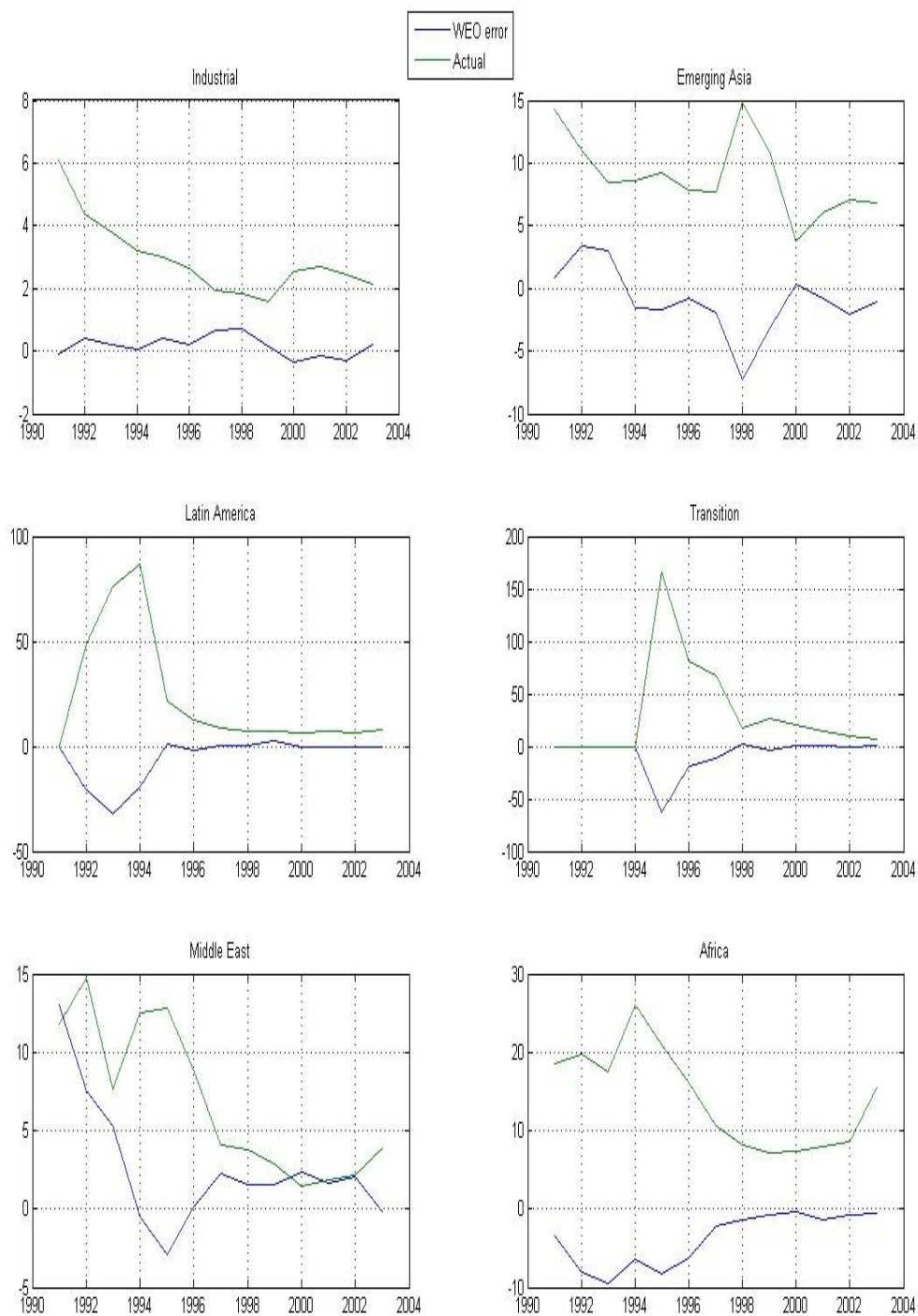


Figure 6.2. Inflation and WEO Forecast Errors in Selected Regions, 1991–2003
(In percent per year)



On the other hand, signs of optimism in WEO forecasts were evident in Latin America and Africa. For Latin America, the optimistic bias was statistically significant for growth but not significant for inflation. The forecasts were inefficient in both cases, likely reflecting the bias in the case of growth and serial correlation in the case of inflation. For Africa, growth optimism is even more evident in the sense that the bias was numerically larger and statistically more significant. A closer look at individual countries (not reported here), however, would show that growth forecasts were optimistic for about half of the countries during 1991–2003, while the forecasts were pessimistic for a handful of countries that were experiencing conflict. The forecasts for inflation were biased but the evidence for efficiency is mixed.

Relative accuracy of WEO forecasts

The relative accuracy of WEO forecasts can be assessed by comparing them to the forecasts produced by other private and public institutions. Table 6.4 reports the MAEs and RMSEs of WEO and Consensus forecast errors. A comparison of MAEs and RMSEs between WEO and Consensus forecasts indicate that the two sets of forecasts were very close, with Consensus performing slightly better for both current year and year-ahead forecasts. WEO forecasts performed better than the means of private forecasts, however, in the case of growth in Europe and inflation in Europe and emerging Asia. A closer examination of forecast errors over time (not reported here) shows that the largest discrepancy was around 0.4 percent for Europe, emerging Asia, and Latin America; a formal test (not reported here) would also indicate that WEO forecasts were not statistically different from Consensus mean forecasts for almost all countries in the sample.

WEO forecasts can also be compared with forecasts produced by other international and regional public institutions in terms of MEs and RMSEs (Table 6.5). For growth, the WEO forecast errors for Africa and Latin America were numerically smaller than those of World Bank, AfDB, and ECLAC forecasts; WEO forecasts for emerging Asia and industrial countries were very close to those produced by the AsDB and the OECD, respectively. For inflation, WEO forecast errors were numerically smaller than those of AsDB forecasts, while neither WEO nor OECD forecasts dominated the other for industrial countries. We may say that, if the IMF forecasts were optimistic for Africa and Latin America, the World Bank, AfDB and ECLAC forecasts were even more so.

Table 6.4. WEO and Consensus Forecasts for Growth and Inflation in Selected Regions, 1991–2003 1/

A. Growth

	MAE				RMSE			
	Current Year		Year Ahead		Current Year		Year Ahead	
	IMF	Consensus	IMF	Consensus	IMF	Consensus	IMF	Consensus
G-7	0.97	0.95	1.28	1.24	1.39	1.36	1.54	1.49
Europe	1.09	1.15	1.40	1.42	1.35	1.39	1.71	1.75
Emerging Asia	1.60	1.47	2.17	1.98	2.14	1.97	3.09	2.87
Eastern Europe	2.93	2.58	3.34	3.06	3.99	3.31	4.47	4.01
Latin America	2.30	1.99	3.25	3.04	2.91	2.48	4.06	3.75

B. Inflation

	MAE				RMSE			
	Current Year		Year Ahead		Current Year		Year Ahead	
	IMF	Consensus	IMF	Consensus	IMF	Consensus	IMF	Consensus
G-7	0.43	0.40	0.63	0.66	0.53	0.48	0.77	0.77
Europe	0.69	0.71	0.92	1.01	0.94	0.94	1.21	1.29
Emerging Asia	1.63	1.80	2.46	3.03	2.23	2.30	3.55	4.29
Eastern Europe	11.15	9.93	29.07	28.04	21.36	17.55	61.63	60.42
Latin America	22.99	21.43	35.81	23.13	47.34	43.01	84.26	56.53

1/ Means of MAEs and RMSEs for individual countries in the respective regions. Regions are defined by Consensus groupings.

Table 6.5. Current-Year Forecast Errors for Growth and Inflation Across Competing Forecasters, 1991–2003 1/

A. Growth

	ME				RMSE			
	IMF	World Bank	Regional Institutions	Consensus	IMF	World Bank	Regional Institutions	Consensus
Africa	0.51	1.01	0.97	-	1.89	2.85	2.91	-
Emerging Asia	0.2	-	0.34	0.21	2.41	-	2.39	2.24
Latin America	0.19	1.96	0.5	-	2.18	3.16	2.35	-
Industrial	-0.12	-	-0.14	-0.19	1.62	-	1.41	1.55

B. Inflation

	ME				RMSE			
	IMF	World Bank	Regional Institutions	Consensus	IMF	World Bank	Regional Institutions	Consensus
Africa	-0.93	-6.01	-11.77	-	4.02	20.29	16.15	-
Emerging Asia	0.35	-	0.44	1.03	2.48	-	3.33	2.63
Industrial	0.06	-	0.17	0.08	1.27	-	1.1	1.15

1/ Regional institutions are the AsDB (for Asia), the AfDB (for Africa), the ECLAC (for Latin America), and the OECD (for industrial countries). Because of data limitations, the sample is restricted to 2002–03 for Africa and Latin America and 1995–2003 for Asia.