IMF Working Paper

Can the IMF's Medium-Term Growth Projections be Improved?

Catia Batista and Juan Zalduendo

INTERNATIONAL MONETARY FUND

IMF Working Paper

Policy Development and Review Department

Can the IMF's Medium-Term Growth Projections be Improved?

Prepared by Catia Batista and Juan Zalduendo[1]

Authorized for distribution by Atish Ghosh

October 2004

Abstract

Numerous reports have noted that the IMF's medium-term growth projections are overly
optimistic, raising questions as to how these can be improved. To this end, we estimate a
growth model and examine its out-of-sample forecasting properties relative to those of IMF
projections. The model's projections outperform those of the IMF in all regions and among
most income groups—projections are less biased (one-quarter of the IMF bias) and have
smaller standard errors (20 percent lower root mean squared errors) even after controlling for
the IMF's macroeconomic assumptions. The paper does not attempt to address the criticisms
that have been leveled against the empirical growth literature, but the results suggest that
benefits can be derived from bringing systematic analysis to bear on cross-country
information.

Contents   Page

I. MOTIVATION

Concerns with over-optimism in the medium-term growth projections used in the IMF's work have been raised in numerous reports and forums.[2] This over-optimism, if present, has three important implications for the design of IMF-supported programs. First, it provides a poor basis for planning fiscal and monetary policies and contributes to fiscal slippages. Second, it distorts the assessment of a country's debt sustainability, thus potentially resulting in the wrong economic policy mix. Finally, over-optimism may also lead to complacency regarding the adequacy of growth-oriented structural reforms pursued by a country.

The evidence presented by Ghosh and Joshi (2003) suggests that an over-optimistic bias exists in medium-term projections—projections have exceeded actual growth rates by an average of about ¾-1 percentage point per year since 1993. An additional concern arises from the current projections for 2003–2007. Specifically, in spite of the deceleration of the world economy that was registered over the past few years, IMF staff projections for low-income countries exceed the performance of these countries during the 1990s by an annual average of 1¾ percentage points and by 1¼ percent relative to the late 1990s (Table 1).

Notwithstanding the evidence on the over-optimistic bias of IMF staff growth projections, it is important to note that these forecasts are meant to be conditional on the implementation of certain policies and on the absence of shocks. Indeed, these projections are not generated by a single model and reflect a varying mix of quantitative analysis and judgment that are, for example, the basis for program design accepted by both government authorities and staff. As such, these projections may entail some degree of compromise. These projections are also complemented by country-specific knowledge. For example, country teams examine one-time factors and sector-specific developments. As a result, we could argue that it is unlikely that short-run growth projections can be improved through the use of analytical tools.[3] Finally, it is important to note that IMF staff projections are prepared as input for the *World Economic Outlook* (WEO), which focuses primarily on the first two years of the 5-year ahead projection horizon of this paper. In this regard, the 5-year ahead projections prepared by IMF staff are not subject to the consistency checks that characterize WEO forecasts.[4]

---

[2] For instance, reports reviewing the Poverty Reducing Strategy Papers (PRSP) and the Poverty Reduction Growth Facility (PRGF), the reports on prolonged users of IMF resources and on fiscal adjustment in IMF-supported programs prepared by the IMF's Independent Evaluation Office, and reports on IMF projections prepared outside the institution, such as the U.S. General Accounting Office (2003) and by Beach et al. (1999).

[3] A review of IMF country reports for 12 countries over a 6-year period (1997–2002) reveals that analytical tools are used sparingly and only in one-half of these countries. It is also plausible that these tools are used but not referred to in the staff reports examined.

[4] Conventions on real exchange rates are established and the output gap is assumed to close by the end of the forecast horizon; i.e., a 2-year ahead projection.

Table 1: Average Annual Growth Rates in Real GDP
(5-Year Periods based on Income Level)

Country income level	1993–1997 Actual	1998–2002 Actual	2003–2007 (Fund proj.) 1/
High-income	3.6	2.7	3.0
Upper middle-income	3.4	2.0	3.8
Lower middle-income	2.7	3.2	3.9
Low-income	2.3	3.5	4.8

1/ Excludes the 5 largest and the 5 lowest projected growth rates.

Against this background, our goal is to develop a simple empirical model that uses information available to IMF staff and that can be tested in its medium-term forecasting qualities against the IMF staff medium-term growth projections for each country. More precisely, the exercise carried out entails the forecast of average annual growth rates for 5-year ahead periods (our medium-run horizon; i.e., 1996–2000) based on an empirical growth model and the comparison of these projections with the ones prepared by Fund staff for the same period and with the caveats noted in the preceding paragraph. Our focus on medium-term projections is in part driven by our belief that 1-year ahead projections by Fund staff are based on an assessment of one-time factors and sector-specific developments that cannot be easily incorporated into formal modeling. Moreover, analytical tools based on growth equations are only useful in improving medium- and long-term projections, which is also consistent with the main focus of the empirical growth literature. The paper also attempts to model the importance of institutions for growth. To this end, we include in one of our econometric specifications the index on country risk known as the ICRG index on institutional factors—i.e., the International Country Risk Guide prepared by the PRS group.

The model developed in this paper has the added advantage of making systematic use of cross-country information. While IMF staff have a detailed understanding of country-specific and one-time factors that may affect growth in a specific country, their projections may not fully take into account the growth experience of other countries that are at similar stages of development and that face similar initial conditions and shocks. There are obvious trade-offs in going from country-specific knowledge to a cross-country approach, but the paper's results suggest that the informational content of cross-country data is useful when the focus is on medium-term growth projections—i.e., 5-year ahead average growth projections.

This paper does not attempt to address the criticisms usually leveled on the empirical growth literature—from the lack of a theoretical foundation to the instability of coefficient estimates. The paper is also limited by the fact that only one out-of-sample 5-year period is examined. Partly owing to these factors the paper does not propose to replace how country teams prepare medium-term growth projections. In fact, the paper's goals are more modest; namely, to stress the usefulness of looking into cross-country experiences when preparing growth projections, particularly among developing countries. Indeed, in more advanced economies where economic relationships are better established, it would be advisable to proceed with more complex analytical tools than those described in this paper.

The paper is organized as follows. The next section briefly reviews the existing literature on economic growth. This is followed by a discussion of the empirical framework used in this paper, the econometric methodology chosen to carry out our empirical work, and the main results from the econometric estimation. The section that follows constitutes the core of the

paper: we specify rules for out-of-sample projections and the latter are then contrasted against those prepared by IMF staff. Concluding remarks follow.

II. REVIEW OF THE LITERATURE

The neoclassical growth model[5] rationalizes the observation of positive, non-decreasing per capita long-term growth as a result of exogenous technological change. This is a direct consequence of the model's assumption of diminishing returns, which implies that the economy should reach a steady state with constant output per effective worker (where the growing number of effective workers reflects population growth and technological change). But the neoclassical models' inability to provide an explicit explanation for long-term growth inspired a new wave of research, which started in the mid-1980s with the seminal papers by Romer and Lucas. These studies made economic growth endogenous by allowing non-diminishing returns as a result of knowledge spillovers and externalities in human capital.[6]

An issue related to this theoretical evolution is the empirical validity of alternative growth models. On the one hand, neoclassical models predict that the level of income per capita should converge toward its steady state value.[7] According to these models, changing an economy's structural characteristics (such as the rate of time preference) generates transitional growth from one steady state to the other. Endogenous growth models, on the other hand, predict that positive economic growth is plausible in the steady state. Therefore, changes in the structural parameters of an economy may cause permanent shifts in its economic growth rate. As discussed by Wacziarg (2002), the distinction between transitional and steady state growth is not easy to make empirically since the time span for which cross-country data are available is small relative to the typical transitional convergence period predicted by neoclassical models.[8] This implies that empirical work on the determinants of economic growth must proceed without a solid theoretical foundation, which has several important implications, as we attempt to make clear in the remainder of this section.

---

[5] The neoclassical growth literature was developed in the 1950s and 1960s by Solow, Swan, Cass, and Koopmans. It was preceded by Harrod and Domar's attempt to fit the study of growth into the Keynesian framework.

[6] For a summary of endogenous growth theories see Barro and Sala-i-Martin (1995) and Aghion and Howitt (1998).

[7] The assumption of diminishing returns to capital implies that the speed of convergence increases with how far a country is from its own steady state. If we control for the determinants of the steady state, we thus have a (conditional) convergence prediction: lower initial values of per capita output generate higher transitional growth rates. Even though diminishing returns are not generally characteristic of the endogenous growth literature, conditional convergence can also be achieved in these models (see Sala-i-Martin, 2002).

[8] Barro and Sala-i-Martin (1995) estimate a half-life of transition to the steady state of around 35–40 years for the Solow model in datasets covering periods until the late 1980s.

One main strand of the empirical growth literature, started by Robert Solow, consists of growth accounting exercises in which growth is broken down into its sources: factor accumulation and a residual reflecting technological change (referred as total factor productivity or TFP). Barro (1999) argues that this exercise may be used as a preliminary step for the analysis of the determinants of growth (such as government policies, household preferences, initial endowments of natural resources, and different types of capital, including human capital) if these are largely independent from the determinants of technological progress. Because this independence is unlikely, however, the growth accounting approach fails to deliver a convincing analysis of the decisive causes for factor accumulation and increased productivity, which constitute the ultimate determinants of economic growth.

An alternative approach in the empirical growth literature has been more directly focused on explaining the fundamental determinants of economic growth. The basic method of dealing with the endogeneity of factor accumulation variables (investment being the most prominently discussed) is to think of a structural model in which growth is decomposed into its basic sources (i.e., factor accumulation and total factor productivity) and these sources are determined by the fundamental determinants of growth (e.g., government policies, private sector choices). While this is the appropriate framework underlying any reduced form growth regression, many studies have followed a less satisfactory approach, in effect combining contemporaneous indicators of factor accumulation with policy variables as regressors. An additional problem that remains even with well-specified reduced form regressions is the difficulty in interpreting the transmission channels through which growth takes place.[9]

One of the first empirical studies using reduced form cross-country growth regressions, while taking into account endogeneity issues, was prepared by Barro (1991), who used a single cross-section of countries with observations averaged over a single period. After this study, time variation has been added by using panel regressions, more efficient simultaneous equation three-stage least square (3-SLS) methods, and even dynamic GMM methods. At the present time, there is an extensive cross-country reduced form regression literature.[10] As discussed by Sala-i-Martin (2002), this has, nevertheless, failed to produce simple and convincing results on the determinants of growth or their quantitative contributions. Structural policies (as openness to trade or institutions related to economic freedom), for instance, have been determined to positively affect growth, but the magnitude of their impact cannot be robustly ascertained using the results in the literature.

Related to this paradoxical failure is the main criticism to the empirical growth literature, summarized by the robustness (or lack thereof) debate on its results. The fact that there is no theoretical model supporting the determination of which regressors to include has opened the floodgates for an empirical literature that includes a wide variety of indicators, many of

---

[9] See Fischer (1993) and Stokey (1994) for a discussion of these problems. Recent studies, such as Tavares and Wacziarg (2001), have estimated structural versions of economic growth models that make these transmission channels more explicit.

[10] The reader is referred to Temple (1999) for a review of the recent empirical growth literature, in particular regarding its numerous shortcomings.

which have been found to contribute significantly to growth. As Wacziarg (2002) poses it, by now there is a "well-established tradition of throwing every variable under the sun into the kitchen sink of growth regressions." In response to this practice, Levine and Renelt (1992) used an extreme bounds analysis to find that very few of the commonly used regressors seemed to explain growth. In a similarly spirited exercise, Sala-i-Martin (1996) was however able to identify robustness in more variables than Levine and Renelt.

III. METHODOLOGY AND RESULTS

A. Empirical Framework

Given the state of the literature reviewed in the previous section, we opted for the construction of a reduced form empirical model that is estimated using both the cross-country and the time variation in the data. As already noted, we do not address the shortcomings of the empirical growth literature, but we attempt to circumvent some of these problems by implementing a carefully structured econometric procedure. Our limited goals are also justified by the fact that the main focus of our paper is, in the end, to assess the model's forecasting qualities vis-à-vis the medium-term projections of Fund staff.

The theoretical model underlying our econometric estimation is taken from Barro (1997) and may be described as

$$g = f(y\,;\,y^*)$$

where y denotes the initial level of per capita output, y^* is the steady state level of per capita output (which depends on an array of control and environmental variables), and g stands for the growth rate of per capita output.[11] We use this framework as a foundation for our reduced form growth regression, since we are interested in the growth effects of some fundamental variables and not in whether they act through factor accumulation or productivity growth.

The model proposed is not a structural model, but is directly inspired by neoclassical growth theories in the sense that it takes growth as a transitional deviation from the steady state level. It is, however, extended to include factors highlighted by endogenous growth theories, as are the case of human capital, government policies, and the diffusion of technology. This approach serves our objective of evaluating the medium-term growth impact of policies, rather than explaining continued long-term positive global per capita growth. Nevertheless, transitional growth seems to be such a slow process that the distinction of whether we are looking at transitional growth or at steady state growth (potentially affected by changes in its structural determinants) is not very relevant for our purposes.

---

[11] The function f is assumed to possess the following properties: $\dfrac{\partial f}{\partial y} < 0$; $\dfrac{\partial f}{\partial y^*} > 0$.

Our specification uses per capita real GDP growth as the dependent variable (g_y using the notation of our model).[12] Unlike most cross-country regression studies, we chose not to use this variable in PPP terms. The reasons for this choice hinge on several factors. First, when a country's price structure changes significantly over time, the real growth rate implied by PPP data is likely to differ from that implied by domestic price data, and this describes more accurately the incentives faced by domestic economic agents.[13] Second, our purpose is to explain growth rates and not to compare GDP levels across countries, which does not require the use of an international comparable currency. Finally, our goal is to compare the model's projections to those of Fund economists, which are not PPP-based.

We consider several indicators of initial real income and human capital levels as state variables y (those which determine, according to our model, the convergence speed toward the steady state). Initial income (or its gap relative to a benchmark country) is used as a measure of initial endowments of physical capital, natural resources, and technology. Lower levels of initial output should contribute to faster (conditional) convergence. Inclusion of initial human capital levels reflects the emphasis of the literature (both at theoretical and empirical levels) on human capital accumulation as key to growth, namely through its crucial role in facilitating technological absorption in a globalizing world with fast technological change. Therefore, although the direct effect of initial y on output growth is expected to be negative, the growth impact of human capital should be positive.

Regarding control and environmental variables (those which determine the steady state level y^* in our model), we considered several different categories of indicators. The control variables include government's choices (macroeconomic management and structural variables) and private sector choices. Environmental variables refer to internal and external conditions that do not result from any agent's choices. The description of which variables were chosen within each category and of the reasons for this choice is developed below.

- Regarding government's choices that relate to macroeconomic management, we followed Fischer (1993) in considering inflation (level and variability) and fiscal balance as complementary indicators of the government's ability to manage the economy towards a path conducive to sustainable growth rates. Non-linear relationships between growth and inflation are also examined.

---

[12] The reader is referred to the appendix on data issues for details on sources and characteristics of the regressors used in the empirical work of this paper.

[13] This is an implicit recognition that economic agents in the non-tradable sector react mainly to domestic developments. This reasoning is attributed by Temple (1999) to Daniel Nuxoll. We were also inclined to use national accounts data both because of availability reasons (i.e., country coverage in Penn data is less complete) and the fact that this data reflects the most up to date statistical information in IMF and World Bank databases. Similarly, Pritchett (1998) argues that the time series variation in the Penn dataset comes primarily from national accounts data given the few benchmark points that exist in the Penn dataset, which is based on purchasing power parities. This is in particular the case among developing countries.

- Still regarding government's choices, on the subject of structural variables relevant to medium-term growth, we examined the role of: the size of government in the economy (as a measure of potential growth losses due to unproductive spending, following Barro, 1991); the depth, development, and intermediation role of the financial sector (which should be a factor in increasing economic efficiency according to King and Levine, 1993); the degree of openness of the economy to trade (a positive contributor to growth by increasing the scale of production and enabling better resource allocation, as studied by Sachs and Warner, 1995); the black market exchange rate premium (as an indicator of price distortions in the foreign exchange market, although it may also be regarded as an indicator of exchange rate policy, thus falling into economic management); and institutional indicators reflecting political stability and the security of property rights (Knack and Keefer, 1995, and Mauro, 1995, helped to empirically establish the relation between institutions and growth).

- One important set of private sector choices is captured by the inclusion of fertility rates, which we think of as a proxy for policies affecting labor supply. These rates impact both the size of the labor force and, potentially, participation rates due to the additional resources spent on child bearing activities (Barro and Becker, 1988).

- Finally, in the category of environmental variables, we want to look at both internal shocks (such as wars and weather disturbances, which are likely to cause major drops in agricultural production) and external shocks (such as change in the terms of trade).

B. Econometric Methodology

We chose a 2-SLS estimation method with GLS correction for heteroscedasticity (Table 2 describes alternative estimation methods) for an unbalanced panel of 5-year periods from 1961 to 2000. Our choice of estimation method was made on the following grounds.

First, using the time variation in the data in addition to the cross-section variation may help improve estimates relative to a standard cross-country growth equation (which, in our case, would involve using observations for 40-year averages in each country). The time variation in our panel dataset is picked up by the use of the 5-year periods in 1961–2000.

Second, our choice of whether to use a limited (LIM) or a full information estimation method (FIM) is subjective—we opted for a LIM method. In our case, the FIM (3-SLS) basically consists of adding a GLS correction to the corresponding LIM (2-SLS) in a way that improves the efficiency of its estimates by controlling for the correlation among error terms in different periods. The drawback of FIMs is that a misspecification anywhere in the system would result in inconsistent estimated coefficients; whereas, when using a LIM, a misspecification for a certain period affects only the coefficients for that period. Because we were worried about such potential misspecifications given the time span of our data (40 years) for which we identify structural change before and after the oil shocks of the 1970s (and because we found no evidence of significant correlation between the error terms of different periods), we decided to trade off potentially small efficiency gains by likely consistency gains. The lack of autocorrelation in the residuals reflects in part the 5-year

periodicity of the data.[14] A GLS efficiency improvement was nevertheless kept to correct for heteroscedasticity, which is likely to occur given that most of the variation is cross-sectional.

Table 2. Alternative Estimation Methods

Source of variation	LIM vs. FIM	Instrument?	
		No	Yes
Cross-section variation only	Limited information methods (LIM)	OLS	2-SLS
Cross-section and time variation 1/	Limited information methods (LIM)	OLS	2-SLS
	Full information methods (FIM)	SUR	3-SLS

1/ Any of the alternatives in this section allows econometric consideration of omitted time-invariant variables through the introduction of fixed or random effects.

Third, the decision to instrument the potentially endogenous regressors included in our sample was straightforward given the evidence in the empirical literature on this matter. Instrumenting was also viewed as a way of reducing temporary measurement errors and cyclical factors that may affect the initial values of the state conditions. We chose to instrument initial income, human capital, and fertility rates by their own 5-year lagged value. The use of 5-year averages reduces the risk of other sources of endogeneity.

Fourth, still regarding our estimation method, a caveat should be raised on the issue of potentially omitted time-invariant regressors. This problem is usually tackled in the panel data literature by introducing fixed or random effects. Since random effects require that the omitted variables be exogenous relative to the included regressors, an unlikely event, the random effects model was not viewed as a valid option. The inclusion of fixed effects is also not exempt of problems as it allows neither for the inclusion of other country-specific time-invariant regressors, nor for the use of lagged value instrumental variables.[15]

Finally, a word or two on the main weaknesses of the proposed econometric approach is warranted. First, there is no structural model supporting our estimation (as is also the case of most of the existing empirical growth literature), which raises robustness concerns of the kind highlighted by the Lucas critique on policy evaluation. We nevertheless attempt to construct a sensible reduced form growth equation in the light of current theories on growth determinants, which are summarized by our choice of state, control, and environmental

---

[14] As a robustness check, we also run 3-SLS on our basic model. This procedure confirmed that the estimated coefficients for the subperiod 1981–2000 remain broadly unchanged. Barro (1997, page 15), who uses a similar econometric approach, also argues that the "estimates are virtually the same if the errors are assumed to be independent over the time periods."

[15] Nevertheless, the coefficient estimates obtained using panel regression methods with fixed and random effects were similar to those obtained using our method of choice. This suggests that there should be no sizeable omitted variable bias, and that there is indeed no significant correlation between the errors corresponding to different periods.

variables. Second, also present are potential selection bias problems. Specifically, the unbalanced panel reflects data availability, which may have consequences for the estimated results—e.g., countries with less data availability might have characteristics relevant for their growth process that are different from those of countries with more available data. The fact that developing and transition economies are under-represented in our sample, in particular in the 1960s–70s, is a factor that deserves to be noted. We mitigate selection bias problems by weighting the periods according to the number of available observations in each period, carrying out the econometric work for periods before and after the oil shocks of the 1970s.

C. Data

The estimations of our model are largely based on data available in the IMF and the World Bank databases. These databases are complemented by the Penn World Tables data on initial income (also known as the Summers-Heston dataset, which is specified in comparable international prices); the Barro-Lee dataset on educational attainment; and a dataset on exchange rate black market premium[16]; and a dataset on institutional factors (ICRG index). Countries were dropped based on data availability and we also specified rules for eliminating non-sensical data.[17] Selection bias arising from data availability might exist. However, given that our goal is to test the model's forecasting qualities vis-à-vis those of Fund staff for the same set of countries, the presence of selection bias is not a major source of concern.

In addition, in one of our specifications we include the political risk component of the ICRG index. This component represents indicators of government stability, socio-economic conditions, investment profile, ethnic tensions, rule of law, internal conflicts, external conflicts, role of religion in politics, role of military in politics, quality of bureaucracy, corruption, and democratic institutions. The paper finds that these factors have an important contribution to growth—about ¼-½ percent in annual growth rates for a one standard deviation change in the average index of a country. A separate paper examines the characteristics of institutions that matter the most for growth; namely, the stability or predictability of institutions and the environment these provide for accountability.[18]

D. Econometric Implementation and Results

The procedure used to select a model specification involves using one regressor for each category described in sub-section A. The least significant regressor in each round was replaced by a regressor in the same category if one existed or, otherwise, the category was

---

[16] The black market exchange rate premium data is based on work by Reinhart and Rogoff.

[17] The IMF database for the 1960s and 1970s, and to a lesser extent the World Bank database, present deficiencies that are addressed by specifying rules for dropping suspect data; e.g., inflation rates that remain constant over many years were eliminated from the dataset. Transition economies were included only for the last 5-year period (1996–2000) to discard biases arising from the large swings in output these countries have experienced.

[18] Zalduendo (2004). "The Design of Macroeconomic Policies and Structural Reforms," IMF, mimeo.

dropped. A complete description of the regressors examined can be found in the data appendix. This procedure was repeated until all regressors were significant at a 10 percent level. The dependent variable is the 5-year annual average growth of real GDP per capita.

This implementation method was carried out for the full sample period (1961–2000). Regressors are thus selected on the basis of their statistical significance in this full period. However, data availability implies that two-thirds of the observations belong to the 5-year periods between 1981 and 2000. Moreover, the results were sensitive to tests that identify the presence of structural changes in the estimated coefficients: a structural break is identified as taking place in the 5-year period that begins in 1981, possibly related to this being the first full 5-year period that follows the oil shocks.[19] Given the presence of these structural breaks, much of the analysis in the paper focuses on the last two decades, thus addressing to some extent the time effects that have been noted in the empirical growth literature. The selected time period has the additional advantage of extending the data coverage of the empirical literature, which has in practice focused on datasets that end in the late 1980s.

The main regression results, summarized in Table 3, are:[20]

• The results confirm the existence of conditional income convergence between poor and rich countries consistent with the findings of the existing empirical literature. The implied speed of convergence, 1¼ percentage points per year, is, however, lower than that reported in other studies.[21] The lower convergence rate reflects, in our view, the extension of the estimation so as to include data for the 1990s and warrants further research. Figure 1shows the familiar contrast between absolute and conditional convergence, the latter in effect after controlling for all other medium-term growth determinants. The horizontal axis plots the country's initial income (in logs) and the vertical axis the per capita growth rates.

---

[19] A Chow test suggests the presence of a structural break in the estimated coefficients that begins with the 5-year period 1981–85. See Ben-David and Papell (1997) and Easterly (2001a) for a discussion on growth trend breaks following the oil shocks of the 1970s.

[20] The use of GLS methods precludes the calculation of a meaningful R-squared statistic. Specifically, an R-square statistic computed from GLS sums of squares need not be bounded between 0 and 1 and does not represent the percentage of total variation in the dependent variable. We address these shortcomings by presenting in our tables the standard error of the regression and the standard deviation of the dependent variable.

[21] The speed of convergence can be derived from the log-linearization of a Solow model around its steady state. The coefficient on initial income level is transformed into speed of convergence φ given that $\beta_1 = -\left[\dfrac{1-e^{\varphi t}}{t}\right]$, where t is the number of years and β_1 is the coefficient on initial income. Our φ is 1¼ percentage points per year, lower than the convergence rate (1½–2½) reported by Mankiw et al. (1992) and Barro (1997). This rate implies that it takes a developing country with one-fourth of the U.S. income per capita (but with the same steady state) about 85 years to bridge by half the income gap. Studies that use data for the early 1990s report equally low convergence rates (Kochhar and Coorey, 1999).

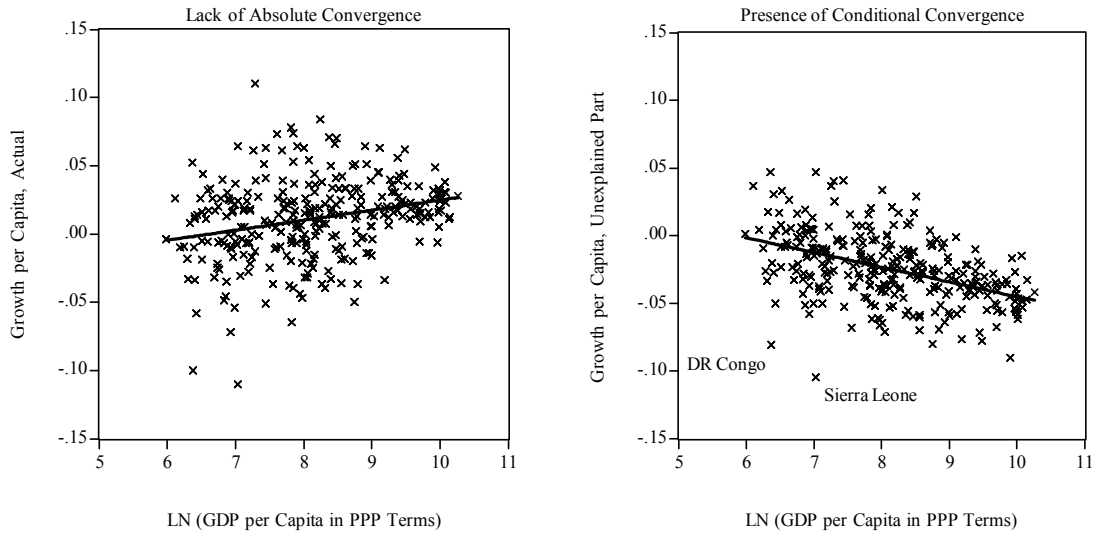**Figure 1: Evidence of Absolute and Conditional Convergence (1981–2000)**

Table 3 also reports the mixed results, in terms of both the sign and statistical significance, of the estimated coefficients for human capital. Specifically, the Barro and Lee data on educational attainment were found to be statistically significant in the period 1961–2000 only because of a large and positive impact on growth during the 5-year periods prior to 1981. The lack of significance in the 1981–2000 period is in all likelihood related to the inclusion of fertility rates as a regressor, which has a high correlation with human capital in more recent 5-year periods.[22] However, since the conclusions in the paper are not affected by dropping either of these two regressors (in particular, the forecasting qualities of the next section remain broadly unchanged), we decided to keep both regressors in our basic model specification, a decision that follows from the econometric implementation procedures described earlier in this sub-section.[23] Moreover, the use of secondary school enrollment rates had the same sign and significance problems as the Barro-Lee education data and the conclusions derived relative to the forecasts from IMF staff did not change. Main results are:

- A number of indicators of macroeconomic management were tested (see data appendix), but only inflation (including a threshold component) and the central government fiscal balance were statistically significant. These results are in line with those of other authors (e.g., Fischer, 1993). We followed Ghosh and Phillips (1998) and Kochhar and Coorey (1999) in modeling a non-linear impact of inflation on growth both through the use of a logarithmic function to reflect "level effects" and through the introduction of a "low-inflation threshold effect". The reasons to proceed in this manner were: first, the fact that an

---

[22] The forecasting qualities of the model, which are discussed in the next section, weakened only marginally when either human capital or fertility was excluded from the regression.

[23] As previously explained, the selection of regressors in our model is based on their performance over the full sample period (i.e., the period 1961–2000).

Table 3. Main Estimation Results

| | 5-year periods in: | | |
	1961–1980	1981–2000	1961–2000
Number of observations	136	298	434
Number of countries	61	89	89
Initial income level	-0.0083 *	-0.0102 *	-0.0128 *
	-4.78	-6.65	-9.28
Initial human capital	0.0039 *	-0.0004	0.0014 *
	4.93	-0.89	2.87
Inflation (CPI)	-0.0008	-0.0046 *	-0.0039 *
	-0.66	-6.69	-4.99
Inflation threshold effect	-0.0022	0.0068 *	0.0052 **
	-0.47	3.27	2.38
Fiscal balance	0.0392 ***	0.1110 *	0.1080 *
	1.77	4.93	5.3
Openness	0.2177 *	0.0866 *	0.0884 *
	7.99	7.08	7.39
Share of private credit	0.0888 *	0.0091 **	0.0260 *
	9.02	2.09	3.96
Internal shocks (weigthed)	-0.0402 *	-0.0292 *	-0.0354 *
	-6.73	-6.96	-7.62
Terms of trade, growth rate	0.0191	0.0225 ***	0.0256 **
	1.34	1.85	2.31
Fertility rates	-0.0012	-0.0107 *	-0.0074 *
	-1.35	-11.7	-8.08
Constant	-0.1042 *	0.0833 *	0.0716 *
	-3.50	4.89	4.45
Wald statistic	3704.34 *	637.18 *	527.04 *
Standard error of regression	0.0271	0.0225	0.0254
Standard dev. dependent varia	0.0313	0.0296	0.0305
R squared (see footnote 19)	0.24	0.42	0.30

* indicates significance at 1 percent; ** indicates significance at 5 percent; and
*** indicates significance at 10 percent.

increase in inflation by 10 percentage points is more negative for growth when inflation
levels are low—an increase in inflation from 10 to 20 percent has a greater negative impact
on growth than an increase from 80 to 90 percent (thus the use of the logarithmic function to
model this convex relation); and, second, the existing empirical evidence suggesting that low
inflation rates (below 5 percent in most studies) have a positive impact on growth. These
results were in line with those of other studies. The level effect had the expected sign
(negative) and was highly significant. The low-inflation threshold effect was also statistically
significant and had the expected positive sign; specifically, inflation rates below 3.5 percent
are positively correlated to per capita growth.[24] The fiscal indicators had the expected impact,

---

[24] The low-inflation threshold effect was identified by finding the inflation rate at which this
regressor becomes statistically significant (1 percent level). This approach also increases the
value of the Wald test on joint significance. It is worth noting that this threshold takes place
at higher inflation rates in the estimation that covers the 5-year periods that end in 1980. The
threshold level we identify is similar to the one mentioned by Ghosh and Phillips (1998), but

a positive relation between fiscal balance and growth in GDP per capita.[25] This reflects the importance of conditions for economic stability and the benefits derived from the availability of resources for private sector activity as a result of strong fiscal positions.

•   Additional regressors were added to reflect other policy choices faced by countries, even though these change less from one year to the next. The estimated coefficients were in all cases significant and had the expected sign. The regressor on the financial sector aims at modeling the role played by deposit money banks in supporting private sector development. Our openness indicator, which controls for scale effects arising from country size, models the importance for growth of exploiting the advantages of trade.[26]

•   Fertility rates were added as a regressor to instrument for policies that affect labor accumulation. The assumption is that the control variables described so far do not adequately represent the determinants of labor accumulation, perhaps because these are more closely linked to the factors that affect total factor productivity and capital accumulation. The negative impact of fertility rates on growth reflects both an increase in the size of the labor force (which reduces the quantity of capital available per worker) and the need for the economy to spend more resources on child bearing activities.

•   The model also includes environmental variables of internal and external origin. Both regressors were found to be statistically significant and had the expected sign. The dummy for internal shocks and the terms of trade growth rate were viewed as central to the growth process, particularly in developing countries that have a narrow export and production base. Thus, their inclusion helps improve the overall fit of the model.

Some growth payoffs from good policies can be summarized by the following calculations—based on the 1981–2000 estimation presented in Table 3 (Table 4 presents descriptive statistics). First, a country that increases its rate of inflation from 2.5 percent to 5 percent will suffer an annual decline in growth rates of 1/10 percentage point. Second, low inflation countries receive a growth payoff for each percentage point in higher inflation achieved below the 3.5 percent threshold level (e.g., increasing the inflation rate from 2 to 3 percent implies an increase in annual growth rates of ¼ percent per year). Third, an additional fiscal

---

smaller than the one referred to by Kochhar and Coorey (1999)—2.5 and 8 percent threshold levels, respectively. The differences could be due to data periods and country coverage. In particular, there has been a marked decline in inflationary experiences over the past two decades, which match the period used in our econometric estimation.

[25] There are in practice no good instruments for the fiscal balance, thus precluding the use of instrumental variables to tackle the link between fiscal balance and growth. In our view, however, this link is diminished by our focus on a panel dataset with 5-year averages.

[26] An index in the interval [0, 1] constructed from the residuals of a regression of export plus imports as a share of GDP on the natural logarithm of population. The goal is to control for country size—large residuals indicate greater openness and vice versa for small residuals.

Table 4. Descriptive Statistics for 1981–2000

	Mean	Standard deviation
Number of countries	89	
Number of observations	298	
Actual growth per capita (in percent)	1.08	2.96
Predicted growth per capita (in percent)	1.12	1.86
Error (in percent)	-0.04	2.25
Initial income level (logs)	8.12	1.06
Initial human capital (number of years)	4.96	2.88
Inflation (logs; absolute value)	-2.34	1.37
Inflation threshold effect (as defined)	-0.10	0.44
Fiscal balance (in percent of GDP)	-3.74	4.46
Openness (defined in interval [0, 1])	0.51	0.07
Credit to private sector (defined in [0, 1])	0.84	0.12
Internal shocks (weigthed; defined in [0, 1])	0.12	0.16
Terms of trade, growth rate (in percent)	0.76	7.64
Fertility rates (number of children per womar	3.89	1.83

deficit equivalent to 1 percentage point of GDP affects annual growth rates by ⅛ percentage points. Fourth, other policy choices (trade openness, financial sector role) also have an impact on growth—respectively, 0.9 and 0.1 percentage points on growth rates for a ten-percentage point change in the defined interval (between 0 and 1) for these regressors.

Table 5 presents the results of three models estimated for the period 1981–95. The estimation was constrained to data until 1995 so that out-of-sample forecasts could be carried out for 1996–2000.[27] We use the model from the baseline specification—referred from now on as basic model—and variants that include institutional factors and proxies for exchange rate policy.[28] The main additional results to those described in the preceding paragraphs are:

- The estimated coefficient on the political risk index—the ICRG index—is statistically significant and positively correlated with growth.[29] We chose the political risk index instead of the composite risk index (which also includes economic and financial risks) because the former mirrors more closely the growth determinants not yet included in our model.[30]

---

[27] The coefficient estimates for terms of trade and openness increase, albeit only slightly, when the period 1996–2000 is excluded from the estimation period. A Chow test rejects, however, the notion that a structural change exists during the last 5-year period of the 1990s.

[28] Data availability precludes testing these regressors for the full period (i.e., 1961–2000).

[29] Linear extrapolations within each period were used to complete the missing data for the ICRG index in the early 1980s—this data is collected annually only since 1984.

[30] This political risk index is a weighted average of several indicators, the highest weight being attributed to institutional factors, such as government stability and investment profile. We interpret this index as a proxy for security of property rights indicators and other institutional factors that support private activity and the efficiency of the government sector.

Table 5. Estimation Results in Models with Human Capital
(5-Year Periods between 1981 and 1995)

	Basic model	Model with ICRG index	Model with BMP
Number of observations	211	195	164
Number of countries	73	69	58
Initial income level	-0.0110 *	-0.0107 *	-0.0105 *
	-8.71	-5.74	-4.94
Initial human capital (HK)	-0.0001	-0.0013	-0.0011
	-0.17	-1.58	-1.34
Inflation (CPI)	-0.0039 *	-0.0009 *	-0.0037 *
	-5.47	-5.51	-4.89
Inflation threshold effect	0.0064 *	0.0066 *	0.0030
	3.43	3.11	0.98
Fiscal balance	0.1261 *	0.0794 *	0.0610 **
	6.84	2.85	2.12
Openness	0.1355 *	0.1070 *	0.1022 *
	6.12	4.97	4.96
Share of private credit	0.0084 **	0.0074	0.0074
	2.08	1.38	1.00
Internal shocks (weigthed)	-0.0300 *	-0.0349 *	-0.0329 *
	-9.81	-7.13	-5.43
Terms of trade, growth rate	0.0349 *	0.0369 *	0.0218 ***
	2.61	2.79	1.73
Fertility rates	-0.0103 *	-0.0092 *	-0.0111 *
	-10.49	-7.33	-8.44
ICRG index		0.0305 *	
		2.79	
Black market exchange rate premium (BMP)			-0.0007 **
			-2.14
Constant	0.0657 *	0.0583 *	0.0874 *
	3.65	2.69	3.50
Wald statistic	609.73 *	397.25 *	527.04 *
Standard error of regression	0.0255	0.0253	0.0246
Standard dev. dependent variable	0.0304	0.0306	0.0290
R squared (see footnote 19)	0.43	0.46	0.50

* indicates significance at 1 percent; ** indicates significance at 5 percent; and ***
indicates significance at 10 percent.

- The black market exchange rate premium (or BMP) was also added. This premium can be interpreted as reflecting the exchange rate policy of a country (a macro interpretation) or as a proxy for price distortions (a micro interpretation). Either way this regressor was found to be statistically significant and negatively related with growth, though the magnitude of the coefficient suggests it has only a limited impact on growth rates.

Once all statistically significant regressors were identified, some robustness checks were carried out. First, an attempt was made to verify if factor accumulation had been adequately accounted for by the model specification. For this purpose, the model was run including the investment rate (instrumented by its lagged value) as a regressor and the coefficient estimate was found not to be statistically significant, suggesting that the other regressors in the equation adequately account for factor accumulation. Second, the model was extended by defining different initial years in our non-overlapping 5-year periods, with the estimated

coefficients remaining broadly unchanged in spite of the resulting reduction in the number of 5-year periods that correspondingly enter in our econometric estimation.[31]

IV.   GROWTH EQUATIONS VERSUS CONDITIONAL FORECASTS BY IMF STAFF[32]

As previously noted, our main goal is to evaluate the forecasting qualities of the proposed model vis-à-vis those used in the IMF's work. The procedures used are fairly straightforward and include an assessment of the characteristics of the implied errors and of the unbiasedness and efficiency of our growth projections.[33] The tasks involved are part of the usual toolkit for forecast evaluation and have been described in other studies.[34] In fact, the usual problem in judging the performance of the projections prepared by IMF staff is the lack of an alternative model against which to compare these projections. This paper fills this gap by extending the empirical growth literature into the forecasting sphere.[35] A caveat is also in order. The projections carried out provide an assessment as to the average usefulness of the model. The model, however, may do very poorly in some countries.

Our assessment also poses some special challenges. For example, within sample prediction would, by construction, imply zero mean errors. Thus, it is necessary to perform out-of-sample forecasts to assess the model. To this end, forecasts were prepared for the period 1996–2000 based on our model with human capital—period 1981–1995, see Table 5.

As to the values of right-hand-side (RHS) variables used in our out-of-sample forecasts, we opted for classifying these variables into four categories and using pre-defined rules for defining their values. The first category, which represents RHS variables for which the values are **quasi-known**, includes the initial levels of human capital and education as well as a country's fertility rates. For the purpose of out-of-sample prediction, the value of these variables for the period 1996–2000 was determined by their lagged initial values adjusted by the growth rates observed over the previous 5-year period (Table 6). The second category, referred to as **unknown** RHS variables, applies to the internal and external shocks received by a country. We assumed that country-specific characteristics, as reflected in the average shocks of the past fifteen years (1981–1995), are the best proxy for these variables. Implicitly, the assumption is that countries prone to shocks (positive or negative) are likely to

---

[31] There are a total of eight 5-year periods between 1961 and 2000, but this is reduced to seven when the first 5-year period begins after 1961 (and before 1966).

[32] The discussion in this sub-section focuses largely on what has been referred as the basic model. The tables, however, also present results that include the ICRG index and also our proxy for exchange rate policy, referred in the rest of this paper as the BMP index.

[33] Errors are defined as the difference between the predicted and actual 5-year growth rates.

[34] See Beach et al. (1999) and Musso-Phillips (2002) for an evaluation of 1-year ahead WEO projections. Ghosh and Joshi (2003) extend the evaluation to 3- and 5-year ahead projections.

[35] IMF staff do not project per capita growth rates, which are therefore derived by adjusting their real GDP growth forecasts by the actual population growth rates.

remain affected by these growth factors.[36] The third category applies to variables with **stable** values, such as the share of credit to the private sector in total credit and the openness of the economy. These variables are assumed to maintain the average values of the previous 5-year period.[37] This is a sensible assumption given the stability of these variables over time. Finally, the last category refers to indicators of macroeconomic management. This category uses the **projections prepared by Fund country teams** in October 1995 for the 5-year period 1996–2000.[38] This assumption, which we later relax, is needed in our view to make fair comparisons between the model's forecasting qualities and those of IMF staff.[39]

Table 6. Rules for Defining the Values of Right-Hand-Side Variables

Regressor type	Regressor name	Rules
Quasi-known	Initial income level, fertility rates, and level of education	Initial level based on lagged values and lagged growth rates; instrumented.
Unknown	Terms of trade and internal shocks	Average for the period 1981-1995; except TEs for which data for the 1990s is used.
Stable	Trade openness and credit to the private sector	Average over the 5 years preceding the estimation period. Similar for ICRG and BMP.
Fund desk macro assumptions	Inflation, including threshold effects, and fiscal balance	Annual average macro assumptions by Fund economists; t+5, as reported in WEO database for October 1995.

Based on the above rules, the main statistical properties of the model's projections are derived and compared to those of IMF staff, focusing in particular on the accuracy of the model's projections vis-à-vis those of the Fund. [Other statistical properties are presented in Table 7[40]; in particular, the mean error of the model is one-third of the mean error in the IMF

---

[36] Arguably, we are drawing on a country's own negative shocks experience to pass judgment on future risks. However, we derive these indicators based on the 15-year average of terms of trade shocks and internal shocks to ensure that this assumption does not unduly restrict the growth projections for any one country.

[37] The ICRG index and the BMP are assumed to behave as other stable regressors; i.e., their RHS values are defined on the basis of the average over the 5-year period 1991–1995.

[38] The cut-off point of October 1995 projections implies that macroeconomic data available to country teams is preliminary when the projections for 1996–2000 are prepared. In contrast, the model is estimated based on actual data for 1981–1995, potentially biasing the results. However, given that the focus of the paper is on average growth rates over a 5-year period, we suspect our forecasts are not heavily affected by this factor.

[39] Although we control for the effects of macro assumptions on growth, we cannot control for the effects of structural reforms that may be assumed by IMF staff when preparing their projections. We sense, however, that these are less important in the overall outcome.

[40] Other statistical properties examined include: (i) *normality*: while the null hypothesis of normality in the distribution of projection errors is rejected in both the proposed growth model as well as the projections by IMF staff, normality in the model's errors is restored if 2 out of 82 countries are excluded from the projections; (ii) *biasedness*: a test suggests that the

(continued…)

staff forecasts.] Specifically, a comparison of root mean squared errors (RMSE) highlights that the model represents, on average, a 20 percent improvement vis-à-vis the IMF staff forecasts. This conclusion is derived from estimating the Theil inequality coefficient (or Theil U-statistic), which represents the ratio of the root mean square errors of the model's projections relative to those prepared by IMF staff, where a value of less than one implies that the model is more accurate than the projections prepared by IMF staff (Table 8).

The accuracy results for different country groups were consistent with our aggregate findings. For example, the model outperforms the projections by IMF staff in all regions (simple averages).[41] The projections in some regions appear to be particularly weak, though this conclusion needs to be qualified by the fact that in some regions the number of countries with forecasts is small. Similarly, the classification per income group suggests that our model performs better among developing countries.

Three extensions are carried out to assess the robustness of the results described and examine additional characteristics of the model's forecasting qualities.[42] First, the model and its forecasts were re-estimated using secondary school enrollment rates, as this is one of the main sources of loss in country coverage (a list of countries can be found in the appendix).[43] The aim of this robustness check is to examine the possible effects that may result from the selection bias problems. The results suggest that the Theil U-statistic remains in line with the previous results and the conclusions on alternative country groups remain valid (Table 9), with the model's performance deteriorating only marginally in regions with small samples.[44] The bias is also lower (¼ percent) than in the model with Barro-Lee data on human capital.

---

median error of the model projections are unbiased while those of the IMF projection errors are biased—in fact, the model's over-optimistic bias (Table 7) is between one-third and one-half of the staff's bias for the same set of countries and the same period (1996–00); and (iii) *efficiency of forecasts*: a regression of the predicted values on a constant and on actual growth rates and a test on the estimated coefficients adding up to 1 cannot be rejected.

[41] Of course, this result does not imply that the model is better in individual countries, but that on aggregate it performs better.

[42] One-year ahead forecasts were also prepared using a set of rules similar to that used for our 5-year out-of-sample projections. These projections outperform those of IMF staff only among countries in the Africa region. This result corresponds to the overly optimistic bias of staff projections for these countries, which has frequently been noted in the WEO.

[43] The two main culprits for the loss of country coverage in our out-of-sample projections are the Summers and Heston Penn data and Barro and Lee data on educational attainment. Replacing the latter by data on secondary school enrollment rates and re-estimating the model increases the country coverage by 34 countries (a 41 percent increase in the number of countries for which projections are prepared) without affecting the results thus far described.

[44] The number of countries in each region or income group, and even the total number of countries, might differ slightly as a result of differences in data availability between the values of RHS variables derived using the rules in Table 6 and their actual values.

Table 7. Biases of Model and Fund Projections

	Basic model with human capital	Model w/ ICRG and w/o BMP	Model w/o ICRG and w/ BMP
Number of countries	82	73	63
Bias/Error (in %)			
Model	0.73	0.49	0.55
Fund desks	1.39	1.50	1.57
Share	53	33	35
Correlation between actual and projected per capita growth rates			
Model	0.31	0.34	0.29
Fund desks	0.18	0.17	0.08

Table 8. Theil U Statistic: Model and Fund Desk with Barro-Lee Human Capital Data

	Basic model with human capital		Model w/ ICRG and w/o BMP		Model w/o ICRG and w/ BMP	
	Number of countries	Value	Number of countries	Value	Number of countries	Value
Total	82	0.83	73	0.79	63	0.77
REGIONS						
North & South America	24	0.97	22	0.95	20	0.87
West and Central Europe	13	0.93	12	0.84	11	0.87
Russia & CIS countries	1	0.30	1	0.20	1	0.06
Middle East	6	0.61	6	0.53	3	0.45
Africa	23	0.73	18	0.70	14	0.72
Asia & Pacific countries	15	1.00	14	0.90	14	0.79
INCOME LEVELS						
High-income	14	1.35	13	1.33	12	1.31
Upper middle-income	16	0.88	13	0.87	12	0.81
Lower middle-income	26	0.87	26	0.76	23	0.78
Low-income	26	0.76	21	0.74	16	0.71

Second, the model's projections with secondary school enrollment rates were modified by using the actual values for each of the categories of RHS variables. The purpose of this exercise is to assess the rules we have specified and, in particular, the macro assumptions for inflation and the fiscal balance used by Fund staff. Worth noting is that using the actual values of macroeconomic variables leads to the elimination of the model's bias (from 0.36 to -0.01) and to a 10 percentage point decline in the standard errors (from 0.85 to 0.74) relative to the forecasts based on the rules in Table 6. Given that our model's growth projections had less bias and lower standard errors than those prepared by Fund staff, we attribute the decline in the RMSE to the bias that originates from using the macro projections of Fund staff (i.e., inflation and fiscal balance forecasts). Thus, to some degree the bias in growth projections is driven by the fact that program targets on the fiscal stance and inflation are not met. Yet, as discussed before, this does not account for the full growth bias. The model's projections also improve in most regions, in particular the Africa and Middle East regions.

Using other actual values for the RHS variables, or even assuming no exogenous shocks (last set of columns in Table 9), leads only to marginal improvements. In particular, the model's bias increases only in a few cases when replacing the rules defined in Table 6 by the actual values, and the standard errors, as reflected in the Theil U-statistic, change little. In sum, it

Table 9. Theil U Statistic: Model and Fund Projections with Secondary School Enrollment Rates

	Fully based on rules		Rule-based excl. WEO macro 1/		Rule-based excl. quasi-known 2/		Rule-based excl. unknown 3/		Rule-based excl. stable 4/		Rule-based but NO shocks 5/	
	Value	Number of countries	Value	Number of countries	Value	Number of countries	Value	Number of countries	Value	Number of countries	Value	Number of countries
Total Theil U	0.85	109	0.74	109	0.85	109	0.85	116	0.84	109	0.86	109
Model error	0.36		-0.01		0.48		0.29		0.49		0.75	
Fund desks error	1.37		1.29		1.37		1.19		1.37		1.37	
Share (abs. value)	26		1		35		24		36		55	
	(Theil U statistic and number of countries in each category)											
REGIONS												
North & South America	0.94	25	0.79	24	0.98	25	0.91	25	0.92	25	0.99	25
West and Central Europe	0.91	15	0.84	15	0.84	15	0.99	23	0.89	15	0.94	15
Russia & CIS countries	0.76	7	0.83	7	0.77	7	0.78	7	0.78	7	0.72	7
Middle East	0.89	10	0.45	10	0.89	10	0.88	10	0.86	10	0.94	10
Africa	0.80	36	0.69	37	0.82	36	0.78	35	0.81	36	0.80	36
Asia & Pacific countries	0.91	16	0.94	16	0.92	16	0.95	16	0.91	16	0.95	16
INCOME LEVELS												
High-income	1.22	14	1.08	13	1.08	14	1.28	22	1.18	14	1.16	14
Upper middle-income	0.99	20	0.88	21	0.99	20	0.95	20	1.03	20	0.92	20
Lower middle-income	0.91	32	0.87	31	0.93	32	0.91	32	0.88	32	0.95	32
Low-income	0.77	43	0.64	44	0.78	43	0.77	42	0.77	43	0.8	43

1/ Actual values for inflation and central government fiscal balance are used instead of the WEO assumptions.
2/ Actual values for initial income level, initial human capital and initial fertility rates are used.
3/ Actual values for internal and external shocks.
4/ Actual values for economy's openness and share of private sector credit.
5/ Assumes no exogenous shocks during the projection period.

can be argued that the model's bias and standard errors are driven primarily by the macroeconomic assumptions of Fund staff, and that the rules specified for quasi-known, unknown, and stable RHS variables perform, all things considered, relatively well.

Finally, we examine the model's performance vis-à-vis the IMF staff forecasts based on a classification that distinguishes between countries with and without a Fund arrangement. The results suggest that the bias and standard errors of IMF staff projections are higher when a Fund arrangement exists (Table 10). Specifically, the model's bias is 35 percent of the bias in the projections used in Fund work among countries without a Fund arrangement, but this bias declines to 22 percent among countries with a Fund arrangement. The Theil U-statistic, which averages 0.85 in the projections that are based on the rules for RHS variables specified above (see Table 9), is broken down into 0.90 in countries without a Fund arrangement and to 0.79 in countries with a Fund arrangement. Further analysis of the data shows that this finding arises from the standard errors that result from using IMF staff macro assumptions, which is consistent with the discussion presented in the previous paragraph.

Table 10. Bias and Theil U in Countries
with and without a Fund Arrangement

	Value	Number of countries
Without a Fund-supported program		
Theil U statistic	0.90	54
Model error	0.28	
Fund desks error	0.81	
With Fund-supported program		
Theil U statistic	0.79	55
Model error	0.43	
Fund desks error	1.93	

V. CONCLUDING REMARKS

This paper estimates a growth equation that covers the growth experience of the most recent two decades while capturing both the cross-country and the time-series growth effects that followed the oil shocks of the 1970s. In contrast to the existing literature, which is based on datasets that extend as far back as the early 1960s and end in the late 1980s, our paper focuses on a more recent time period (i.e., 1981–2000). Although this provides for some results that are different from those in the existing growth literature, indicating that some established results do not come up as strongly (e.g., the speed of convergence appears to be lower than suggested by the literature and typical human capital results are less striking than in other studies), it also provides coefficient estimates that are in line with the post-oil shock era. Together with our panel dataset of 5-year periods, this partly sidesteps the concerns raised by Easterly regarding the use of growth equations for forecasting purposes.[45]

The paper's objective, however, is not to address the many problems of the empirical growth literature, but to examine the scope for improving IMF staff projections by making use of cross-country data. While IMF staff have a detailed understanding of country-specific and one-time factors that affect a country's growth, their projections do not take into account the experience of other countries. Although there are obvious trade-offs in going from country-specific knowledge to a cross-country approach, the paper concludes that the informational content of cross-country data might be worth incorporating into medium-term projections.

Specifically, the forecasting qualities of the model outperform, on average, those of Fund staff when using a simple set of rules for out-of-sample forecasts. The model displays:

- a Theil U-statistic in out-of-sample projections that indicates our model's forecasts have, on average, a 20-percent lower RMSE than the projections prepared by Fund staff (same countries and 5-year period; Table 8),

- a bias that is between one-fourth and one-half of the bias in the IMF staff projections when using their 5-year ahead macroeconomic assumptions (Table 7), and

- a RMSE and a bias that improve when actual RHS values are used for the macro variables of the model instead of the IMF staff's 5-year ahead projections (Table 9), suggesting that closer attention should be given to the inflation and fiscal projections prepared by Fund staff.

The results for different country groups are consistent with these aggregate results. The model shows that IMF staff medium-term projections perform poorly in some regions, though in some of these cases the number of observations is small, thus warranting a

---

[45] Easterly's paper (2001a) raises valid concerns regarding the usefulness of growth equations given that the improved policy package in the developing world during the 1980s and 1990s was not mirrored by improvements in growth rates—in fact, per capita growth rates weakened. Our paper deals with this issue by deriving coefficient estimates using panel regressions based on the 5-year periods that follow the oil shocks of the 1970s.

qualified interpretation. However, extensions of the forecasting exercises to increase country coverage (Table 9) confirm that IMF staff medium-term projections in many of these regions can be improved. The results also show that the model outperforms the projections of IMF staff among upper middle-income, lower middle-income, and low-income countries, but performs poorly among the few high-income countries in the sample. The difference in performance between developed and developing countries cautions against replacing the detailed work carried out by country teams in countries where more sophisticated models can be applied. Finally, the model's biases and standard errors vis-à-vis those of Fund staff are smaller among those countries that have a Fund arrangement (Table 10).

In interpreting the results, the limitations of the methodology used should be borne in mind. First, the model still suffers from the criticisms made to other growth equations, such as lack of theoretical foundation and instability of coefficient estimates (see Section II). Second, the results presented constitute evidence for only one out-of-sample 5-year period, suggesting that more research is warranted and that cross-country models need further testing.

Yet, the paper's goals are more modest. We attempt to test the usefulness of growth equations in improving growth forecasts. Given the simplicity of the analytical tools used, the results obtained are likely to represent a lower bound to possible improvements arising from cross-country information. Similarly, it is also clear that the benefits arise largely among developing countries, suggesting that more sophisticated analytical tools should be used when stable economic relationships and data are available—e.g., this is the case in developed countries and among some emerging market economies.

Data Appendix

Our dataset was constructed based on seven databases downloaded mostly from EDSS, the IMF's data sharing system. The econometric program STATA was used to merge these databases using country and year codes. Rules for dropping data were programmed into the merging process. The IMF data was assumed to dominate over the World Bank database whenever both datasets contained the same economic series for any given year.

<u>IMF WEO Database</u>: GDP (capita and levels), fiscal indicators (government size and fiscal balance), investment and consumption data, population, inflation, terms of trade, broad money.

<u>IMF IFS Database</u>: Government and total credit from DMBs. Other monetary statistics.

<u>World Bank WDI Database</u>: GDP (capita), government consumption, openness, inflation, agricultural GDP, fertility rates, secondary school enrollment rates.

<u>Barro-Lee Database</u>: Years of educational attainment of the adult population.

<u>Summers-Heston Database (version PWT 6.1)</u>: Initial level of income in PPP terms.

<u>ICRG Database</u>: Data on institutional factors.

<u>Black Market Exchange Rate Premium Database</u>: Provided by Carmen Reinhart (University of Maryland).

The variables used in our econometric work are described as follows. Variables in **bold** were included in the final specification, which is referred in this paper as the basic model.

A. Dependent Variables

Real GDP Growth per Capita: Growth rates from national accounts data at constant prices.

B. Regressors

(i) State/Initial Conditions

Initial Income Level: Natural logarithm of the initial level of income (in PPP terms); instrumented by its 5-year lagged value.

Income Share to U.S. Income: Share of country income to U.S. income based on PPP values.

Human Capital: Years of schooling (Barro-Lee database); instrumented by its 5-year lagged value.

Secondary Enrollment Rates: Gross rates considering students of all ages; i.e., including students whose age exceeds the officially set age group for that class (e.g. repeaters). Available from the WDI database only since 1970; instrumented by its 5-year lagged value. Capacity to Absorb Technologies: Product of the human capital regressor and a country's initial income level; instrumented by the 5-year lagged values of its components.

(ii) Control Variables

Macroeconomic Management

Inflation Rate: Consumer price inflation rates (CPI). GDP deflators were used whenever CPI data was not available. Annual average of 5-year periods. Defined as LN(abs(inflation)).

Threshold on Inflation Rates (TIR): This regressor is weighted by the number of high inflation occurrences that exist in our 5-year period (implying it takes a minimum value of zero if the inflation rate is always above the threshold in any 5-year period and a maximum value of 1 if it is always below the threshold level). TIR = D * [LN(Inflation)-LN(0.035)] and D acquires a value of 1 if inflation is below the 3.5 percent threshold and 0 otherwise. Several indicators of inflation uncertainty and inflation expectations were modeled but were found not to be statistically significant. Among them:

Variability of Inflation: Standard deviation of annual inflation rates in each country (cross-country effects only).

Uncertainty of Inflation: Standard deviation of annual inflation rates in each country, per period (cross-country and time series effects).

Inflation Expectations: Standard deviation of a n-2 to n+2 series of annual inflation rates (cross-country and time series effects).

Fiscal Balance: Fiscal balance (central government) as a share of GDP.

Exchange Rate Policy/Market Distortions: Black market exchange rate premium. Provided by Ms. Carmen Reinhart (University of Maryland). The exchange rate premium could also be interpreted as a proxy for price distortions in the economy, thus falling in the sphere of structural issues.

Structural Issues

Government Size: Share of central government expenditures in GDP.

Government Consumption: Share of general government consumption in GDP.

Indicator of Openness: An index in the interval [0, 1] constructed from the residuals of a regression of export plus imports as a share of GDP on LN(population). The goal is to control for country size. Large residuals indicate greater openness (and vice versa).

Financial Sector Development: Broad money as a share of GDP.

Financial Sector Depth: Currency as a share of broad money.

Financial Sector Intermediation Role: Share of private sector credit in total credit from deposit money banks (DMBs).

ICRG Index: Index on institutional factors; only the political index component is used.
Private Sector Choice Variables

Fertility Rate: Number of children per woman of child bearing age; instrumented by its 5-year lagged value.

(iii) Environmental Variables

Internal Shocks: Average over each 5-year period of a dummy variable that takes a value of 1 when agricultural production falls by more than 5 percent in any year.

External Shocks: Annual average over each 5-year period of the change in terms of trade.

(iv) Factor Accumulation

Investment: Share of total capital formation in GDP over each 5-year period; instrumented by its 5-year lagged value.

Population Size: Data from the World Bank WDI database. Data in logs.

A1: List of Countries Included in Out-of-Sample Projections
(Basic Model with Secondary School Enrollment Rates; 109 Countries)

Income level	Country	Income level	Country
High-Income	Australia	Lower Middle-Income	Morocco
	Austria	(continued)	Paraguay
	Canada		Peru
	Denmark		Philippines
	Iceland		Romania
	Japan		Russia
	Korea, Rep. of		South Africa
	New Zealand		Sri Lanka
	Norway		Syrian Arab Republic
	Singapore		Thailand
	Slovenia		Tunisia
	Sweden		Turkey
	United Kingdom		
	United States	Low-Income	Armenia
			Bangladesh
Upper Middle-Income	Argentina		Benin
	Barbados		Burkina Faso
	Botswana		Burundi
	Brazil		Cameroon
	Chile		Central African Rep.
	Costa Rica		Chad
	Czech Republic		Comoros
	Estonia		Congo, Dem. Rep. Of
	Hungary		Congo, E4of
	Latvia		Côte d'Ivoire
	Lebanon		Ethiopia
	Lithuania		Gambia
	Malaysia		Ghana
	Mauritius		Guinea
	Mexico		Guinea-Bissau
	Panama		India
	Slovak Republic		Indonesia
	Trinidad and Tobago		Kenya
	Uruguay		Lesotho
	Venezuela		Madagascar
			Malawi
Lower Middle-Income	Albania		Mali
	Algeria		Mauritania
	Belarus		Mozambique
	Belize		Nepal
	Bolivia		Nicaragua
	Bulgaria		Niger
	Cape Verde		Nigeria
	China		Pakistan
	Colombia		Papua New Guinea
	Dominican Republic		Rwanda
	Ecuador		Senegal
	Egypt		Sierra Leone
	El Salvador		Tanzania
	Guatemala		Togo
	Guyana		Uganda
	Honduras		Ukraine
	Iran		Vietnam
	Jamaica		Yemen
	Jordan		Zambia
	Macedonia, FYR		Zimbabwe

REFERENCES

Aghion, Philippe, and Peter Howitt, 1998, *Endogenous Growth Theory* (Cambridge, Massachusetts: MIT Press).

Barro, Robert, 1991, "Economic Growth in a Cross Section of Countries," *The Quarterly Journal of Economics*, Vol. 106, No. 2, pp. 407–43.

———, 1997, *Determinants of Economic Growth: A Cross-Country Empirical Study* (Cambridge, Massachusetts: MIT Press).

———, 1999, "Notes on Growth Accounting," *Journal of Economic Growth*, Vol. 4, pp. 119–37.

———, and Gary Becker, 1989, "Fertility Choice in a Model of Economic Growth," *Econometrica*, Vol. 57, No. 2, pp. 481–501.

———, and Xavier Sala-i-Martin, 1995, *Economic Growth* (New York, New York: McGraw-Hill).

Beach, William, Aaron Schavoy, and Isabel Isidro, 1999, *How Reliable Are IMF Economic Forecasts?* (Washington, D.C.: The Heritage Foundation).

Ben-David, D., and D. Papell, 1997, "Slowdowns and Meltdowns: Postwar Growth Evidence from 74 Countries," NBER Working Paper 6266 (Cambridge, Massachusetts: NBER).

Easterly, William, M. Kremer, L. Pritchett, and L. Summers, 1993, "Good Policy or Good Luck: Country Growth Performance and Temporary Shocks," *Journal of Monetary Economics,* Vol. 32, pp. 459–83.

———, 2001a, "The Lost Decades: Developing Countries' Stagnation in Spite of Policy Reform, 1980–1998," *Journal of Economic Growth*, Vol. 6, No. 2, pp. 135–57.

———, 2001b, *The Elusive Quest for Growth: Economists' Adventures and Misadventures in the Tropics* (Cambridge, Massachusetts: MIT Press).

Fischer, Stanley, 1993, "The Role of Macroeconomic Factors in Growth," *Journal of Monetary Economics*, Vol. 32, No. 2, pp. 485–512.

Ghosh, Atish, and B. Joshi, 2003, "How Good (or Bad) are Fund Projections?" (unpublished; Washington, DC: International Monetary Fund).

———, and S. Phillips, 1998, "Inflation, Disinflation, and Growth," IMF Working Paper 98/68 (Washington, DC: International Monetary Fund).

International Country Risk Guide, 2003, *Brief Guide to the Ratings System of the International Country Risk Guide* (Chicago, Illinois: PRS Group).

Keefer, Philip, and S. Knack, 1995, "Institutions and Economic Performance: Cross-Country Tests Using Alternative Institutional Measures," *Economics and Politics*, Vol. 7, pp. 207–27.

King, Robert, and R. Levine, 1993, "Finance and Growth: Schumpeter Might Be Right," *The Quarterly Journal of Economics*, Vol. 108, No. 3, pp. 717–37.

Kochhar, Kalpana, and S. Coorey, 1999, "Economic Growth: What Has Been Achieved and How?" in Hugh Bredenkamp and Susan Schadler (editors), *Economic Adjustment and Reform in Low-Income Countries* (Washington, DC: International Monetary Fund).

Levine, Ross and D. Renelt, 1992, "A Sensitivity Analysis of Cross Country Growth Regressions," *American Economic Review*, Vol. 82, No. 4, pp. 942–63.

Mankiw, N., D. Romer, and D. Weil, 1992, "A Contribution to the Empirics of Economic Growth," *The Quarterly Journal of Economics*, pp. 407–37.

Mauro, Paolo, 1995, "Corruption and Growth," *The Quarterly Journal of Economics*, Vol. 110, No. 3, pp. 681–712.

Musso, Alberto, and Steven Phillips, 2002, "Comparing Projections and Outcomes of IMF Supported Programs," *IMF Staff Papers*, Vol. 49, No. 1.

Pritchett, Lant, 1998, "Patterns of Economic Growth: Hills, Plateaus, Mountains, and Plains," World Bank Working Paper 1947 (Washington, DC: The World Bank).

Reinhart, C., and K. Rogoff, 2002, "The Modern History of Exchange Rate Arrangements: A Reinterpretation," NBER Working Paper 8963 (Cambridge, Massachusetts: NBER).

Sachs, Jeffrey, and Andrew Warner, 1995, "Economic Reform and the Process of Global Integration," *Brookings Papers on Economic Activity*, pp. 1–118.

Sala-i-Martin, Xavier, 1996, "I Just Ran Four Million Regressions", NBER Working Paper 6252 (Cambridge, Massachusetts: NBER).

_____, 2002, "15 Years of New Growth Economics: What Have We Learnt?," Columbia University Discussion Paper 102-47 (New York, New York: Columbia University).

Stokey, Nancy, 1994, "Comment on Barro and Lee's Sources of Economic Growth," *Carnegie-Rochester Conference Series on Public Policy*, Vol. 40, pp. 47–57.

Tavares, Jose, and Romain Wacziarg, 2001, "How democracy affects growth," *European Economic Review*, Vol. 45, pp. 1341–78.

Temple, Jonathan, 1999, "The New Growth Evidence," *Journal of Economic Literature*, Vol. 37, pp. 112–56.

Wacziarg, Romain, 2002, "Review of Easterly's 'The Elusive Quest for Growth," *Journal of Economic Literature*, Vol. 40, pp. 907–18.