Macroeconomic Stability and the Distribution of Growth Rates

Vatcharin Sirimaneetham
The World Bank, 30th Floor, Siam Tower
989 Rama 1 Road, Bangkok 10330, Thailand

Jonathan Temple†
Department of Economics, University of Bristol
8 Woodland Road, Bristol BS8 1TN, UK

October 3, 2008

Abstract
We examine the view that macroeconomic instability can form a binding constraint on economic growth. Using a new index of stability, we show that developing countries can be divided into two distinct growth regimes, depending on a threshold level of stability. For the relatively stable group of countries, the output benefits of investment are greater, conditional convergence is faster, and measures of institutional quality have more explanatory power, suggesting that instability forms a binding constraint for the less stable group. We also show that macroeconomic stability dominates other candidates for identifying distinct growth regimes.

JEL classifications: O23, O40
Keywords: macroeconomic stability, economic growth, threshold estimation, Bayesian Model Averaging

*We are grateful to the editor, three anonymous referees, David Ashton, Holger Breinlich, Edmund Cannon, Jan Fidrmuc, and Andreas Leukert for comments and discussion that have improved the paper, together with seminar participants at Bristol, and a Brunel University workshop on Aspects of Growth and Macroeconomic Policy. The usual disclaimer applies. Temple thanks the Leverhulme Trust for financial support under the Philip Leverhulme Prize Fellowship scheme.

†Corresponding author. Email jon.temple@bristol.ac.uk
1 Introduction

It is widely believed that economic growth requires macroeconomic stability. At the broadest level, stability could help to explain the sustained growth of East Asian countries between the early 1960s and the late 1990s. In contrast, sub-Saharan Africa and Latin America have often endured both macroeconomic disarray and slow growth. Economic mismanagement could also help to explain why some developing countries became heavily indebted, in which case the relatively slow growth of the 1980s and 1990s might be attributed to the macroeconomic policies of earlier periods.

Although macroeconomic stability could be important for growth, the strength of the empirical relationship remains uncertain. One argument is that any observed correlation between stability and growth is mainly due to a small number of countries with the very worst macroeconomic outcomes. Once a certain threshold level of stability has been achieved, the marginal benefits of additional stability could be minimal. Another argument, which dates back to at least Sala-i-Martin (1991), is that macroeconomic disarray could be a symptom of deeper problems. Recent research, especially after the work of Acemoglu, Johnson and Robinson (2001), Acemoglu et al. (2003) and Easterly and Levine (2003), argues that macroeconomic policies lack explanatory power relative to institutions.

This paper revisits the growth effects of macroeconomic stability. This is well-worked ground, and a new paper on this topic must try hard to justify its existence. One innovation is that we construct a new composite index of the extent of macroeconomic stability. A more fundamental aim of the paper, however, is to sharpen the link between statistical modeling and informal commentary on policy and growth. Much of that commentary reduces to a simple idea: sound policy is a necessary but not sufficient condition for rapid growth, while bad policy may often be a sufficient condition for slow growth. Perhaps growth performance is only as strong as the weakest link in a set of policy outcomes.

This view conceives growth in terms of the absence or presence of binding constraints, as in recent work by Hausmann, Rodrik and Velasco (2008) and Rodrik (2007). It contrasts with the linear regressions usually adopted in the empirical growth literature, which implicitly assume that different growth determinants can smoothly substitute for one another. Only rarely have empirical studies allowed for the distinctive implications of binding constraints, in which failure in one dimension may limit the growth benefits of favorable characteristics elsewhere.
To take such effects into account, we adopt a number of approaches. Initially, we use direct comparisons of growth rate distributions, where countries are classed into groups according to our new policy indicator. We find that, looking at unconditional distributions, macroeconomic instability is not always a binding constraint. In particular, even when a country ranks low in terms of macroeconomic stability, this is not a sufficient condition for slow growth. We do find, however, that the highest long-run growth rates are confined to countries with stable macroeconomic outcomes.

Next, we examine how regressions can accommodate the binding constraints view. We first use standard regressions to quantify the effects of macroeconomic stability over the period 1970-99, restricting the sample to developing countries. But these linear models implicitly assume that adverse effects of instability can be offset by other factors, rather than allowing instability to form a binding constraint. So we also experiment with threshold estimation, using the methods of Hansen (1996, 2000). We find evidence that our index of macroeconomic stability can be used to split the sample into two groups. In the relatively stable group of countries, the elasticity of output with respect to the investment rate is relatively high, conditional convergence is relatively rapid, and the standard growth determinants of the Solow model (together with a measure of institutional quality) can explain between 75% and 90% of the cross-section variation in growth rates, a remarkably high proportion. In the less stable group, instability reduces growth, while the Solow variables have less explanatory power, investment is less effective, and the residual variance is much higher. We also show that fundamentals such as good institutions are not strongly associated with growth unless macroeconomic stability is also in place. These results suggest that instability can indeed form a binding constraint on growth.

Our work acknowledges an important criticism of past research, which is that policy outcomes are likely to be endogenous in both an economic and a statistical sense. Rodrik (2005) points out that observed policies are decision variables that must be endogenous to social and economic circumstances. The implication for our analysis is that macroeconomic stability is not randomly assigned, and will almost certainly be correlated with omitted country characteristics, and hence with the error term of the growth regression. In the microeconometric literature, the availability of control variables is often limited, but there may be plausible candidates for instrumental variables through ‘natural experiments’. Here, we face a slightly different problem: there are many possible control variables, but relatively few plausible candidates for instruments. We try two different approaches. The first,
to which we give less emphasis, is to follow Barro (1996) in exploiting the observed association between French colonial heritage and macroeconomic stability, linked to the membership of many former French colonies in the CFA Franc zone. This implies that French colonial heritage could be a suitable instrument, but it would not be difficult to criticise the necessary exclusion restriction. For example, French colonial heritage is likely to have influenced the legal system, a debate reviewed in La Porta, Lopez-de-Silanes and Shleifer (2008).

We therefore give more emphasis to an alternative approach, which is to consider an unusually wide range of possible control variables, including various indicators of geographic characteristics and institutions. Our comprehensive approach means that we have some chance of finding controls that influence the extent of stability. In that case, the correlation between macroeconomic stability and the error term will be lessened, even though macroeconomic stability is not randomly assigned. This is essentially the idea of “selection-on-observables” from the treatment effects literature, and will be appropriate if the central endogeneity problem is omitted variables rather than simultaneity bias.1 Our approach is based on Bayesian methods for model averaging, and hence can address the model uncertainty problem highlighted by Levine and Renelt (1992). The evidence that stability matters varies with the sample of countries, but in the largest sample we consider, the benefits of stability are robust across a wide range of specifications.

Finally, we use our results to construct counterfactual distributions of growth rates and steady-state levels of GDP per capita. We can then see what might have happened, had all developing countries achieved macroeconomic stability throughout the last thirty years of the twentieth century. To the extent that the estimated role of stability can be interpreted as a causal effect, the variation in stability exerts a major influence on the distributions of growth rates and steady-state GDP per capita.

The rest of the paper is organized as follows. Section 2 briefly reviews the existing literature on policy and growth, and then discusses the empirical analysis of binding constraints in more detail. Section 3 will describe our new measure. Section 4 looks at the relationship between stability and growth in a variety of ways, with an emphasis on threshold estimation. Section 5 examines robustness using Bayesian methods. Section 6 uses the core

---

1Simultaneity bias will be relevant if policy outcomes are directly dependent on growth outcomes, which may be plausible in the short run, but less so over the thirty-year period that we consider. It is more plausible that growth and policy are jointly influenced by other variables, such as institutions, hence our emphasis on the omitted variable problem rather than simultaneity.
growth regressions to generate counterfactual distributions of growth rates and steady-state levels of income. Finally, section 7 concludes.

2 Relation to existing literature

The literature on policy and growth has often studied trade regimes, and more recently, factors such as entry barriers and regulation. It is therefore worth emphasizing that our paper is about macroeconomic stability, and not about market-led development or the Washington Consensus. As initially summarized by Williamson (1990), the Consensus reflected principles that went well beyond macroeconomic policies. These included tax reform, financial liberalization, liberalized trade policy, openness to foreign direct investment, privatization, deregulation, and the protection of private property rights. Rather than investigate these, we examine a narrower question, whether the Washington Consensus was right to emphasize the benefits of stable macroeconomic outcomes. Attempts to promote stable outcomes can be controversial in practice, not least when reductions in fiscal deficits are involved. Moreover, it is rarely clear how much stability is “enough”.  

Motivated by these considerations, empirical studies such as Bleaney (1996) and Fischer (1991, 1993) investigated the role of macroeconomic stability, and concluded that it matters for sustained growth. In recent years, however, it has been argued that the growth benefits were sometimes exaggerated. Macroeconomic policies have generally improved over time, whereas many developing countries grew more slowly in the 1980s and 1990s than previously. This has led to the conclusion that the growth dividend of greater macroeconomic stability has been disappointing, an argument reviewed in Montiel and Servén (2006). The reasons for the post-1980 growth collapse in developing countries are discussed in Easterly (2001b) and Rodrik (1999) and seem likely to go beyond macroeconomic policy decisions.

Other evidence casts further doubt on the role of stability. Improvements in policy indicators explain relatively few growth accelerations (Hausmann, Pritchett and Rodrik 2005) and in general policy indicators are far more persistent than growth rates, suggesting that policy will usually leave the medium-run variation in growth unexplained (Easterly et al. 1993). Perhaps most fundamentally, empirical studies such as Easterly and Levine (2003)

---

2 In this respect, it is worth noting that our paper does not address the subtler and much more difficult questions that relate to short-run policy activism such as demand management. Our results concern macroeconomic outcomes (rather than policies) assessed over the long run and should be interpreted in that light; they do not imply, for example, that budget deficits must always be avoided.
have found that growth and policy variables are not robustly correlated in the cross-country data when controlling for institutional development. The survey by Easterly (2005) concludes that “the long run effect of policies on development is difficult to discern once you also control for institutions” (page 1055). This highlights a problem for the empirical literature, which is that economic disarray usually extends across a range of outcomes. It may be hard to disentangle the effects of specific macroeconomic outcomes from one another, and from other growth determinants. Perhaps bad macroeconomic outcomes are best seen as symptoms of deeper underlying problems, including institutional weaknesses.

Although some claims for the importance of policy may have been exaggerated, there is a common-sense view which commands wide support. There is likely to be a threshold level in the quality of macroeconomic management, below which growth becomes difficult or impossible. A clear and persuasive exposition of this view can be found in Easterly (2001a). He indicates that governments may not be able to initiate growth, but can certainly destroy any prospect of growth if macroeconomic policies are bad enough. He illustrates the consequences of policy errors using several historical examples, showing that the worst policy outcomes - hyperinflation, high black market premia, large budget deficits - are typically associated with slow growth or even collapses in output. None of this implies, however, that getting macroeconomic policy right is a sufficient condition for rapid growth. It is not difficult to find examples of countries with sound macroeconomic policies and slow growth, with Bolivia in the 1990s as just one example, discussed in Kaufmann et al. (2003).

The common-sense view has come to dominate recent assessments of the role of policy, but is yet to have much impact on the empirical literature. Traditionally, cross-country research on policy and growth has been dominated by simple linear models of the form

\[ g = \eta + \alpha P + \beta' Z + \varepsilon \]  

(1)

where \( g \) is the growth rate, \( P \) indicates the quality of macroeconomic policy, \( Z \) is a vector of other growth determinants, \( \eta \) and \( \alpha \) are parameters and \( \beta \) is a parameter vector. This linear specification assumes that bad policies can be offset by other factors or, put differently, the different variables can smoothly substitute for one another. Yet many informal accounts of growth are phrased in terms of necessary conditions, which cannot be captured by a linear regression of the form above. There are relatively few papers which consider necessary conditions in a formal way, with the notable exceptions of

A simple way to address this hypothesis is to examine the unconditional distribution of growth rates. If macroeconomic instability can form a binding constraint, we should expect to see that unstable countries have growth rates that are tightly distributed around a low mean, because instability is a sufficient condition for slow growth. In contrast, for stable countries, we should observe wider dispersion in growth rates around a higher mean. The wide dispersion arises because stable countries may not have other growth preconditions in place, leading to variation in performance across these countries, driven by variation in other growth determinants. If we divide countries into groups, the distributions of growth rates across countries might look like the hypothetical example in the first panel of figure 1: the solid line represents a possible distribution for unstable countries, while the dashed line represents a possible distribution for stable countries.

The binding constraints view also has implications for the specification of empirical growth models. One way to capture the idea is a simple nonlinear model with two regimes:

\[
\begin{align*}
g &= \eta_1 + \varepsilon_1 & \text{if } P \leq \gamma \\
g &= \eta_2 + \alpha P + \beta'Z + \varepsilon_2 & \text{if } P > \gamma
\end{align*}
\]  

(2)

where again \(P\) is a policy indicator, \(Z\) a vector of other growth determinants, \(\eta_1, \eta_2\) and \(\alpha\) are parameters and \(\beta\) is a parameter vector. The model implies that if the policy indicator \(P\) falls below a threshold value \(\gamma\), governments effectively destroy any prospect of growth, given a low value of \(\eta_1\) and a low variance of the error term \(\varepsilon_1\), and regardless of other country characteristics. In section 4 of the paper, we will use the methods of Hansen (1996, 2000) to estimate more general versions of (2), and show that macroeconomic stability appears to be a more important threshold variable than other candidates, such as geographic and institutional variables.

The analysis in this paper is based solely on cross-section variation. For the present research question, this has some advantages over a panel analysis. One drawback of panel data is that short-run deterioration in policy outcomes may be associated with a short-term growth slowdown, even if macroeconomic stability and growth are not associated in the long run (Bruno and Easterly 1998). The panel data approach could easily capture these short-run responses rather than a genuine long-run effect. Moreover, the cross-section variation may be more informative than panel data about the effects of the ex ante prospects for stability, since a panel data analysis
could be driven mainly by the effects of the realizations of outcomes. Given the spans of data currently available, there is a case for using cross-section data to identify the long-run effects of macroeconomic stability. To the extent that stability and growth are strongly associated in the international cross-section, that would shift the burden of proof in the debate, placing new demands on those who argue that macroeconomic stability is largely irrelevant.

3 Measuring macroeconomic stability

In this section, we introduce our new measure of macroeconomic stability. We combine several indicators in order to measure the average extent of stability over thirty years, 1970-99. The use of a combination has several advantages. From a statistical point of view, it will tend to lessen the outlier problems associated with skewed distributions. From an economic point of view, the new index aims to capture an underlying latent variable, the quality of the macroeconomic decision-making process, rather than seeking to rely on more specific ‘symptoms’ like high inflation. Using several proxies for this latent variable helps to limit measurement error and makes sense if, as suggested by Sala-i-Martin (1991), macroeconomic disarray is associated with undesirable outcomes across a range of indicators. Our approach acknowledges that it will be difficult to identify the separate effects of fiscal discipline, inflation control and exchange rate management in small cross-country data sets. Instead, it makes sense to reduce the dimensions of the problem and focus on a single index of policy outcomes. Arguably, there is more hope of answering questions about policy outcomes and growth if the relevant hypotheses are deliberately characterized in broad terms, given the limitations of the available data.

Our composite measure is based on fiscal discipline, inflation and exchange rate management. Our preferred index is based on an outlier-robust version of principal components analysis, using the minimum covariance determinant method of Rousseeuw (1984). Our empirical analysis will focus on developing countries with available data, but excluding transition economies, and also excluding countries with small populations (where the 1970 population size was below 250,000). There are 78 countries in the sample we use to construct our main indicators; data availability means that our later growth regressions will use between 60 and 70 observations, while the Bayesian model averaging in section 5 will be carried out mainly on a sample of 72 observations.
We now describe the individual policy indicators. To capture fiscal discipline, we use data on the average central government budget surplus as a share of GDP \((\text{SURPLUS})\) over 1970-99.\(^3\) Some countries, notably Guyana and Sudan, have extreme negative values for this variable, reflecting persistently high budget deficits. Our principal components analysis, and hence our later results, are robust to excluding these countries, or replacing \(\text{SURPLUS}\) with the monotonic but bounded transformation, \(\arctan(\text{SURPLUS})\).\(^4\)

To measure success in keeping inflation low, we construct a variable \(\text{INFLA}\). This is the natural logarithm of one plus the median inflation rate over 1970-99, computed from the GDP deflator. We use the median inflation rate to capture success in keeping inflation low on average. Relative to the more common use of the mean, this measure is less at risk of being dominated by a small number of short-lived hyperinflations.

We measure exchange rate management in three different ways. These are the black market premium \(\text{(BMP)}\), an index of currency overvaluation or real exchange rate distortion \(\text{(OVERVALU)}\) and a measure of the variability in exchange rate distortions \(\text{(ERATE)}\). The black market premium reflects departures of an illegal, market-determined exchange rate from the official exchange rate. To lessen outlier problems, our variable \(\text{BMP}\) is defined as the natural logarithm of one plus the mean value of the black market premium over the period.

The two variables \(\text{OVERVALU}\) and \(\text{ERATE}\) were introduced in Dollar (1992) and, in the case of \(\text{OVERVALU}\), extended forwards and backwards by Easterly and Sewadeh (2002). The first of these measures is based on evaluating price levels in a common currency, after correcting for the possible effects of factor endowments on the prices of non-tradeables. This correction is achieved by using the component of price levels that is orthogonal to GDP per capita and its square, population density and two regional dummies. If a country’s price level is higher than predicted by these controls, this indicates the domestic prices for tradeables may be relatively high, and so high values of \(\text{OVERVALU}\) could indicate some combination of real overvaluation and

\(^3\)We also experimented with including the stock of central government debt relative to GDP, but \(\text{SURPLUS}\) is available for more countries.

\(^4\)This transformation is a natural choice, given that the variable is a ratio which can take on extreme values in either direction, positive or negative. The \(\arctan(x)\) function maps \(x\) into the smallest or most basic angle with tangent \(x\). When the angle is expressed in radians, the values of the \(\arctan\) function will be restricted to the interval \((-\pi/2, \pi/2)\) and this will limit the effect of outlying observations. When the transformation is applied to \(\text{SURPLUS}\), the lowest value is less than one standard deviation below the mean, compared to five standard deviations below in the raw data.
trade restrictions. We discuss its interpretation in the appendix.

ERATE is Dollar’s measure of variability in the overvaluation index for 1976-85 (see Table A1 in his 1992 paper). It can be seen as measuring instability in exchange rate management. Given the likely role of inflation in generating movements in the overvaluation index, it may also be an indicator of more general forms of macroeconomic instability (Rodriguez and Rodrik 2000).

Although sometimes we will use the five outcome indicators individually, for most of our analysis we aggregate them into a composite index. The best-known policy index in the recent literature is that of Burnside and Dollar (2000). They construct an aggregate measure of policy quality based on three indicators: inflation, the budget surplus and the Sachs and Warner (1995) indicator of openness to trade. Since their central focus is a possible interaction between the growth effects of aid and the quality of policy, they weight the policy indicators using the coefficients in a simple regression of growth on the indicators, and controls including initial GDP, regional dummies and proxies for political stability. Note that this procedure is not as well-suited to the aims of the current paper. In their procedure, growth will typically be correlated with the aggregate policy index by construction. Here we want to compare distributions of growth rates across countries with good and bad policies, and this requires a composite policy index that makes no use of information on growth rates.

We aggregate the five separate variables using a principal components analysis. First of all, we check that the correlations between these variables are high enough to justify using principal components: in the extreme case, where the variables were all pairwise uncorrelated, a principal components analysis would not make sense. Testing for this “sphericity” case, allowing for sampling variability in the correlations, is a standard problem in multivariate analysis. A likelihood ratio test comfortably rejects sphericity at the 1% level; see the Supplemental Appendix for more details.

We always normalize the first principal component in such a way that high values indicate good policy. The structure of our first index can be seen in Table 1, which shows the correlations between the policy indicators and the first two principal components. In terms of standardized indicators

\footnote{They also experiment with the use of government consumption as a share of GDP, but find this variable to be negatively correlated with the budget surplus, and insignificant when the budget surplus is included. See Burnside and Dollar (2000, p. 850).}
(all with mean zero and unit variance) we can write our first index as

\[
MACRO = 0.334 \times SURPLUS - 0.447 \times INFLA - 0.585 \times BMP \\
- 0.347 \times OVERVAL - 0.475 \times ERATE
\]

This index places most weight on the black market premium and the Dollar (1992) measure of variability in exchange rate distortions. The first principal component explains 42 per cent of the total variance in the standardized data. According to this index, the governments that were most successful in achieving macroeconomic stability during 1970-1999 were Singapore, Thailand, Malaysia, Panama and Benin. In contrast, the analysis suggests that Nicaragua, Guyana, Sudan, Uganda and Zambia were characterized by long-term instability.

A drawback of principal components analysis, especially in a small sample, is the inherent sensitivity to outlying observations. As Hubert et al. (2005) note, a classical principal components analysis maximizes the variance and decomposes the covariance matrix, and both the variance and the covariance matrix can be highly sensitive to outliers. This is an important concern when aggregating measures of macroeconomic policy. Easterly (2005) points out that the empirical distributions of policy indicators are often heavily skewed, with a small number of countries experiencing policies that are unusually bad (several standard deviations from the mean) relative to other developing countries.

For this reason, our main index is based on an outlier-robust principal components analysis. Since we have relatively few dimensions, we can implement the minimum covariance determinant (MCD) method. This is based on identifying the particular subset of \( h < n \) observations, among the many possible subsets of the total set of \( n \) observations, for which the classical covariance matrix has the smallest determinant (a method due to Rousseeuw 1984, p. 877; see also Rousseeuw and van Driessen 1999). We can then use the covariance matrix for just these \( h \) observations to represent the associations among the variables, and to compute the eigenvectors associated with the principal components. We use the standard choice \( h = 0.75n \) so that the method effectively discards the least representative 25% of the cases in estimating the correlations, building in a high degree of robustness.\(^6\)

Using this approach to estimating correlations, we can extract outlier-robust principal components. The correlations between the first two of these

\(^6\)We use the \texttt{ROBPCA} program to implement the MCD approach. Note that the simpler alternative of identifying outliers from bivariate scatter plots is flawed, because it will not always detect observations that are outliers in a multidimensional space.
new principal components and the individual policy indicators are shown in column (2) of Table 1. In terms of loadings on the individual variables, the robust index can be written as:

\[
RMacro = 0.101 \times SURPLUS' - 0.578 \times INFLA' - 0.693 \times BMP' - 0.219 \times OVERVALU' - 0.357 \times ERate'
\]

where each variable has been centred using a robust estimate of its location. Relative to the classical PCA, the outlier-robust PCA places less weight on \textit{SURPLUS}, \textit{OVERVALU} and \textit{ERATE}, and more weight on \textit{INFLA} and \textit{BMP}. Although the weights in the two cases may look rather different, the simple correlation between \textit{MACRO} and \textit{RMACRO} is 0.98, reflecting high correlations between some of the individual components. Using the \textit{RMACRO} index, the five best performers are Singapore, Thailand, Panama, Malaysia and Togo, and the five worst performers are Nicaragua, Uganda, Ghana, Argentina and the Democratic Republic of Congo.

An alternative approach would be to use the diagnostic plot suggested by Hubert et al. (2005). Using this plot we can identify possible outliers, which are then excluded from an otherwise standard principal components analysis. This method indicated that Guyana, Nicaragua and Sudan might be anomalous observations. On excluding them, we obtain the results in column (3) of Table 1. The proportion of the variance explained by the first principal component falls slightly, but the correlations between this component and the different indicators are similar to those reported in columns (1) and (2).

It is interesting to note briefly the correlations between our new measures and those previously used in the literature. Table 2 shows the correlations between \textit{MACRO}, \textit{RMACRO}, the Burnside-Dollar index, and an updated Burnside-Dollar index for 1970-97 due to Easterly, Levine and Roodman (2004). The correlations are sufficiently high that the various indices may be capturing an underlying latent variable. This is the case even though the Burnside-Dollar and Easterly-Levine-Roodman measures use a different weighting strategy and the Sachs-Warner measure of liberal policies, including trade policies.

4 Macroeconomic instability: the weakest link?

We now use our preferred index, \textit{RMACRO}, to examine how growth varies across countries with good and bad policy outcomes. We order the countries
by \textit{RMACRO}, and split the sample at the 33rd and 66th percentiles. This
gives us three groups of countries. We want to investigate how the mean and
standard deviation of the growth rate varies across these groups. We mea-
sure growth as the annual growth rate in GDP per capita (chain-weighted)
over 1970-99, using data from version 6.1 of the Penn World Table, due to
Heston et al. (2002).

In figure 2, we present Tukey box-plots of the growth rate. The upper and
lower limits of the enclosed box correspond to the 75th and 25th percentiles
of the growth rate, while the horizontal line within the box corresponds to
the median. Looking at the first panel of figure 2, we can see that the median
growth rate is substantially lower in the relatively unstable group 1 than in
groups 2 and 3. There is less support for the idea that macoconomic
stability always destroys long-term growth prospects, because even in group
1, the 75th percentile of the growth rate is 1.4%. We find similar patterns
(not shown) if we measure growth in terms of GDP per worker rather than
per capita, and if we classify countries according to \textit{MACRO} rather than
\textit{RMACRO}.

In the second panel of figure 1, we have used kernel density plots to
summarize the same information in a slightly different way.\footnote{The samples
are relatively small to apply these methods, and the choice of bandwidth
becomes important. We discuss this choice in the Supplemental Appendix.}
We exclude the intermediate group, for clarity. The solid line in the figure shows the
distribution of growth rates for the least stable group, while the dashed line
shows the distribution for the relatively stable group. The figure shows that
stable countries have higher growth on average, but also that instability does
not necessarily preclude growth. There is substantial variation in growth
across the countries with unstable outcomes, and a significant fraction of
them display positive growth rates over the thirty-year period. Nevertheless,
there are no countries growing at more than 3.5\% a year in the unstable
group, whereas there are seven countries that grew at least this rapidly
in the stable group (Cyprus, Indonesia, Malaysia, Mauritius, Singapore,
South Korea and Thailand). On this evidence, macroeconomic stability is
a necessary condition for sustaining high growth rates over a long period.

We have examined similar figures (not shown) for all five individual indi-
cators, \textit{SURPLUS, INFLA, BMP, OVERVALU} and \textit{ERATE}. The patterns
in these figures are generally less supportive of the idea that stability pro-
motes growth. This suggests that combining the separate indicators into an
overall index is worthwhile. The evidence that policy matters is strongest
for the Dollar index of exchange rate distortions (\textit{OVERVALU}) and the
black-market premium (BMP). In the second panel of figure 2, we present box-plots for countries grouped according to the black-market premium.

4.1 Growth regressions

In this section, we use growth regressions to examine the relationship between macroeconomic stability and growth. We begin with conventional linear models, and briefly present the results of OLS and 2SLS estimation. Our starting point is the version of the Solow model introduced by Mankiw, Romer and Weil (1992), henceforth MRW. This is arguably the leading structural model to be found in the literature, and reduces the arbitrariness otherwise present in the choice of specification. We estimate the model using data for 1970-1999 rather than the 1960-1985 period used in MRW. We show that, even conditional on the investment rate, population growth, initial income, and regional dummies, there is a significant partial correlation between growth and macroeconomic stability.

In more detail, our specification relates the log difference in GDP per capita to the log of the investment rate, the log of initial GDP per capita, the log of population growth plus 0.05, and a human capital variable, as in MRW. There are two main departures in our specification. First, we use regional dummies to proxy for the initial level of efficiency, as in Temple (1998). Second, we do not use the rate of investment in human capital, but instead a measure of the initial level of educational attainment.\footnote{The use of a stock measure rather than a flow can be justified formally as a proxy for the steady-state level of educational attainment, as in equation (12) in Mankiw, Romer and Weil (1992, p. 418).}

We use the natural logarithm of either the 1970 literacy rate, or average years of schooling in 1970, where the literacy data are from the World Bank's World Development Indicators (2004) and the schooling data are from Barro and Lee (2001). In each case, the figure relates to the population aged 15 and over.

First of all, we look at a growth regression that excludes the policy indicators; this can be seen in column (1) of Table 3. These results show that the MRW regression works well over a different time period; the explanatory power is similar, although the effect of population growth is imprecisely estimated. We also report the elasticity of steady-state income with respect to the investment rate, which is 1.18 in column (1), within the range spanned by the estimates of MRW. In the next column, we look at a regression that includes only initial income, regional dummies and the robust policy index RMACRO. The policy index is significant at the 5% level and the associ-
ation is strong: if interpreted as a causal effect, a one-standard-deviation improvement in stability would have raised the annual growth rate by 0.71 percentage points over this thirty-year time period. In column (3) we control for the effects of investment and population growth, as in MRW. The effect of $RMACRO$ is slightly weaker, as might be expected, but significant at 12%. The reduction in the size of the coefficient indicates that macroeconomic stability raises investment, an idea that we explore later.

In column (4) we add the logarithm of the 1970 literacy rate, which increases the explanatory power of the regression. $RMACRO$ is once again significant at the 5% level. This result is robust to replacing the literacy rate with the logarithm of average years of schooling in 1970, as in column (5). This reduces the size of the sample by 10 observations, so (4) is our preferred specification in what follows.

The partial correlations between growth and macroeconomic stability do not appear to be driven by anomalous observations. The results are robust to the deletion of potential outliers, as identified by least absolute deviation regressions. Our findings are similarly robust to using single-case diagnostics such as DFITS and DFBETA, which identify a similar set of outliers to the LAD method in this case. We have also used added-variable plots (not shown) to identify potential outliers. On excluding Nicaragua and the Democratic Republic of Congo, the results are slightly less strong, in that $RMACRO$ is now significant only at the 8 per cent level. Finally, we also carry out some simple diagnostic tests. These suggest the models do not suffer from omitted non-linearities (based on Ramsey’s RESET test) or heteroskedasticity (based on versions of the Breusch-Pagan and White tests) except in the regression that includes investment but not a measure of human capital (column 3).

We now turn to an instrumental variables approach which exploits the observed association between French colonial heritage and macroeconomic stability, as in Barro (1996). Many former French colonies maintained a fixed exchange rate with the French franc, and this appears to have been associated with lower inflation rates. Our sample contains 15 former French colonies, and for these countries the mean and standard deviation of $RMACRO$ are 0.52 and 0.72 respectively. This compares to a mean of 0.01 and standard deviation of 1.03 for former British colonies and, since $RMACRO$

\footnote{To identify potential outliers, we estimate the models using the LAD estimator, and then define outliers as countries whose LAD residuals are more than two standard deviations from the mean value.}

\footnote{The results are available upon request. See Cook and Uchida (2003, p. 153-54) for a brief discussion of how DFITS and DFBETA are computed and used.}
is standardized, to a mean of zero and unit standard deviation for the sample as a whole.

The results of instrumenting \textit{RMACRO} using a dummy for former French colonies are shown in column (6) of Table 3. We test the significance of \textit{RMACRO} using the Anderson and Rubin (1949) test statistic, which is optimal for models that are just-identified (Moreira 2003, p. 1031) and should be robust to weak-instrument problems. The p-value associated with the test is 0.02, so \textit{RMACRO} is significant even in the 2SLS estimates. Relative to the OLS point estimates, the 2SLS coefficients assign more weight to macroeconomic stability and less to investment. The finding that the 2SLS coefficient for \textit{RMACRO} is considerably higher than the OLS coefficient could be due to measurement error or sampling variability, as Acemoglu et al. (2001) and Frankel and Romer (1999) argue in other contexts. But it could also be due to a failure of the exclusion restriction, and hence we should treat these results cautiously. The small number of observations reinforces this point. As a way to lessen endogeneity problems arising from the non-random assignment of policy, the approach we use in the next section, namely a comprehensive search through a wide range of control variables and specifications, may be preferable.\footnote{Moreover, the applicability of IV approaches to cross-country growth data may have been exaggerated. When the instrument is correlated with the error term, even weakly, the inconsistency of the IV estimator can be worse than that of the OLS estimator, particularly if the instrument is not strongly correlated with the endogenous explanatory variable (see Cameron and Trivedi 2005, section 4.9.2). There are good reasons to doubt many of the exclusion restrictions adopted in the literature, since most candidates for instruments might be correlated with omitted growth determinants; see Durlauf, Johnson and Temple (2005) for more discussion.}

In summary, there is an association between macroeconomic stability and growth, even conditional on investment rates. Taking the results at face value, a one-standard-deviation improvement in stability translates into an annual growth rate that is between 0.5 and 0.7 percentage points higher over a thirty-year period. Increasing the annual growth rate by 0.7 percentage points would leave GDP per capita higher by 23\% at the end of the thirty years. Later in the paper, we examine the implications for the location and shape of the distribution of growth rates, and the steady-state distribution of GDP per capita.

\subsection*{4.2 Threshold estimation}

In the remainder of this section, we adopt the methods of Hansen (1996, 2000) for sample splitting and threshold estimation. We are interested in
estimating nonlinear models of the following type:

\[
g = \eta_1 + \alpha_1 P + \beta_1' Z + \varepsilon_1 \quad \text{if } P \leq \hat{\gamma}
\]

\[
g = \eta_2 + \alpha_2 P + \beta_2' Z + \varepsilon_2 \quad \text{if } P > \hat{\gamma}
\]

where \(\hat{\gamma}\) is a threshold estimated jointly with the other parameters in the model. The variable \(P\) could be an indicator of macroeconomic outcomes, or some other variable, such as a measure of institutional quality. This specification nests our earlier example, (2) above, since the intercept, slope coefficients and error variances are allowed to vary across the two regimes.

A particular strength of the Hansen approach is that we can consider alternative candidates for the threshold variable \(P\) and compare their merits. We can also see whether macroeconomic instability forms a binding constraint on growth, in the sense of limiting the benefits of favourable fundamentals, such as geographic and institutional characteristics.

It is possible to test for the existence of a threshold, and hence multiple regimes, using the Hansen (1996) bootstrapped LM test. Hansen (2000) develops an asymptotic approximation to the least-squares estimate of the threshold \(\hat{\gamma}\) which allows construction of a (possibly asymmetric) confidence interval. These methods can therefore reveal the extent to which a proposed sample split is estimated with precision, and whether the proposed nonlinearity is supported by the data.\(^{12}\)

We will consider seven possible candidates for the threshold variable \(P\), namely \(R MACRO\), and six indicators of either geographic or institutional ‘fundamentals’. We want to see whether differences in macroeconomic stability are associated with distinct growth regimes, or whether fundamentals provide a better way to divide the sample. Two of the fundamentals we consider are standard measures of geographic characteristics. The first variable is the logarithm of the Frankel-Romer (1999) measure of natural openness to trade, which is partly based on proximity to large markets, and denoted \(FR\) in the following. The second variable is absolute latitude (that is, distance from the equator) which is denoted \(ABSLAT\). In both cases, the data are taken from Hall and Jones (1999).

The remaining four candidates for threshold variables are all measures of institutional quality. These are \(GOVKKM\), a composite index of the quality of governance for 1996-2000, due to Kaufmann, Kraay and Mastruzzi (2005); \(POLITY\), the extent of democracy, based on the POLITY IV database of

\(^{12}\)Previous applications of these methods to growth regressions include Hansen (2000) and Papageorgiou (2002). In our emphasis on institutions as a potential threshold variable, the current analysis is especially close to the work of Minier (2007), but we consider the role of macroeconomic stability in more detail.
Marshall and Jaggers (2000), and averaged over 1970-99; POLCON, a measure of the extent of political constraints due to Henisz (2000), again averaged over 1970-1999; and EXPRISK, the measure of average expropriation risk for 1985-95 used in Acemoglu, Johnson and Robinson (2001). It is worth noting that several of these measures are partly based on observed outcomes rather than constraints, and this could lead us to exaggerate the effects of institutions, and understate the effects of macroeconomic stability.\footnote{See Glaeser et al. (2004) on the general desirability of using measures of constraints or rules, rather than measures closely related to outcomes.}

Our procedure is as follows. For each of the six fundamental variables, we estimate a regression which relates growth to that variable, the Solow variables, and RMACRO. We omit regional dummies to avoid overfitting problems when the sample is subdivided. We then test for a threshold associated with RMACRO, and alternatively with the fundamental variable. We summarize the results in Table 4. It is immediately apparent that RMACRO dominates all the other candidates as a threshold variable. In all but one case, the null of no threshold is rejected for RMACRO at the 10\% level, while it is never rejected for any of the other six measures of fundamentals. These results suggest that the data are well described by two regimes, where the classification of countries into the two regimes depends on macroeconomic stability rather than geography or institutions. The estimated threshold for RMACRO is also reported in the table, along with its 95\% confidence interval (which may be asymmetric) and the number of countries in each subsample. Since RMACRO is normalized to have a zero mean and unit standard deviation in the 70-country sample, it is clear that the estimated threshold is precisely estimated and relatively stable across the various specifications summarized in Table 4.

An especially interesting result is that, when the sample is divided using the estimated threshold, the standard growth variables have much higher explanatory power for the relatively stable countries. For this group, the model typically accounts for between 75\% and 90\% of the variation in growth rates, while the $R^2$ for the less stable countries is typically between 40\% and 50\%. This is consistent with our overall message: if macroeconomic stability is achieved, growth is well explained by a standard regression, but the Solow variables (and measures of geographic or institutional fundamentals) have less explanatory power when instability forms a binding constraint on growth, since this switches off the benefits of favorable characteristics. The main departure from our earlier hypothesis is that the cross-section residual variance is higher, not lower, for countries which experience macroeconomic instability.
We show a representative set of results for the two groups in columns (7) and (8) of Table 3. These are based on a model containing the Kaufmann et al. measure (GOVKKM) and hence where the candidate variables for a threshold were GOVKKM and RMACRO. As in the other cases, the Hansen (1996) test favoured the use of macroeconomic stability to split the sample. The estimated threshold for RMACRO, \( \hat{\gamma} = 0.297 \), is slightly above the mean and divides the sample into 42 relatively stable countries (column 7) and 28 unstable countries (column 8). Comparing these two columns, it is clear that macroeconomic stability is associated with a higher elasticity of output with respect to the investment rate, faster conditional convergence, and perhaps stronger growth benefits of good institutions. Overall, the explanatory power of the growth regression is much higher for the stable group, and the specification tests more favourable. In contrast, a Ramsey RESET test for the less stable group rejects the Solow specification.

Across the six specifications used as the basis for Table 4, a less plausible result is that, in the subsamples with relatively stable macroeconomic outcomes, RMACRO is often negatively signed, and sometimes significant at conventional levels, as in column (8) of Table 3. The result that stability has a significantly negative effect in this particular group should be interpreted with caution. It does not arise when the control variable is ABSLAT, POLITY, or EXPRISK. Where a significantly negative relationship does emerge, it may be related to a conditional convergence effect. By construction, all countries in the second regime must have achieved a certain degree of stability, but some may combine instability (relative to other members of the stable group) with strong potential for rapid growth. Simply including initial income as an explanatory variable may not be enough to eliminate such effects. This interpretation of the evidence would be consistent with the idea that, once a certain degree of stability has been achieved, the benefits of greater stability may be limited.\(^{14}\)

Finally, we look at the role of the ‘fundamentals’ (geography and institutions) in more detail. The last two rows of Table 4 report the p-values associated with these variables for the unstable and stable groups of countries respectively. This shows that the posited fundamentals usually lack explanatory power in the less stable countries, but often emerge as signif-

\(^{14}\)We should also emphasize that the difference in signs for RMACRO between the two regimes does not drive our evidence for the existence of a threshold. If we repeat the exercise but remove RMACRO from the models, the p-value for a threshold based on RMACRO is generally similar to our Table 4 results except in the case of GOVKKM, and even there the null of no threshold is still rejected at the 12% level.
icant for the more stable countries. This supports an account in terms of binding constraints.

5 Robustness

In this section, we use Bayesian methods to examine the robustness of the partial correlation between growth and macroeconomic stability. Levine and Renelt (1992) showed that partial correlations in the empirical growth literature may not be robust to changes in specification. This is a serious problem for growth researchers, because the list of candidate predictors is long and it is not easy to rule out particular variables or specifications using prior reasoning. Put differently, there is a model uncertainty problem, and the standard errors in any specific regression will tend to understate the extent of uncertainty about the parameters. In this section, we address this problem using Bayesian model averaging, as in Brock et al. (2003), Durlauf et al. (2008), Fernandez et al. (2001), Malik and Temple (2005), Raftery et al. (1997) and Sala-i-Martin et al. (2004). In what follows we refer to the last of these papers as SDM.\footnote{More recently, Crespo Cuaresma and Doppelhofer (2007) and Eicher et al. (2007) have developed approaches which allow joint consideration of model uncertainty and sample splits or thresholds. The application of these to macroeconomic stability would be an interesting area for further work, although in samples of the present size, it would be important to allow for outliers.}

We discuss the main ideas only briefly, drawing heavily on the original presentation in Raftery (1995). Bayesian approaches treat parameters as random variables, and aim to summarize uncertainty about these parameters in terms of a probability distribution. The natural extension to model uncertainty is to regard the identity of the true model as unknown, and summarize a researcher’s uncertainty about the data generating process in terms of a probability distribution over the model space. By explicitly treating the identity of the true model as inherently unknowable, but assigning probabilities to different models, it is possible to summarize the ‘global’ uncertainty about parameters in a way that incorporates model uncertainty.

We consider the case of $K$ possible models, and assume throughout that one of these models generated the observed data $D$. We denote the models by $M_1...M_K$ and their corresponding parameter vectors by $\theta_k$. The Bayesian approach to model uncertainty is to assign a prior probability to each model, $p(M_k)$, as well as a prior probability distribution $p(\theta_k \mid M_k)$ to the parameters of each model. Using this structure a Bayesian can then carry out inference on a quantity of interest, such as a slope parameter, by using the
full posterior distribution. In the presence of model uncertainty, this distribution is a weighted average of the posterior distributions under all possible models, where the weights are the posterior probabilities that a given model generated the data (for example, Leamer 1978).

To illustrate in the case of just two possible models, the full posterior distribution of a parameter of interest $\Delta$ can be written as:

$$p(\Delta \mid D) = p(\Delta \mid D, M_1)p(M_1 \mid D) + p(\Delta \mid D, M_2)p(M_2 \mid D)$$

where $p(\Delta \mid D, M_k)$ are the conventional posterior distributions obtained under a given model and the terms $p(M_k \mid D)$ are the posterior model probabilities, namely the probability, given a prior and conditional on having observed $D$, that model $M_k$ is the one that generated the data. This approach requires the evaluation of posterior model probabilities. Briefly, as in Raftery (1995), Raftery et al. (1997) and Sala-i-Martin et al. (2004), we use the Bayesian Information Criterion (BIC) to approximate the Bayes factors that are needed to compute the posterior model probabilities. We can then implement a systematic form of model selection, and conduct inference in a way that acknowledges model uncertainty. For example, we can easily investigate the hypothesis that a slope coefficient $\beta_z$ is non-zero, by summing the posterior model probabilities for all models in which $\beta_z \neq 0$; we call this a posterior inclusion probability.

As the list of candidate predictors becomes longer, there quickly comes a point where estimation of all the possible models is not feasible, and attention must be restricted to a subset. We use the approach of Raftery et al. (1997), where a branch-and-bounds search algorithm is used to identify a subset of models with high posterior probability; for a longer discussion and references, see Malik and Temple (2005) and Sirimaneetham and Temple (2006).

We also use the more complex approach of Hoeting et al. (1996). This is because outliers could be a serious problem for our analysis. In general, any procedure for dealing with model uncertainty or model selection may be influenced by outliers. Even if steps are taken to identify these observations, the final results can easily depend on the order in which model selection and outlier detection is carried out. Hoeting et al. (1996) suggest a procedure for addressing this issue. First, the full model, containing all the candidate predictors, is estimated by an outlier-robust method due to Rousseeuw (1984), and the standardized residuals used to identify possible outliers. Next, model averaging is carried out. As in Hoeting et al. (1996), a ‘model’ is now defined in terms of (1) a set of predictors and (2) a set of
observations identified as outliers, where the latter are some or all of those identified in the initial stage. (This restriction on the number of candidate outliers is needed to keep the dimensionality of the problem manageable.) Then a Markov chain Monte Carlo model composition ($MC^3$) approach is used to approximate the posterior model probabilities. For more details, see Hoeting et al. (1996).

Here, we are interested in seeing whether $RMACRO$ is a robust determinant of growth. Our list of candidate predictors is taken from SDM, who seek to explain differences in growth rates over 1960-1996 for 88 countries (developing and developed). We modify their analysis by measuring growth over 1970-99, and replacing their measure of initial GDP for 1960 with a measure for 1970. Despite the change in time period, we can continue to use the same candidate predictors as SDM, since the majority of their explanatory variables were chosen precisely because they are fixed over time or likely to change only slowly. In practice, to keep the application of BMA methods manageable, we focus on the 31 variables in SDM that have a posterior inclusion probability greater than 4% (based on their Table 2, p. 824). It is worth noting that one of these variables is Dollar’s original index of real exchange rate distortions, measured for 1976-85. This has a low posterior inclusion probability, just 8.2%, in the main results of SDM.

With this set of control variables, we can consider the effects of stability at the same time as a wide range of other hypotheses. For example, the SDM variables include several measures related to geographic characteristics: these include the fraction of land area in the tropics, the fraction of population in the tropics, population density, population density in coastal areas in the 1960s, and the prevalence of malaria in the 1960s. Other variables that are included in the SDM data include regional dummies, the relative price of investment goods, life expectancy in 1960, indicators of religion, ethnic diversity, the relative importance of primary exports, and the share of public investment in GDP. The SDM data span a wide range of the hypotheses investigated in the growth literature, and hence our robustness tests are unusually systematic.\(^\text{16}\)

For the purpose of the BMA, and given the high number of candidate predictors, we would like to have as many developing countries in the samp-

\(^{16}\)One change we make relative to SDM is that we sometimes transform some of the explanatory variables to reduce outlier problems. The variables concerned are the relative price of investment goods, population density in coastal areas in 1965, and overall population density in 1960, all of which have highly skewed distributions. In some of our analysis, we use the natural logarithms of these variables, rather than simply entering them in levels.
ple as possible. Our measure $RMACRO$ is available for 78 countries but, when we merge it with the SDM data set, our sample is reduced to just 63 developing countries. In what follows, we extend the country coverage by imputing missing values for a small number of variables in the SDM data. This allows us to increase the number of countries to 72. The decision to impute missing values involves a trade-off: we introduce measurement error, but at the same time bring to bear some additional information, and lessen the biases that arise when data are missing in non-random ways. Here, the number of imputed values in the data matrix for the explanatory variables is just 21, representing less than 1% of the total number of cells.

The evidence that policy has explanatory power is always much stronger in the 72-country sample than in the 63-country sample, as we document in Sirimaneetham and Temple (2006). The reason for this is clear, if we inspect the values of $RMACRO$ for the nine countries that are added in moving to the 72-country sample. These nine countries include four that are in the bottom decile for $RMACRO$ (Guyana, Iran, Nicaragua and Sierra Leone) and three that are in the top two deciles (Cyprus, Chad and Fiji). Hence, in moving to the larger sample, we are increasing the representation of countries at the extreme ends of the distribution of macroeconomic outcomes. This clearly adds identifying variation to the data set. At the same time, we must have considerable faith that policy outcomes and growth are reliably measured for these countries.\footnote{This is related to a more general debate about the appropriate response to ‘good’ and ‘bad’ leverage points, those observations with unusual values for the independent variables; see Temple (2000). Here we concentrate on the 72-country sample, with the caution that it may contain a number of leverage points that will affect inference and the posterior model probabilities.}

We do not report the full BMA results, but instead focus on the posterior inclusion probability associated with $RMACRO$. This is the sum of the posterior model probabilities for all models in which the variable appears. When we combine $RMACRO$ with the 31 variables from SDM, and carry out model averaging, the posterior probability of inclusion is 100%. This implies that $RMACRO$ appears in every model that is assigned non-zero posterior probability (note that the Raftery et al. 1997 procedure effectively rounds posterior probabilities down to zero for the weakest models). The relevant posterior mean - that is, the weighted average of the coefficients on $RMACRO$ across all models, where the weights are the posterior model probabilities - is 0.51. This is close to the estimate found in our earlier growth regressions.

When we use the outlier-robust $MC^3$ approach of Hoeting et al. (1996)
the results are less strong, but still supportive. Dollar’s original index of real exchange rate distortions has a high posterior inclusion probability, 99%. The evidence for a separate effect of $RMACRO$ is weak, but becomes much stronger if we exclude Dollar’s index. The inclusion probability for $RMACRO$ then rises to 69%.

We can also examine whether macroeconomic stability matters, conditional on institutions. To investigate this, we add four measures of institutions to the BMA exercises. These are the same measures we used earlier, namely $GOVKKM$, $POLITY$, $POLCON$ and $EXPRISK$. Initially, we exclude $EXPRISK$ because it reduces the sample of countries substantially. When we add the other three measures to our previous BMA, the posterior inclusion probability of $RMACRO$ is 97.4%. If we use the outlier robust $MC^3$ approach, the inclusion probability of $RMACRO$ falls to 53%. Incidentally, the results strongly support the hypothesis that growth and institutions are highly correlated. The measure $GOVKKM$ dominates the others, with an inclusion probability of 100%. The inclusion probabilities for the extent of democracy ($POLITY$) and political constraints ($POLCON$) never exceed 35%.

When we also include the expropriation risk measure, the sample is reduced to 56 countries. The posterior inclusion probability of $RMACRO$ is very high in this sample (96.8%) while $GOVKKM$ continues to outperform the other measures of institutional quality. The $POLITY$ and $POLCON$ measures have inclusion probabilities in the 40%-50% range, while expropriation risk adds little in terms of explanatory power, with an inclusion probability of just 0.1%.

To summarize, when we allow for a wide range of candidate growth predictors, the evidence that $RMACRO$ matters is sensitive to the inclusion of leverage points, which explains why our results are much stronger for the larger sample based on imputed data. In that sample, there is always a high inclusion probability for either $RMACRO$ or Dollar’s index of real exchange rate distortions. Expressed differently, nearly all the best-performing models include at least one of these variables, regardless of how the rest of the specification varies. We also find some evidence that stability matters even when controlling for institutional quality. This is a demanding test, given that some of these institutional measures are likely to reflect a wide range of outcomes, rather than simply rules and constraints.

$^{18}$To keep the number of candidate predictors manageable, this sometimes requires us to use slightly fewer of the original SDM variables. We then drop those with relatively low posterior inclusion probabilities in the SDM paper.
6 Counterfactual distributions

This section examines the role of macroeconomic stability in broader perspective. One way to assess the effect is to construct a counterfactual distribution, for either growth rates or steady-state levels of income. We can then see what might have happened if all countries had achieved the same level of macroeconomic stability over 1970-99. To do this, we use regression estimates to construct kernel density estimates of counterfactual distributions.\footnote{Kernel density estimates of counterfactual distributions are associated in particular with the work of DiNardo, Fortin and Lemieux (1997) on wage distributions. These methods have also been applied in growth economics, by Desdoigts (1996, 2004).} An advantage of this approach is that we can see where in the distribution the role of stability may have been especially important, information that is not directly apparent from regression estimates.

Whether we look at the counterfactual distribution of income levels or growth rates, it should be noted that the effects - in terms of changes in the location and shape of the distribution - will not be uniform throughout the growth rate distribution. For example, when we look at growth rates, the changes observed in the shape of the counterfactual distribution will depend on the full joint distribution of macroeconomic stability and the growth rate. This is easy to see by considering a hypothetical example. If all countries with intermediate growth rates or better also had stable outcomes, but countries with low growth did not, then imposing macroeconomic stability throughout the sample would only affect the lower end of the distribution. Changes in the growth rate distribution cannot be summarized simply by a set of regression coefficients, and looking at the whole distribution can add useful information.

First of all, we look at actual and counterfactual distributions of growth rates. The basic idea is to work out what each country’s growth rate would have been, had all countries achieved the same level of macroeconomic stability over 1970-99. First of all, we estimate a growth regression similar to those in section 4, relating growth to the Solow variables, regional dummies, and $RMACRO$. The coefficient on $RMACRO$ in this regression is 0.64. We then compute a counterfactual growth rate $g_i^*$ which is equal to

$$g_i^* = g_i + 0.64(M^* - RMACRO_i)$$

where $g_i$ is the observed growth rate, and $M^*$ is the value of $RMACRO$ at the 95th percentile in our sample, corresponding to the value for Malaysia.

The first panel of figure 3 shows the distribution of the actual growth rate (the solid line) and the counterfactual distribution (the dashed line).
This clearly shows how the distribution of growth rates would have shifted to the right if macroeconomic stability had been widely achieved. It is worth noting that the shift takes place throughout the distribution.

This exercise holds the rate of investment constant, but macroeconomic stability may help to raise investment. To examine this, we carry out a growth regression which excludes investment: hence it now measures the overall effect of macroeconomic stability. The second panel of figure 3 is the relevant diagram. Again, the counterfactual distribution lies to the right. As might be expected, the effect of stability is now slightly stronger, and continues to be observed throughout the distribution.

Our growth regressions include initial income, and hence can be seen as modelling the level of the steady-state growth path, as in Mankiw, Romer and Weil (1992). Under the assumption that all countries grow at the same rate in the long-run steady state, we can use the estimated coefficients for 1970-99 to compute the implied steady-state distribution of GDP per capita. Again, we can also construct the counterfactual distribution that would have obtained under universal macroeconomic stability.

The first panel of figure 4 shows the actual and counterfactual steady-state distributions of log GDP per capita. Note that the actual distributions are not necessarily expected to have the familiar ‘twin peaks’ pattern of Quah (1996), because our sample is restricted to developing countries. The figure shows how better policy might have moved the distribution of steady-state income levels rightwards, and the potential magnitude of this effect is clearly substantial. In the second panel of figure 4, we extend the analysis by taking into account the effect of $RMACRO$ on investment. Relative to the first panel, the counterfactual distribution is slightly further to the right, as would be expected if macroeconomic stability is associated with higher investment. Overall, our results indicate that macroeconomic stability could be a major influence on the steady-state distribution of income levels.

7 Conclusions

This paper has examined the relationship between macroeconomic stability and growth in developing countries. The paper introduces a new index of the extent of macroeconomic stability, based on aggregating five policy

---

20To do this, we run a simple regression of the logarithm of the investment rate on initial income, initial human capital, regional dummies and $RMACRO$, and then use this to calculate a (country-specific) counterfactual investment rate that would have obtained if all countries had achieved macroeconomic stability. We can then construct a counterfactual steady-state distribution, shown in the second panel of figure 4.
indicators using an outlier-robust version of principal components analysis. Using this index, we show that growth is positively associated with macroeconomic stability in a sample of 70 developing countries. If this is interpreted as a causal effect, a one-standard-deviation improvement in the index would raise the annual growth rate by somewhere between 0.5 and 0.7 percentage points over a thirty-year period. We have also used Bayesian methods to examine the robustness of this result. Consistent with previous work on this topic, the strength of the evidence depends on the sample of countries. In the largest sample we consider, the effect is generally robust across a range of specifications.

Perhaps more fundamentally, we have also uncovered some evidence of interesting non-linearities, using threshold estimation. Formal tests indicate that our stability index can be used to divide the sample into two groups. In the relatively stable group of countries, investment has a strong effect on output, and the standard growth determinants of the Solow model, together with a measure of institutional quality, can explain between 75% and 90% of the cross-section variation in growth rates. In the less stable group, instability clearly reduces growth, the Solow variables have less explanatory power, investment is less effective, and the residual variance is much higher. We also find that good institutions are not strongly associated with growth unless macroeconomic stability is also in place. These patterns support the common-sense view that some degree of stability is a necessary condition for rapid growth, even when a separate role for institutions is taken into account. Our work suggests that the conclusion of some recent papers, that macroeconomic stability is largely irrelevant, may be premature.

8 Appendix

This appendix briefly discusses the Dollar (1992) measure of exchange rate overvaluation, which can be interpreted in a variety of ways. One issue is whether Dollar’s procedures can reliably control for the determinants of non-tradeables prices. This has been analyzed by Falvey and Gemmell (1998, 1999). They suggest that Dollar’s approach can be a reasonable approximation on average, at least when GDP per capita is a good proxy for relative factor endowments.

Assuming for now that the Dollar procedure is effective in modeling non-tradeables prices, a remaining question is whether differences in tradeables prices reflect trade restrictions or exchange rate policies. Rodriguez and Rodrik (2000) provide an especially useful discussion of the strict assump-
tions that are needed for Dollar’s approach to capture trade restrictions. They argue that international variation in price levels will be partly driven by trade costs, which in turn could reflect geographic characteristics. They show that around half the variation in the original Dollar measure can be explained by a combination of the black market exchange rate premium, regional dummies, and two geographic indicators - one measuring the ratio of coastal length to land area, and the other a dummy for tropical countries. Overall they conclude that the cross-section variation in price levels is likely to be driven by a combination of nominal exchange rate policies and geographic characteristics, rather than variation in trade barriers.

This provides partial support for our own use of the OVERVALU variable as a measure of macroeconomic policy outcomes. With the above discussion in mind, our maintained assumption will be that the cross-section variation in OVERVALU primarily reflects differences in national exchange rate policies. Given that other interpretations are possible, we briefly examine what happens if we omit OVERVALU from the set of indicators developed in section 3 of the paper. If we recalculate the principal components for four indicators rather than five, we obtain the following index:

\[
MACROND = 0.332 \times SURPLUS - 0.516 \times INFLA - 0.615 \times BMP - 0.495 \times ERATE
\]

again in terms of standardized variables. This composite indicator is very highly correlated with our preferred measures MACRO \(r = 0.97\) and RMACRO \(r = 0.98\). Hence, our main results are unlikely to be sensitive to omission of OVERVALU from the policy index. This robustness is likely to reflect, at least in part, the high correlation that Rodriguez and Rodrik (2000) note between OVERVALU and a variable with a much clearer interpretation, the black market exchange rate premium, BMP.

References


[40] Heston, Alan, Robert Summers and Bettina Aten (2002), Penn World Table Version 6.1, Center for International Comparisons at the University of Pennsylvania (CICUP), October.


Source: Authors’ calculations. In the first panel, the solid line shows a hypothetical growth rate distribution for countries with unstable macroeconomic outcomes, the dashed line for countries with stable outcomes. The second shows the distributions in the data; see text for more details.

Figure 2 – Box-plots for growth rates, countries grouped by stability

Source: Authors’ calculations. This graph shows three Tukey box-plots for three groups, from least stable (group 1) to most stable (group 3). The horizontal line in the box indicates the median growth rate for that group. The stability measure is RMACRO in the first panel and BMP in the second.
Figure 3 – Actual and counterfactual distributions of growth rates

Source: Author’s calculations. The solid line shows the actual distribution, the dashed line the counterfactual. The second panel allows for the effect of macroeconomic stability on investment.

Figure 4 – Actual and counterfactual distributions of steady-state log income

Source: Authors’ calculations. The solid line shows the actual distribution, the dashed line the counterfactual. The second panel allows for the effect of macroeconomic stability on investment.
### Table 1
Principal Component Analysis

<table>
<thead>
<tr>
<th>Variable</th>
<th>Expected Sign</th>
<th>(1) MACRO</th>
<th>(1) RMACRO</th>
<th>(1) MACROOL</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>1st PC</td>
<td>2nd PC</td>
<td>1st PC</td>
</tr>
<tr>
<td>SURPLUS</td>
<td>+</td>
<td>0.484</td>
<td>0.579</td>
<td>0.340</td>
</tr>
<tr>
<td>INFLA</td>
<td>-</td>
<td>-0.647</td>
<td>0.437</td>
<td>-0.744</td>
</tr>
<tr>
<td>BMP</td>
<td>-</td>
<td>-0.848</td>
<td>0.184</td>
<td>-0.888</td>
</tr>
<tr>
<td>OVERVALU</td>
<td>-</td>
<td>-0.503</td>
<td>-0.633</td>
<td>-0.395</td>
</tr>
<tr>
<td>ERATE</td>
<td>-</td>
<td>-0.688</td>
<td>0.232</td>
<td>-0.653</td>
</tr>
</tbody>
</table>

Number of countries: 78

% Variance explained: 41.94 20.29 41.27 24.00 37.29 23.10

Notes: Numbers shown are the correlations between principal components (PCs) and the corresponding variables. Numbers in bold indicate the highest correlations between a given principal component and corresponding variables. Column (2) uses an outlier-robust version of PCA; see main text. Column (3) is based on a classical PCA but excluding Guyana, Nicaragua and Sudan. These are the outliers suggested by the diagnostic graph recommended in Hubert et al. (2005).

### Table 2
Correlations between GDP growth and various policy indices

<table>
<thead>
<tr>
<th></th>
<th>RGDP7099C</th>
<th>MACRO</th>
<th>RMACRO</th>
<th>MACROOL</th>
<th>BD</th>
<th>ELR7097</th>
</tr>
</thead>
<tbody>
<tr>
<td>RGDP7099C</td>
<td>1.0000</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>MACRO</td>
<td>0.4715</td>
<td>1.0000</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>RMACRO</td>
<td>0.4194</td>
<td>0.9759</td>
<td>1.0000</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>MACROOL</td>
<td>0.4087</td>
<td>0.9952</td>
<td>0.9913</td>
<td>1.0000</td>
<td></td>
<td></td>
</tr>
<tr>
<td>BD</td>
<td>0.6618</td>
<td>0.6658</td>
<td>0.6226</td>
<td>0.5848</td>
<td>1.0000</td>
<td></td>
</tr>
<tr>
<td>ELR7097</td>
<td>0.5570</td>
<td>0.6031</td>
<td>0.6208</td>
<td>0.6450</td>
<td>0.8498</td>
<td>1.0000</td>
</tr>
</tbody>
</table>

Notes: This table shows the correlations between growth (RGDP7099C), our indicators MACRO, RMACRO and MACROOL (see Table 1), the Burnside-Dollar policy index BD, and the Easterly et al. (2004) policy index.
<table>
<thead>
<tr>
<th>Column</th>
<th>(1)</th>
<th>(2)</th>
<th>(3)</th>
<th>(4)</th>
<th>(5)</th>
<th>(6)</th>
<th>(7)</th>
<th>(8)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Estimation</td>
<td>OLS</td>
<td>OLS</td>
<td>OLS</td>
<td>OLS</td>
<td>2SLS</td>
<td>OLS</td>
<td>OLS</td>
<td></td>
</tr>
<tr>
<td>Regime</td>
<td>All</td>
<td>All</td>
<td>All</td>
<td>All</td>
<td>All</td>
<td>1</td>
<td>2</td>
<td></td>
</tr>
<tr>
<td>Observations</td>
<td>70</td>
<td>70</td>
<td>70</td>
<td>70</td>
<td>60</td>
<td>70</td>
<td>42</td>
<td>28</td>
</tr>
<tr>
<td>RMACRO</td>
<td>0.71</td>
<td>0.49</td>
<td>0.64</td>
<td>0.64</td>
<td>1.35</td>
<td>0.70</td>
<td>-1.20</td>
<td></td>
</tr>
<tr>
<td>(0.30)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Initial income</td>
<td>-1.10</td>
<td>-0.26</td>
<td>-0.80</td>
<td>-1.04</td>
<td>-1.15</td>
<td>-0.98</td>
<td>-0.83</td>
<td></td>
</tr>
<tr>
<td>(0.37)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Investment</td>
<td>1.07</td>
<td>1.10</td>
<td>0.83</td>
<td>0.84</td>
<td>0.56</td>
<td>0.45</td>
<td>1.52</td>
<td></td>
</tr>
<tr>
<td>(0.32)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Population growth</td>
<td>-0.21</td>
<td>-0.19</td>
<td>-0.12</td>
<td>-0.10</td>
<td>-0.02</td>
<td>-0.15</td>
<td>0.08</td>
<td></td>
</tr>
<tr>
<td>(0.23)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Literacy</td>
<td>0.68</td>
<td>0.88</td>
<td>1.12</td>
<td>0.72</td>
<td>0.41</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>(0.31)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Average schooling</td>
<td>0.79</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>(0.27)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>GOVKKM</td>
<td>1.06</td>
<td>1.91</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>(0.98)</td>
<td></td>
<td>(0.44)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Investment elasticity</td>
<td>1.18</td>
<td>1.69</td>
<td>0.96</td>
<td>0.89</td>
<td>0.69</td>
<td>0.66</td>
<td></td>
<td></td>
</tr>
<tr>
<td>R²</td>
<td>0.51</td>
<td>0.37</td>
<td>0.51</td>
<td>0.57</td>
<td>0.55</td>
<td>n/a</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Regression s.e.</td>
<td>1.56</td>
<td>1.75</td>
<td>1.57</td>
<td>1.47</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Heteroscedasticity</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Breusch-Pagan</td>
<td>0.32</td>
<td>0.02</td>
<td>0.07</td>
<td>0.27</td>
<td>0.18</td>
<td>0.48</td>
<td></td>
<td></td>
</tr>
<tr>
<td>White</td>
<td>0.66</td>
<td>0.19</td>
<td>0.03</td>
<td>0.64</td>
<td>0.35</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Ramsey RESET</td>
<td>0.90</td>
<td>0.58</td>
<td>0.02</td>
<td>0.68</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Anderson-Rubin</td>
<td>0.02</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Note: MacKinnon-White heteroskedasticity-consistent (hc3) standard errors reported in parentheses, for all columns except (6). The dependent variable is the annual growth rate over 1970-99, in percentage points. The explanatory variables are standardized to have a standard deviation of one in the 70-country sample. Regressions (1)-(6) include five regional dummies. The estimated constant and coefficients on regional dummies are not reported. The ‘Investment elasticity’ is the elasticity of the steady-state income level with respect to the investment rate. ‘Heteroscedasticity’ reports p-values associated with two tests for heteroscedasticity. ‘Ramsey RESET’ is the p-value associated with a RESET test. For column (6), ‘Anderson-Rubin’ is the p-value associated with the Anderson-Rubin test for the significance of the endogenous explanatory variable (RMACRO).
## Table 4
Threshold estimation

<table>
<thead>
<tr>
<th>Z:</th>
<th>FR</th>
<th>ABSLAT</th>
<th>GOVKKM</th>
<th>POLITY</th>
<th>POLCON</th>
<th>EXPRISK</th>
</tr>
</thead>
<tbody>
<tr>
<td>RMACRO threshold</td>
<td>0.068</td>
<td>0.167</td>
<td>0.030</td>
<td>0.002</td>
<td>0.013</td>
<td>0.012</td>
</tr>
<tr>
<td>Z threshold</td>
<td>0.324</td>
<td>0.320</td>
<td>0.271</td>
<td>0.523</td>
<td>0.600</td>
<td>0.354</td>
</tr>
<tr>
<td>γ - RMACRO</td>
<td>0.309</td>
<td>0.180</td>
<td>0.297</td>
<td>-0.185</td>
<td>0.297</td>
<td>0.309</td>
</tr>
<tr>
<td>95% C. I., Lower</td>
<td>-0.375</td>
<td>-0.300</td>
<td>-0.520</td>
<td>-0.374</td>
<td>-0.520</td>
<td>-1.241</td>
</tr>
<tr>
<td>95% C. I., Higher</td>
<td>0.309</td>
<td>0.309</td>
<td>0.714</td>
<td>0.309</td>
<td>0.618</td>
<td>0.324</td>
</tr>
<tr>
<td>N [RMACRO&lt;=γ]</td>
<td>43</td>
<td>36</td>
<td>42</td>
<td>29</td>
<td>42</td>
<td>36</td>
</tr>
<tr>
<td>N [RMACRO&gt;γ]</td>
<td>27</td>
<td>34</td>
<td>28</td>
<td>40</td>
<td>28</td>
<td>19</td>
</tr>
<tr>
<td>R² [RMACRO&lt;=γ]</td>
<td>0.43</td>
<td>0.60</td>
<td>0.47</td>
<td>0.40</td>
<td>0.42</td>
<td>0.48</td>
</tr>
<tr>
<td>R² [RMACRO&gt;γ]</td>
<td>0.80</td>
<td>0.76</td>
<td>0.90</td>
<td>0.77</td>
<td>0.78</td>
<td>0.82</td>
</tr>
<tr>
<td>Z p-val [RMACRO&lt;=γ]</td>
<td>0.90</td>
<td>0.00</td>
<td>0.29</td>
<td>0.57</td>
<td>0.85</td>
<td>0.49</td>
</tr>
<tr>
<td>Z p-val [RMACRO&gt;γ]</td>
<td>0.29</td>
<td>0.09</td>
<td>0.00</td>
<td>0.07</td>
<td>0.51</td>
<td>0.04</td>
</tr>
</tbody>
</table>

Note: ‘RMACRO threshold’ is the p-value for the Hansen (1996) test of a threshold in RMACRO, in a model which includes RMACRO, the Solow variables, and the Z variable shown at the top of the corresponding column. ‘Z Threshold’ is the p-value for the Hansen (1996) test of a threshold associated with the Z variable. The tests indicate that RMACRO can be used to divide the sample into two regimes. The lower rows show the threshold γ for RMACRO estimated using the Hansen (2000) procedure; the 95% confidence interval for the threshold (which need not be symmetric); the number of observations in the two regimes either side of the threshold; the R² of the separate growth regressions for the two regimes; and the p-value of the Z variable for each of the two regimes. Note that the growth regression always has highest explanatory power in the subsample with greater macroeconomic stability; for an example, see ‘Regime 1’ and ‘Regime 2’ in columns (7) and (8) of Table 3.
Appendix

Table A-1
Countries in the main (70-country) regression sample
ordered by RMACRO, from worst to best

Nicaragua, Uganda, Ghana, Argentina, Democratic Republic of Congo, Guyana, Iran,
Zambia, Bolivia, Brazil, Peru, El Salvador, Niger, Uruguay, Egypt, Syria, Venezuela,
Jamaica, Zimbabwe, Turkey, Mauritania, Costa Rica, Paraguay, Chile, Malawi, Haiti,
Rwanda, Israel, Honduras, Burundi, Dominican Republic, Guatemala, Bangladesh,
Ethiopia, Sri Lanka, Mexico, Madagascar, Lesotho, Colombia, Kenya, Trinidad and
Tobago, Nepal, India, Botswana, Pakistan, Nigeria, Papua New Guinea, the Philippines,
Indonesia, South Korea, Tunisia, Ecuador, Mauritius, Republic of Congo, Morocco,
Mali, Chad, Cameroon, Cyprus, Central African Republic, Burkina Faso, Senegal, Benin,
Fiji, Jordan, Togo, Malaysia, Panama, Thailand, Singapore.
Table A-2
Variables and definitions

<table>
<thead>
<tr>
<th>Variable</th>
<th>Description</th>
<th>Sources</th>
</tr>
</thead>
<tbody>
<tr>
<td>INVEST</td>
<td>Natural logarithm of investment share in GDP, 1970-99</td>
<td>Penn World Table 6.1</td>
</tr>
<tr>
<td>POPG</td>
<td>Natural logarithm of average annual growth rate of population aged 15-64,</td>
<td>World Bank (2004)</td>
</tr>
<tr>
<td></td>
<td>1970-99, plus 0.05.</td>
<td></td>
</tr>
<tr>
<td>SCHOOL70</td>
<td>Natural logarithm of average years of schooling at all educational levels of</td>
<td>Barro and Lee (2001)</td>
</tr>
<tr>
<td></td>
<td>population aged over 15 in 1970.</td>
<td></td>
</tr>
<tr>
<td>RGDPPC70</td>
<td>Natural logarithm of real GDP per capita (rgdpcp) in 1970.</td>
<td>Penn World Table 6.1</td>
</tr>
<tr>
<td>RGD7099C</td>
<td>Natural logarithm of real GDP per capita (rgdpcp) in 1999 minus same variable for 1970. This is divided by 29, to obtain annual growth rates.</td>
<td>Penn World Table 6.1</td>
</tr>
<tr>
<td>RGD7099W</td>
<td>Natural log of real GDP per worker (rgdpwok) in 1999 minus that of 1970. This is divided by 29, to obtain annual growth rates.</td>
<td>Penn World Table 6.1</td>
</tr>
<tr>
<td>Regional dummies</td>
<td>Five regions: East Asia and the Pacific (RGNNEAP), Middle East and North Africa (RGNMENA), South Asia (RGNSA), Sub-Saharan Africa (RGNSSA), and Latin America and Caribbean (RGNLAC)</td>
<td>Easterly and Sewadeh (2002)</td>
</tr>
<tr>
<td>MACRO</td>
<td>The first principal component from a classical principal components analysis of SURPLUS, INFLA, BMP, OVERVALU, and ERATE. Higher values mean better policy outcomes.</td>
<td>see text</td>
</tr>
<tr>
<td>RMACRO</td>
<td>As above, but from a robust principal components analysis.</td>
<td>see text</td>
</tr>
<tr>
<td>INFLA</td>
<td>Natural log of (1+inflation rate based on median GDP deflator)</td>
<td>World Bank (2004)</td>
</tr>
<tr>
<td>REALI</td>
<td>Mean lending rate adjusted by GDP deflator</td>
<td>World Bank (2004)</td>
</tr>
<tr>
<td>BMP</td>
<td>Natural log of (1+mean black market premium)</td>
<td></td>
</tr>
<tr>
<td>ERATE</td>
<td>Variation of the Dollar real exchange rate measure around its mean.</td>
<td>Dollar (1992)</td>
</tr>
<tr>
<td>POLITY</td>
<td>Measures degree of democracy. The POLITY score is the democratic score minus autocratic score. 0-10 scale, where higher values mean higher degree of democracy. We use the mean value 1970-1999.</td>
<td>Marshall and Jaggers (2000)</td>
</tr>
<tr>
<td>EXPRISK</td>
<td>Protection against expropriation risk. Higher values mean lower risk.</td>
<td></td>
</tr>
<tr>
<td>GOVKKKM</td>
<td>A composite index of overall quality of governance. We use the mean of indices for voice and accountability, political stability, government effectiveness, regulatory quality, rule and law, and corruption, during the period 1996-2000. Higher values mean higher-quality governance.</td>
<td>Kaufmann, Kraay and Mastruzzi (2005)</td>
</tr>
</tbody>
</table>

Note: For a description of the SDM controls used as candidate predictors in our implementation of Bayesian Model Averaging, see Table 1 of Sala-i-Martin et al. (2004). As discussed in the main text, we restrict attention to the 31 variables with a posterior inclusion probability greater than 4% in their Table 2.