Macroeconomic policy and the distribution of growth rates

Vatcharin Sirimaneetham
Jonathan Temple

Discussion Paper No. 06/584

March 2006

Department of Economics
University of Bristol
8 Woodland Road
Bristol BS8 1TN

Macroeconomic policy and the distribution of growth rates*

Vatcharin Sirimaneetham
Department of Economics, University of Bristol
8 Woodland Road, Bristol BS8 1TN, UK

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

March 27, 2006

Abstract

We examine the view that high-quality macroeconomic policy is a necessary, but not sufficient, condition for economic growth. We first construct a new index of the quality of macroeconomic policy. We then directly compare growth rate distributions across countries with good and bad policies; use Bayesian methods to examine the partial correlation between policy and growth; and outline how growth and steady-state income levels might have differed, had all countries achieved good policy outcomes. One finding is that bad macroeconomic policies can be offset by other factors, but the fastest-growing countries in our sample all shared high-quality macroeconomic management.

JEL classifications: O23, O40

Keywords: macroeconomic policy, economic growth, Washington Consensus, Bayesian Model Averaging, counterfactuals

---

*We are grateful to David Ashton, Holger Breinlich, Andreas Leukert and Ian Painter for helpful contributions or discussion on related topics. We are responsible for any errors. 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

The belief that economic growth requires sound macroeconomic policy is a central element of development orthodoxy. Even those sceptical about that orthodoxy will often agree that macroeconomic stability is a precondition for successful economic development. At the broadest level, the macroeconomic stability of East Asian countries between the early 1960s and the late 1990s could help to explain why East Asian countries have sustained high growth rates. In contrast, sub-Saharan Africa and Latin America have endured a painful combination of macroeconomic disarray and slow growth.

As most economists would expect, macroeconomic mismanagement could explain not only slow growth, but also why some developing countries have become heavily indebted. Even if slow growth is attributed to problems with external debt, the origins of a debt crisis can typically be traced back to policy decisions. Easterly (2002) finds that the group of highly indebted poor countries (the HIPCs) had worse macroeconomic policies over 1980-97 than other developing countries, even after controlling for income levels.

Although these various observations may seem convincing, the strength of the empirical relationship between macroeconomic policy and growth continues to be disputed. One argument is that, even where there is some evidence of a correlation, it arises only because of slow growth in countries with terrible policies. Once the quality of policy is above a certain threshold, the marginal effect of better policy on growth 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 following the work of Acemoglu et al. (2001) on institutions, has argued that policies typically lack explanatory power relative to institutional weaknesses.

In this paper, we revisit the relationship between macroeconomic policy and growth. This is well-worked ground, and a new paper on this topic must try hard to justify its existence. We depart from the existing literature in at least four ways. First, we seek to avoid some of the weaknesses of past work by constructing a new composite index of the quality of macroeconomic policy. We use an outlier-robust version of principal components analysis to aggregate five different policy indicators. Our new composite index is correlated with previous measures of policy quality, such as the well-known index used by Burnside and Dollar (2000) in their study of the effectiveness of foreign aid. But the new index also has some advantages over previous measures for the questions we investigate here, and perhaps in future research on a variety of related topics.

Second, we examine a popular and important hypothesis in more detail than usual. Much of the commentary on macroeconomic policy and growth can be reduced to a very simple hypothesis: sound policy is a necessary but not sufficient condition for rapid growth; bad policy is a sufficient condition for slow growth. We call this the 'weakest link' view, because growth performance is only as strong as the weakest link in a set of policy outcomes. Rapid growth relies on simultaneously achieving a number of goals and, if any one of these is absent, growth will swiftly come to a halt.

Although this is a common view of the role of policy, previous research has not explored its empirical implications in any detail. We use direct comparisons of growth rate distributions, where countries are classed into groups according to our new policy indicator. We find that the extent of support for the popular view is less than overwhelming. In particular, even when a country ranks low in terms of macroeconomic policy, this is not a sufficient condition for slow growth. We do find, however, that sustained rapid growth is confined to a set of countries with high-quality macroeconomic management.

Third, we examine in more detail how growth and policy are related. We first use orthodox growth regressions to quantify the effects of macroeconomic policy over the period 1970-99, restricting the sample to developing countries. This approach has been widely used, but suffers from a number of important weaknesses, not least the model uncertainty problem highlighted by Levine and Renelt (1992). We therefore supplement these simple regressions with Bayesian methods for model averaging, which allow us to investigate sensitivity to the regression specification in a systematic way. Drawing on a set of many other candidate predictors, we find some evidence that the new macroeconomic policy indicator matters for growth, and strong evidence that government decisions matter when a wider range of decisions are taken into account.

More precisely, in a BMA exercise that closely follows Sala-i-Martin et al. (2004), we find that the best-performing models nearly always contain at least one of three macroeconomic policy variables (one of which is our new indicator) regardless of variation in the rest of the specification. Hence, we establish relatively strong evidence that macroeconomic policies help to explain variation in growth rates across developing countries, even when model uncertainty is taken into account. We also look at whether the effect of macroeconomic policy is robust to the inclusion of measures of institutional quality.

Finally, we use our main regression 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 shared the same quality of policy throughout the last thirty years of the twentieth century. One simple device is to set the macroeconomic policy index for all countries to the value for Malaysia, the country at the 95th percentile of policy quality, and then compute counterfactual growth rates and steady-state income levels. This shows the extent to which the distributions of growth rates and steady-state income levels might have looked different, if countries had shared the same quality of policy. We can also see whether bad policy accounts for the shape of the international distribution of output per worker, including the "twin peaks" pattern identified by Quah in a series of papers (for example, Quah 1996).

The main conclusions of our analysis are as follows. In general, countries with good macroeconomic policies appear to grow more quickly. Nevertheless, some countries have grown reasonably quickly despite weak policy. This implies that bad macroeconomic management is not a sufficient condition for slow growth, and can sometimes be offset by strengths elsewhere. At the same time, we show the fastest growth rates are confined to countries with high-quality macroeconomic policies. In this respect, there is some support for the macroeconomic orthodoxies of the Washington Consensus, at least as a long-run proposition.[1]

[1] It is worth noting that the paper does not address the subtler and much more difficult questions that relate to short-run policy activism such as demand management. Our

The rest of the paper is organized as follows. Section 2 briefly reviews the existing literature on policy and growth, and then discusses the 'weakest link' view in more detail. Section 3 will describe our construction of a new measure of the quality of policy. Section 4 will then use this measure to group countries into those with good and bad policies, before comparing the unconditional distributions of growth rates across these two groups, to examine the 'weakest link' view. Section 5 presents a more conventional analysis based on growth regressions, while section 6 examines the robustness of the policy effect to changes in the regression specification, using recently developed Bayesian methods. Section 7 uses the core growth regressions to generate counterfactual distributions of growth rates and steady-state levels of income. Finally, section 8 concludes.

2 Relation to existing literature

The literature on policy and growth has traditionally emphasized short-run macroeconomic management and trade policy. We will focus throughout on the effects of macroeconomic policy, partly because it is relatively easy to measure, and partly because attempts to impose macroeconomic orthodoxy are often controversial in practice. Motivated by these considerations, cross-country studies such as Bleaney (1996) and Fischer (1991, 1993) investigated the role of macroeconomic stability in sustaining growth, and tended to argue that policy plays a vital role.

The account of this literature in Easterly (2005) emphasizes the shift that has taken place in recent years. Initially, commentators on development often argued that policy differences could account for most of the post-1960 variation in developing country growth rates. This belief has been undermined in several ways. It is not difficult to find examples of countries that have stagnated despite orthodox macroeconomic policies. The time series variation in policy and growth is not especially supportive, either. Improvements in policy indicators explain relatively few growth accelerations (Hausmann, Pritchett and Rodrik 2005) and in general policy indicators are

results concern macroeconomic policy 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.

far more persistent than growth rates, suggesting that policy will usually leave at least the medium-run variation in growth unexplained (Easterly et al. 1993).

Looking at the broader picture, macroeconomic policy has generally improved over time, whereas developing country growth performance was weaker in the 1980s and 1990s than previously. The reasons for the post-1980 growth collapse in developing countries are discussed in Easterly (2001b) and Rodrik (1999) and appear more complicated than a simple 'bad policy' story.[2] Finally, some studies, notably Easterly and Levine (2003), have found that growth and policy variables are not robustly correlated in the cross-country data when controlling for institutional development.

With all this in mind, some recent contributions have suggested that the role of macroeconomic policy is typically overstated. Perhaps bad policies are best seen as a symptom of deeper underlying problems, such as institutional weaknesses. Since macroeconomic disarray is often associated with several problems at once, it is often argued that it will be hard to disentangle the effects of specific policies from one another, let alone other growth determinants. This is an important motivation for constructing an aggregate index of the quality of policy, an approach that is central to this paper.

Although some of the claims for the importance of policy may be exaggerated, another hypothesis continues to have more general support: it is often argued that high quality macroeconomic management is a necessary, but not sufficient, condition for rapid growth. Performance is only as strong as the weakest link in a set of policy outcomes. This view has come to dominate assessments of the role of policy (for example, Easterly 2005) but has had relatively little impact on theory and empirical methods.[3]

An especially clear and persuasive exposition of the 'weakest link' 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

[2] Although it could be argued that macroeconomic mismanagement, as well as bad luck, played a significant role in the external debt crisis of the 1980s. The slow growth of that decade might then be attributed partly to policy decisions in the 1970s, as well as the oil shocks and the rise in world real interest rates.

[3] An exception is Hausmann, Rodrik and Velasco (2005), who provide an analytical framework for isolating the binding growth constraints in a particular setting.

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.

If good macroeconomic policies are necessary but not sufficient for growth, this has implications for the kinds of empirical model that should be estimated. For example, the idea could be captured by a simple nonlinear model with two regimes:

$$\begin{aligned} g &= \gamma_0 + \varepsilon_0 \text{ if } P < P^* \\ &= \gamma_1 + \alpha P + \beta' Z + \varepsilon_1 \text{ if } P \geq P^* \end{aligned} \tag{1}$$

where g is the growth rate, P indicates the quality of macroeconomic policy, Z is a vector of other growth determinants, γ_0, γ_1 and α are parameters and β is a parameter vector. High values of P indicate good macroeconomic policies. If the quality of policy falls below a threshold value P^*, governments effectively destroy any prospect of growth (given a low value of γ_0 and a low variance of the error term ε_0) regardless of other country characteristics.

This contrasts with the linear models that dominate the growth literature:

$$g = \gamma + \alpha P + \beta' Z + \varepsilon \tag{2}$$

in which bad policies can be offset by other factors, and growth varies smoothly with the policy indicator, without any kind of threshold effect. Clearly there will be circumstances in which the alternative models (1) and (2) are not greatly different. But it is worth noting that the conventional model (2) builds in a policy-growth relationship somewhat different from the 'necessary condition' or 'weakest link' view that is embedded in many informal accounts.

We do not examine models with thresholds, but instead consider a potentially more general testable implication. If the 'weakest link' view is right, we should expect to see that countries with bad policies have growth rates that are tightly distributed around a low mean, because bad policy is a sufficient condition for slow growth. In contrast, where countries have good policies, we should observe much wider dispersion in growth rates around a higher mean. The wide dispersion arises because countries with good policies 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 with good and bad policy, the distributions of growth rates across countries might look like the hypothetical example in figure 1: the solid line represents a possible distribution for countries with bad policy, while the dashed line represents a possible distribution for countries with good policy.

With this in mind, our paper will pay less attention to the conditional mean of the growth rate than is usual. Instead, we compare the shape of the entire growth rate distribution across countries with good and bad policies. In some of our empirical work, we will use box-plots and kernel density estimates to examine whether the patterns look similar to the hypothetical pattern we sketched in figure 1. In general, good policies are associated with higher growth rates, but there is substantial variation for both groups. In other words, some countries have grown moderately quickly despite weak policy on average. But it is also worth noting that the highest growth rates in our sample are confined to countries with high-quality macroeconomic policies.

It is important to emphasize that our work shares important deficiencies with other empirical research on policy and growth. One especially important criticism, articulated in Rodrik (2005), is that policy outcomes - ultimately representing decision variables - must be endogenous to social and economic circumstances, calling into question the usual exogeneity assumptions. In terms of the statistical and microeconometric literature on treatment effects, the assignment of treatments (government policies) is not randomized, and nor is it likely to be "ignorable" in the technical sense of that term. This implies that policy indicators will almost certainly be correlated with country characteristics that are not observed by the econometrician.

In cross-section research of the kind we pursue here, the solutions to this problem are limited, not least because the literature has arguably failed to identify a genuinely convincing instrument for the quality of macroeconomic policy. Nevertheless, it remains interesting to see whether policy and growth are related in the cross-country data. This is especially so, given some recent claims that macroeconomic policies are only a small part of the development story.

3 Measuring macroeconomic policy

In this section, we discuss our indicators of the quality of macroeconomic policy, representing an important element of the Washington Consensus. We will combine several indicators to measure the overall quality of policy over 1970-99. This has a number of advantages. From a statistical point of view, it will tend to lessen the outlier problems associated with skewed distributions, and help to alleviate measurement error. From an economic point of view, the new index aims to measure an underlying latent variable, the quality of the macroeconomic decision-making process, rather than seeking to rely on more specific possible 'symptoms' like high inflation. This approach may work especially well when, as suggested by Sala-i-Martin (1991), macroeconomic disarray tends to be associated with undesirable outcomes on a range of indicators.

It is important to note that we are not seeking to examine all elements of the Washington Consensus. As initially summarized by Williamson (1990), the Consensus enshrines principles that go well beyond macroeconomic policies. These include the need for tax reform, financial liberalization, liberalized trade policy, openness to foreign direct investment, privatization, deregulation, and the protection of private property rights.[4] An examination of these dimensions is well beyond the scope of this paper, and would take us into many different controversies. Fischer (2003) argues that, although an over-simplification, the policies associated with the Washington Consensus

[4] Modern development orthodoxy, as promoted by the Washington institutions, extends to even more aspects of policy. Rodrik (2002) characterizes an 'augmented' Washington consensus.

are still "a useful shorthand description of a major part of a desirable basic policy orientation" (2003, page 6).

We focus on the potential importance of macroeconomic policies in that "basic policy orientation". We will consider, in particular, the roles of fiscal discipline, inflation and exchange rate management.[5] We measure average performance in these areas over thirty years, 1970-1999. We combine the individual indicators to obtain an index of policy quality for this period, using either a classical principal components analysis (PCA) or a robust extension of PCA based on the minimum covariance determinant method of Rousseeuw (1984). Our empirical analysis will focus on developing countries with available data, but excluding transition economies, and small countries where the 1970 population size was below 250,000.

We now describe the individual policy indicators that we will use to construct a composite index. To capture fiscal disclipine we use data on the average central government budget surplus as a share of GDP ($SURPLUS$) over 1970-99.[6] 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 either excluding these countries, or replacing $SURPLUS$ with the monotonic but bounded transformation $\arctan(SURPLUS)$.[7]

To measure success in keeping inflation low, we construct a variable $INFLA$. 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 standard use of the mean, this measure is less likely to be dominated

[5] For a related approach, including additional dimensions of the Washington Consensus, see Berr et al. (2005).

[6] Although we have also experimented with including the stock of central government debt relative to GDP, the latter variable is available for a smaller number of countries, and so we use $SURPLUS$ in what follows.

[7] 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 $SURPLUS$, the lowest value is less than one standard deviation below the mean, compared to five standard deviations below in the raw data.

by a small number of hyperinflation episodes, at least where these are short-lived.

We use three measures that relate to various aspects of exchange rate management. These are the black market premium (*BMP*), an index of currency overvaluation or real exchange rate distortion (*OVERVALU*) and a measure of the variability in exchange rate distortions (*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 *BMP* is defined as the natural logarithm of one plus the mean value of the black market premium over the period.

The two variables *OVERVALU* and *ERATE* were introduced in Dollar (1992) and, in the case of *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 *OVERVALU* could indicate real overvaluation and trade restrictions. In contrast, low values of *OVERVALU* may be associated with outward orientation. This measure has sometimes been criticised, particularly as an index of trade restrictions; we discuss these issues in Appendix 1.

The final measure of exchange rate management we use is denoted *ERATE*. This 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 but, given the likely role of inflation in generating movements in the overvaluation index, may also be an indicator of more general forms of macroeconomic instability (Rodriguez and Rodrik 2000). A possible danger here is that *ERATE* may partly reflect political instability, with effects that are distinct from macroeconomic mismanagement. We do not adopt other measures of policy variability for this reason.

Although sometimes we will use the five policy indicators individually, for most of our analysis we aggregate them into a composite index. This index is designed to capture the overall quality of macroeconomic policy. Constructing such an index helps in several ways: it reduces the measurement error associated with taking a single indicator as a proxy for the quality of policy, and helps to limit the influence of outlying observations. The use of a composite index also acknowledges a basic limitation of the cross-country data. 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 quality. Arguably, there is more hope of answering questions about policy and growth reliably when policy is deliberately characterized in these broad terms.

The best-known macroeconomic 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-Warner (1995) indicator of openness to trade.[8] 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 much less well-suited to our purposes than to those of Burnside and Dollar. 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 for this it makes sense to use a composite policy index that makes no use of information on growth rates.

To construct our composite index, we use a principal components analysis. This takes p specific indicators and yields new indices (the principal components) $P_1, P_2, ...P_p$ that are mutually uncorrelated and capture different dimensions of the data. In our work we standardize the indicators to have unit variance; equivalently, we base the principal components on

[8] 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).

the correlation matrix rather than the covariance matrix. We use solely the first principal component. Formally, this is defined by a vector of weights $a = (a_1, a_2, ..., a_p)'$ on the (standardized) indicators $X_1, X_2, ...X_p$ such that the linear combination

$$P_1 = a_1X_1 + a_2X_2 + ... + a_pX_p$$

has the maximum variance for any possible choice of weights, subject to the sum-of-squares normalization that $a'a = 1$. We use this method to aggregate different sets of components into a new measure of policy quality. A key assumption here is that a well-measured aggregate index can be written as a linear function of the policy indicators.

To recap, our policy indicators are the budget surplus, inflation, the black market premium, real exchange rate distortions and real exchange rate volatility.[9] 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 any sense. Testing for this "sphericity" case, allowing for sampling variability in the correlations, is a standard problem in multivariate analysis. We implement a Bartlett-type likelihood ratio test, where our precise test statistic is:

$$S_p = (n-1)p \log \frac{\frac{1}{p}\sum \lambda_j}{\left(\prod \lambda_j\right)^{1/p}}$$

as in Flury and Riedwyl (1988, p. 203). Here n is the number of cases (here, countries) and the λ_j are the eigenvalues associated with the p principal components $j = 1, ..., p$. Under the assumption of multivariate normality, this statistic is asymptotically distributed as chi-squared with $p(p+1)/2-1$ degrees of freedom. For the first principal components analysis we report below, the test statistic is 60.97 with 14 degrees of freedom, and so we comfortably reject sphericity at the 1% level.

[9] In a previous analysis we also used data on real interest rates, since low rates often indicate financial repression. The component loading for the real interest rate was effectively zero, and we therefore exclude it from the construction of our preferred measure.

We always normalize the first principal component in such a way that high values indicate good policy. The structure of our first composite 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 variables (all with mean zero and unit variance) we can write our first composite index as

$$\begin{aligned} MACRO &= 0.334 * SURPLUS - 0.447 * INFLA - 0.585 * BMP(3) \\ &\quad -0.347 * OVERVALU - 0.475 * ERATE \end{aligned}$$

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 managing their macroeconomic conditions during 1970-1999 were Singapore, Thailand, Malaysia, Panama and Benin. In contrast, the analysis suggests that the quality of policy was unusually low in Nicaragua, Guyana, Sudan, Uganda and Zambia.

A drawback of any principal components analysis, especially in a sample of the current size, is that it may be highly sensitive to outlying observations. As Hubert et al. (2005) note, a classical principal components analysis is maximizing the variance and decomposing the covariance matrix, and both the variance and the covariance matrix can be highly sensitive to anomalous observations. 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, we also use methods for constructing outlier-robust principal components. Since we have five policy indicators and so relatively few dimensions, we can easily 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.[10]

Using this approach to estimating correlations, we can extract outlier-robust principal components. The correlations between the first two of these new principal components, and the policy indicators, are shown in column (2) of Table 1. In terms of loadings on the individual variables, the robust composite indicator can be written as:

$$\begin{aligned} RMACRO &= 0.101 * SURPLUS' - 0.578 * INFLA' - 0.693 * BMP' \\ &\quad -0.219 * OVERVALU' - 0.357 * ERATE' \end{aligned} \tag{4}$$

where the $'$ on the individual variables indicates that each has been centred using a robust estimate of their location. Relative to the classical PCA, the outlier-robust PCA places less weight on *SURPLUS*, *OVERVALU* and *ERATE*, and more weight on *INFLA* and *BMP*. Although the weights in (3) and (4) may look rather different, the simple correlation between *MACRO* and *RMACRO* is 0.98. Using the *RMACRO* index, the five best performing countries are Singapore, Thailand, Panama, Malaysia and Togo, and the five worst performing countries are Nicaragua, Uganda, Ghana, Argentina and the Democratic Republic of Congo.

An alternative approach for a PCA is to use the diagnostic plot suggested by Hubert et al. (2005). Using this plot we can identify possible outliers, which are then excluded from a classical 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 variance explained by the first principal component falls slightly, but the correlations between this component and the

[10] We use the `ROBPCA` program for S-Plus 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.

different indicators are very similar to those reported in column (1). This suggests that, although in principle there is some potential for variation in the aggregate policy indicators to be driven by unusual cases, the indices we use are robust to outlying observations.

It is interesting to note briefly the correlations between our new measures of policy quality, and those previously used in the literature. Table 2 shows the correlations between $MACRO$, $RMACRO$, 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 could be measuring similar aspects of performance. The is the case even though the Burnside-Dollar and Easterly et al. measures use a very different weighting strategy and additional information, from the Sachs-Warner measure of trade policy.

4 When is policy the weakest link?

First of all, we look at how growth varies across countries with good and bad policies. We order the countries by their values of the composite policy indicator $RMACRO$, 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 measure 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 across figure 2, we can see that the median growth rate is substantially lower in group 1 (with the worst policies) than in groups 2 and 3. There is less support for the idea that bad policy 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 $MACRO$ rather than $RMACRO$.

In figure 3 we use kernel density plots to summarize the same information in a slightly different way.[11] We exclude the intermediate group, for clarity. The solid line in the figure shows the distribution of growth rates for the group with the worst policies, while the dashed line shows the distribution for the group with the best policies.

The figure shows that growth is systematically higher with good policies (since the distribution is further to the right for countries with good policies). Contrary to the "weakest link" view we discussed earlier, bad policy does not necessarily preclude growth. There is substantial variation in growth across the countries with bad policy, 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 bad policy group, whereas there are seven countries that grew at least this rapidly in the good policy group (Cyprus, Indonesia, Malaysia, Mauritius, Singapore, South Korea and Thailand). On this evidence, avoiding bad macroeconomic policy outcomes is a necessary condition for sustaining high growth rates over a long period.

We have examined similar figures (not shown) for all five individual indicators, *SURPLUS, INFLA, BMP, OVERVALU* and *ERATE*.[12] The patterns in these figures are generally less supportive of the idea that good policy promotes growth. This suggests that combining the separate indicators into an overall index is a useful step. The evidence that policy matters is strongest for the Dollar index of exchange rate distortions (*OVERVALU*) and the black-market premium (*BMP*). In figure 4 we present box-plots for countries grouped according to the black-market premium.

5 Macroeconomic policy and growth regressions

In this section, we examine the relationship between macroeconomic policy and growth in a more orthodox way, using growth regressions. Later in the paper, we will use these simple OLS regressions to compute counterfactual distributions of growth rates and steady-state income levels.

[11] The samples are relatively small to apply these methods, and the choice of bandwidth becomes important. We discuss this choice in Appendix 2.

[12] Easterly (2005) also looks at this issue, using bar charts, and using short-run variation in the data to a greater extent than here.

A natural starting point for any growth regression is the empirical model introduced by Mankiw, Romer and Weil (1992). We estimate a version of their model using data for 1970-1999 (their original sample was for 1960-1985). We show that, conditional on the 'Solow determinants' of growth, notably investment shares, population growth, and initial income, growth has a robust partial correlation with our indices of macroeconomic policy. The effect is not only robust, but also sizeable. In our main regressions, an improvement in the quality of policy by one standard deviation would have raised the annual growth rate by around 0.5-0.7 percentage points over this period.

The Mankiw, Romer and Weil regression model is standard and we do not discuss it in detail here. Briefly, we regress the log difference in GDP per capita on the log of the investment share, the log of initial GDP per capita, the log of population growth plus 0.05, and a human capital variable. There are two main departures in our specification. First, we include regional dummies in all specifications; these can be motivated partly as proxies for the unobservable variable initial 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.[13] We measure educational attainment using 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 (2000). 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 original MRW results are fairly robust to using data over a different time period; the explanatory power is similar to the MRW regressions, although the effect of population growth is imprecisely estimated in this developing country sample.

In column (2) we look at a simple regression that includes only initial in-

[13] The use of a stock measure 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).

come, regional dummies and the robust policy index *RMACRO*. The index is significant at the 5% level and has a sizeable effect: a one-standard-deviation increase in the quality of policy 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; the effect of *RMACRO* is weaker, but remains significant at 12%. The reduction in the size of the coefficient suggests that the effects of macroeconomic management may work partly via the investment rate, an idea that we explore in more detail below.

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, and a one-standard-deviation increase in this variable would raise the growth rate by 0.65 percentage points. This result is robust to replacing the literacy rate with average years of schooling in 1970, as in column (5). This reduces the size of the sample by 10 observations.

The partial correlations between growth and our policy indicator do not appear to be driven by anomalous observations. The results are robust to the deletion of potential outliers, as identified by median (least absolute deviation) regression.[14] 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.[15] 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 structure (based on Ramsey's RESET statistic) 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).

Overall these results suggest that, conditional on the Solow growth de-

[14] 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.

[15] 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.

terminants and regional dummies, the quality of macroeconomic policy has some explanatory power for growth rates. An increase in the policy index of one standard deviation 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 22% at the end of the thirty years. Later in the paper, we will use these results to explore the role of policy in more detail, including the effects on the location and shape of the distributions of growth rates and steady-state income levels. Before then, we examine the robustness of the partial correlation between growth and policy, using recently developed Bayesian methods.

6 Robustness

Since Levine and Renelt (1992) it has often been argued that partial correlations in the empirical growth literature are not 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 variables on a priori grounds. 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 methods for model averaging as in Brock et al. (2003), Fernandez et al. (2002), Raftery (1995), Raftery et al. (1997) and Sala-i-Martin et al. (2004). In what follows we refer to the latter paper as SDM.

Brock et al. (2003), Malik and Temple (2005) and SDM all discuss applications of Bayesian model averaging to economic problems, and so we discuss the main ideas only briefly, drawing on the presentation in Raftery (1995).[16] Recall that Bayesians 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 our uncertainty

[16] For those interested in learning more about the key ideas, the discussion in Raftery (1995) is highly recommended.

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 acknowledges 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 θ_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 (Leamer 1978).

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

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

Here $p(\Delta \mid D, M_k)$ 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) of Schwarz (1978) 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 β_z is non-zero, by summing the posterior model probabilities for all models in which $\beta_z \neq 0$.

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. We provide some additional details on our implementation of BMA methods in Appendix 3.[17]

We complement the standard BMA methods (based on the BIC approximation) with the more sophisticated approach of Hoeting et al. (1996). This is because a potentially serious problem in the empirical study of growth data is that outliers may be present. Where some observations are unrepresentative, this could easily lead to some variables being assigned a high posterior probability of inclusion, and others not, where the majority of the data would point to a different conclusion. In general, any procedure for dealing with model uncertainty (or even just 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 estimator due to Rousseeuw (1984), and the standardized residuals used to identify possible outliers. Then model averaging is carried out but, as in Hoeting et al. (1996), a 'model' is defined as (1) a joint set of candidate predictors and (2) a set of observations identified as outliers, where the latter are some or all of those identified in the first stage. (This restriction is used to keep the dimensionality of the problem manageable.) Then a Markov chain Monte Carlo model composition (MC^3) approach, as in Madigan and York (1995), is used to approximate the posterior model probabilities. For more details of this approach, 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

[17] For a more general summary of how the approach is implemented and used to compute posterior model probabilities, and a discussion of the necessary assumptions, see Appendix 1 of Malik and Temple (2005).

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 probability of inclusion 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.

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.

When we combine the SDM data set with the data on our policy measure, *RMACRO*, our sample is reduced to 63 developing countries. In some of what follows, we extend the country coverage by imputing missing values for a small number of variables. This allows us to increase the number of countries to 72.[18] The decision to impute missing values involves a trade-off: we introduce measurement error, but at the same time we bring to bear additional information (for the extra countries) and lessen the biases that occur when data are missing in non-random ways. Here, the number of imputed values in the design matrix (containing the data on the explanatory variables) is just 21, representing fewer than 1% of the total number of cells in the design matrix ($32 \times 72 = 2304$).

In thinking about which sample is most appropriate, it is worth anticipating one aspect of our findings. The evidence that policy has explanatory power is always much stronger in the 72-country sample than in the 63-country sample. 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-

[18] The country missing from our earlier 73-country growth regression sample is Burkina Faso, which is not included in the SDM data set.

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 macroeconomic policy quality distribution.

At one level, the addition of the extra countries, with their relatively extreme outcomes for policy, clearly adds a great deal of identifying variation to the data set. At the same time, we must have considerable faith that growth and policy are reliably measured for these countries, if we are to give more weight to the results for the larger sample. Otherwise, there is a risk that the results will be driven by a misleading set of observations. 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 for more discussion). Rather than attempt to take sides in this debate, we present results for both samples. Readers can then draw their own conclusions from the evidence presented.

We do not report the full BMA results in detail, and 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. This is very low in the sample of 63 countries: just 1% or 2.5%, depending on whether we transform the three explanatory variables mentioned previously. The posterior probability of inclusion becomes 100% when we move to the full sample of 72 countries, however. The relevant posterior mean - 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 coefficients estimated in the previous section using growth regressions based on the MRW model.

The evidence for an effect of macroeconomic policy also becomes much stronger, even in the 63-country sample, if we exclude government investment as a share of GDP (measured for 1970-74) from the candidate predictors. This suggests that, at least conditional on other growth determinants, the macroeconomic policy index and government investment each capture, in slightly different ways, aspects of government decision-making that are relevant to growth.[19] It is also worth noting that SDM's government investment variable, which tends to hide any effect of $RMACRO$ in the 63-country sample, has a highly skewed distribution. The potential distorting effect of outliers on the regression surface may explain why $RMACRO$ is highly sensitive to its inclusion.

Given this point and the overall sensitivity to the sample, a more robust version of BMA may be preferable. We therefore implement the outlier-robust MC^3 approach of Hoeting et al. (1996) and find a general pattern similar to the previous results. One interesting feature of these results is that Dollar's original index of real exchange rate distortions has a high posterior inclusion probability, 87%, in the 63-country sample, and even higher (99%) in the 72-country sample. The evidence for a separate effect of $RMACRO$ is weak, but becomes much stronger if we exclude Dollar's index. The posterior probability of inclusion of $RMACRO$ then rises to 69% in the 72-country sample.[20]

Finally, we look at the question of whether the effects of macroeconomic policy are robust to the inclusion of measures of institutions. Acemoglu et al. (2003) and Easterly and Levine (2003) have argued that macroeconomic disarray may be a symptom of institutional weaknesses; after conditioning on measures of institutional quality, these authors find the effects of macroeconomic policy to be weak. Easterly (2005) concludes that "the long run effect of policies on development is difficult to discern once you also control for institutions" (page 1055).

To investigate this, we consider the relationship between growth and policy, when adding four measures of institutions to the BMA exercises.[21]

[19] At first glance, the role of the government investment share is consistent with the general view of Easterly (2001a) that slow growth is often associated with short-term behaviour by governments. The problem here is that the government investment share is usually *negatively* signed: high government investment is associated with low growth. When we use outlier-robust methods, however, the posterior inclusion probability of the government investment variable is much lower - below 10%.

[20] A remaining possibility is that the effect of policy on growth is nonlinear. But the BMA results are similar when we replace $RMACRO$ with a dummy variable that is equal to one for the countries in the lowest third of the $RMACRO$ distribution, and zero otherwise.

[21] 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.

These are the extent of democracy, based on the POLITY IV database of Marshall and Jaggers (2000), and averaged over 1970-99; a measure of the extent of political constraints due to Henisz (2000), again averaged over 1970-1999; a composite index of the quality of governance for 1996-2000, due to Kaufmann, Kraay and Mastruzzi (2005); and 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 based on outcomes rather than constraints, and this could lead us to exaggerate the effects of institutions, and understate the effects of policy.[22]

Initially, we exclude the expropriation risk measure because it reduces the sample of countries. When we add the other three measures of institutions to our previous BMA, the posterior inclusion probability of $RMACRO$ is again very sensitive to the sample. In the 63-country sample, the posterior inclusion probability for $RMACRO$ is just 0.8%, but in the 72-country sample it is 97.4%. If we use the outlier robust MC^3 approach, the inclusion probability of $RMACRO$ is 3% in the 63-country sample, and 53% in the 72-country sample. Incidentally, the results also strongly support the hypothesis that growth and institutions are highly correlated. The Kaufmann et al. (2005) measure dominates the other institutions measures, with an inclusion probability of 100%. The inclusion probabilities for the democracy and Henisz measures 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%) and the Kaufmann et al. measure (100%) continues to outperform the other measures of institutional quality. The Henisz and democracy 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%.

Past work, notably Levine and Renelt (1992), has argued that the partial correlation between policy and growth is not robust to the choice of control variables. In some ways, our findings are consistent with that widely held view. When we allow for a wide range of possible growth determinants, the

[22] See Glaeser et al. (2004) on the general desirability of using measures of constraints or rules, rather than measures closely related to equilibrium outcomes.

evidence that the specific policy index $RMACRO$ matters for growth is less than overwhelming. But if we take a broader view of government policy, there is a stronger case that it makes a difference. In our Bayesian model averaging exercises, at least one of three variables - $RMACRO$, the share of government investment in GDP, and Dollar's index of real exchange rate distortions - always has a high posterior inclusion probability.

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 tentative evidence that macroeconomic policy matters even when taking into account institutional quality (or perhaps more properly, measures of outcomes that are primarily associated with institutions). This suggests that policy helps to explain some of the variation in growth rates across developing countries. The next section will examine this in more detail.

7 Counterfactual distributions

This section attempts to place the size of the macroeconomic policy effect in broader perspective. One interesting way to assess the effects of macroeconomic policy 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 followed the same macroeconomic policies over 1970-99.[23]

That is the task we undertake in this section, using estimates of the effects of policy obtained from the previous regression results. An advantage of this approach is that we can see where in the distribution the role of policy may have been especially important, information that is not directly apparent from regression estimates. We can also see the extent to which bad macroeconomic policies might account for the pattern of "twin peaks", or bimodality, sometimes identified in the distribution of income per capita levels (Quah 1996).

[23] Kernel density estimates of counterfactual distributions are associated in particular with the work of DiNardo et al. (1997) on wage distributions. These methods have also been applied in growth economics by Desdoigts (1996, 2004).

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 the macroeconomic policy indicator 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 high quality policy, but countries with low growth did not, then imposing high-quality policy throughout the sample might only affect the lower end of the distribution. All this implies that 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 experienced the same quality of macroeconomic policy over 1970-99. First of all, we estimate a growth regression similar to those in the previous section. This controls for the MRW determinants of growth, regional dummies, and the policy index $RMACRO$. 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 the macroeconomic policy index at the 95th percentile in our sample, corresponding to Malaysia.

Figure 5 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 the quality of macroeconomic policy had been higher. It is worth noting that the shift takes place throughout the distribution.

This exercise holds the rate of investment constant, but some effects of better macroeconomic management might occur through the investment channel. To examine this, we carry out a growth regression which excludes investment: hence it now measures the overall effect of the quality of policy, including effects that work via investment. Figure 6 shows how this approach modifies the previous diagram. Again, the counterfactual distribution lies to the right. As might be expected, the effect of policy has become stronger, and continues to be observed throughout the distribution.

Our growth regressions include a role for initial income, and hence build in a model of the level of the steady-state growth path as in Mankiw, Romer and Weil (1992). Under the maintained assumption that all countries grow at the same rate in a long-run equilibrium, we can use the estimated coefficients for 1970-99 to compute the steady-state distribution of income per capita. As well as inferring the steady-state distribution of income per capita implied by the growth regression, we can also construct counterfactual steady-state distributions that would obtain if all countries shared the same level of the policy indicator.

Figure 7 shows the actual and counterfactual steady-state distributions of log output. Note that the actual distributions are not necessarily expected to have the familiar 'twin peaks' pattern, because our sample is restricted to developing countries. Figure 7 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 figure 8 we extend the analysis by taking into account the effect of $RMACRO$ on investment. We first run a simple regression of the logarithm of investment on initial income, initial human capital, regional dummies and $RMACRO$, and then use this regression to calculate a counterfactual investment rate under the assumption of high-quality policy. This altered investment rate is used in constructing the steady-state distribution shown in figure 8. Relative to figure 7, the counterfactual distribution is further to the right, reflecting the finding that better macroeconomic policy is associated with higher investment. Overall, the figures show how better macroeconomic policy could have substantial effects on the steady-state distribution of income levels.

8 Conclusions

This paper has re-examined the important question of how macroeconomic policy and growth are related in developing countries. The paper introduces a new index of the quality of macroeconomic policy, based on aggregating five policy indicators using an outlier-robust version of principal components analysis. By relating growth rates to the new index, we show that growth is positively associated with the quality of policy, and the effect is sizeable: a one-standard-deviation change in the index is found to raise the annual growth rate by somewhere between 0.5 and 0.7 percentage points over a thirty-year period.

We have also investigated the robustness of this effect using recently introduced Bayesian methods. Consistent with previous work on this topic, the evidence that our specific measure of policy matters for growth, when we allow for a wide range of other growth determinants, is not always robust. The strength of the evidence depends on the sample of countries, and particularly on the exclusion or inclusion of other proxies for government policy in the set of candidate predictors. But taken as a whole, the evidence presented in this paper suggests that government policy does help to explain differences in growth rates among developing countries.

9 Appendix 1

This appendix briefly discusses the Dollar (1992) measure of outward orientation, which has sometimes been criticized, at least as a measure of trade policies. One issue is whether Dollar's procedures can reliably control for the determinants of non-tradeables prices. This has been discussed by Falvey and Gemmell (1998, 1999). They suggest that Dollar's approach can be a reasonable approximation on average, the main exceptions occuring when the GDP per capita of a country is a weak proxy for its relative factor endowments.

Assuming for now that the Dollar procedure is effective in modelling non-tradeables prices, a remaining question is whether differences in tradeables prices should be attributed to trade restrictions, or to other factors.

Rodriguez and Rodrik (2000) provide an especially useful discussion of the strict assumptions 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.

In background work for this paper, we have found that the partial correlations between Dollar's index, geographic characteristics and measures of market access are generally fragile, but the index is strongly correlated with tropical location, for reasons that are not immediately clear. There is no obvious reason why closeness to the equator should be associated with unusually high transport costs, and more direct measures of market access lack explanatory power. This suggests that the correlation between Dollar's index and tropical location is driven by something other than the geography of transport costs, perhaps the association between tropical location and relatively weak institutions that was identified by Hall and Jones (1999) and Acemoglu et al. (2001).

The $OVERVALU$ variable in our paper is essentially Dollar's measure for 1976-85 extended forwards and backwards using real exchange rate movements, and then averaged. 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 used in section 3 of the paper. If we recalculate the principal components for four indicators rather than five, we obtain the following index:

$$
\begin{aligned}
MACROND &= 0.332 * SURPLUS - 0.516 * INFLA \\
&\quad -0.615 * BMP - 0.495 * ERATE
\end{aligned} \tag{5}
$$

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 will all be robust 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 clearer interpretation, the black market exchange rate premium, *BMP*.

10 Appendix 2

In this appendix we briefly discuss the method used to obtain the kernel density plots in the paper. The plots we present use the Epanechnikov kernel but, in general, density estimates are not sensitive to the precise choice of kernel. The choice of bandwidth is more important, and especially so in our application, given the small number of observations.

We start by estimating each density using the bandwidth that would minimize the mean integrated square error if the data were Gaussian and a Gaussian kernel were used (this is the default setting for the `kdensity` command in Stata 8.2). This is well-known to sometimes lead to oversmoothing and can obscure important structure - for example, it can make a bimodal distribution appear unimodal. To investigate this we then repeatedly lower the halfwidth of the kernel (the width of the density window around each point) to see if there is any further structure. The final plots that we present in the paper use a halfwidth that is chosen to give a reasonably smooth density plot without obscuring structure such as distinct modes. Given the small sample sizes we are often using, we prefer this approach to automated methods of bandwidth selection.

11 Appendix 3

The 72-country sample in our BMA uses imputed data for a small set of variables. We use a regression-based approach to imputation, in which data on other variables is used to estimate a regression for the variable of interest. Where the variable is missing for a particular country, the empty cell is filled using the fitted value for that cell. Our default predictors are three regional dummies from the SDM data set (for sub-Saharan Africa, Latin America and East Asia) and absolute latitude; occasionally other SDM variables are used, depending on the context. The full details of the predictors used for specific variables are shown in Appendix Table 4.

We now discuss the computational aspects of BMA in more detail. A key problem in applying these methods is the vast range of possible models. For example, with 30 candidate predictors, there are more than a thousand million possible models (the exact figure is $2^{30} \approx 1.074 \times 10^{9}$). Thus, most applications of BMA to sizeable data sets do not average over all possible models, but use a search algorithm to identify the subset of models with greatest relevance. We use two methods to establish this subset. The first is the Occam's Window technique described in Madigan and Raftery (1994) and Raftery et al. (1997). This excludes from the averaging procedure any model that is much less likely than the model with the highest posterior model probability. Our application of this excludes all models that have a posterior model probability lower than 1/100 the posterior model probability of the leading model. Hence, models that fall into this category are treated as if their posterior model probability can be rounded down to zero.

This tends to reduce massively the number of models used in the averaging process, but does not in itself solve the problem of identifying the models that are likely to lie within Occam's Window. In the case of linear regression, however, a branch-and-bound algorithm can be used to identify quickly a set of leading models; see Miller (2002, p. 53-55) for a description of how this works. To implement this procedure, we use a version of the `bicreg` software written for the `R` statistical language. The `bicreg` code was originally written for the language `S` by Adrian Raftery and revised by Chris Volinsky, and then modified for `R` by Ian Painter. The code we use is available online at http://www.research.att.com/~volinsky/bma.html

Some of the BMA results mentioned in the main text make use of an alternative approach. This uses the `MC3.REG` code written by Jennifer Hoeting and again translated to `R` by Ian Painter. Rather than using a model selection algorithm to identify leading models, this code visits different models using Markov Chain Monte Carlo methods, as in Madigan and York (1995), and uses the results of the chain to compute posterior model probabilities. The approach is described in more detail in Hoeting et al. (1996). Our implementation of this approach uses 40,000 iterations of the sampler and sets the parameter π to 0.10, which corresponds to the probability any given observation is an outlier. (Hoeting et al. only use a value as high as 0.10 for samples below 50 observations, but we judge outliers to be likely in this particular application, especially given the skewness of some of the candidate predictors.) All other parameter settings follow the default choices recommended by Hoeting et al. (1996).

References

[1] Acemoglu, D., S. Johnson, and J. Robinson (2001): "Colonial origins of comparative development: an empirical investigation," *American Economic Review,* 91, 1369-1401.

[2] Acemoglu, D., S. Johnson, J. Robinson, and Y. Thaicharoen (2003): "Institutional causes, macroeconomic symptoms: volatility, crises, and growth" *Journal of Monetary Economics,* 50.

[3] Barro, Robert J. and Jong-Wha Lee (2001). "International Data on Educational Attainment: Updates and Implications," *Oxford Economic Papers*, 53(3): 541-63.

[4] Berr, E., Combarnous, F. and Rougier, E. (2005). "Too much consensus could be harmful: measuring the degree of implementation of the Washington consensus and its impact on economic growth," Working paper no. DT/115/2005, Universite Montesquieu-Bordeaux IV.

[5] Bleaney, M. (1996). "Macroeconomic stability, investment and growth in developing countries," *Journal of Development Economics*, 48, 461-477.

[6] Brock, W. A., Durlauf, S. N. and West, K. D. (2003). "Policy evaluation in uncertain economic environments," *Brookings Papers on Economic Activity*.

[7] Burnside, Craig and David Dollar (2000). "Aid, policies, and growth", *American Economic Review*, 90(4), 847-68.

[8] Cook, Paul and Yuchiro Uchida (2003), "Privatisation and economic growth in developing countries", *Journal of Development Studies*, 39(6), 121-54.

[9] Desdoigts, A. (1996). "Smoothing techniques applied to a key economic issue: the 'convergence' hypothesis," *Computational Statistics*, 11, 481-494.

[10] Desdoigts, A. (2004). "Neoclassical convergence versus technological catch-up: a contribution for reaching a consensus," *Problems and Perspectives in Management*, 3, 15-42.

[11] DiNardo, John, Fortin, N. M. and Lemieux, T. (1996). "Labor Market Institutions and the Distribution of Wages, 1973-1992: A Semiparametric Approach," *Econometrica*, 64(5): 1001-44.

[12] Dollar, David (1992), "Outward-oriented developing economies really do grow more rapidly: Evidence from 95 LDCs, 1976-1985", *Economic Development and Cultural Change*, 40(3), 523-44.

[13] Easterly, William (2001a). *The elusive quest for growth: economists' adventures and misadventures in the tropics.* MIT Press, Cambridge MA.

[14] Easterly, William (2001b). "The Lost Decades: Developing Countries' Stagnation in Spite of Policy Reform 1980-1998," *Journal of Economic Growth*, 6(2): 135-57.

[15] Easterly, William (2002). "How did heavily indebted poor countries become heavily indebted? Reviewing two decades of debt relief," *World Development*, 30(10), 1677-1696.

[16] Easterly, William (2005). "National policies and economic growth: a reappraisal," In P. Aghion and S. N. Durlauf (eds.) *Handbook of Economic Growth*, Volume 1A, North-Holland: Amsterdam, 2005, pp. 1015-1059.

[17] Easterly, William, Kremer, Michael, Pritchett, Lant and Lawrence Summers (1993). "Good Policy or Good Luck? Country Growth Performance and Temporary Shocks," *Journal of Monetary Economics*, 32(3): 459-83.

[18] Easterly, William and Levine, Ross (2003). "Tropics, Germs, and Crops: How Endowments Influence Economic Development," *Journal of Monetary Economics*, 50(1): 3-39.

[19] Easterly, William, Levine, Ross and David Roodman (2004). "Aid, Policies, and Growth: Comment," *American Economic Review*, 94(3): 774-80.

[20] Easterly, William and Mirvat Sewadeh (2002), Global Development Network Growth Database, World Bank. http://www.worldbank.org/research/growth/GDNdata.htm

[21] Falvey, R. and Gemmell, N. (1998). "Why are prices so low in Asia?" *World Economy*, 21(7): 897-911.

[22] Falvey, R. and Gemmell, N. (1999). "Factor endowments, nontradables prices and measures of 'openness' ", *Journal of Development Economics*, 58, 101-122.

[23] Fernandez, C., E. Ley, and M. F. J. Steel (2002): "Model uncertainty in cross-country growth regressions," *Journal of Applied Econometrics*, 16(5), 563-576.

[24] Fischer, S. (1991). "Growth, Macroeconomics, and Development," in Blanchard, Olivier Jean and Fischer, Stanley (1991). *NBER Macroeconomics Annual* 1991. Cambridge and London: MIT Press, 329-64

[25] Fischer, S. (1993). "The Role of Macroeconomic Factors in Growth," *Journal of Monetary Economics*, 32(3): 485-512.

[26] Fischer, S. (2003). "Globalization and Its Challenges," *American Economic Review*, 93(2): 1-30.

[27] Flury, B. and Riedwyl, H. (1988). *Multivariate statistics: a practical approach.* Chapman and Hall, London.

[28] Glaeser, E. L., La Porta, R., Lopez-de-Silanes, F. and Shleifer, A. (2004). "Do institutions cause growth?" *Journal of Economic Growth*, 9(3), 271-303.

[29] Hall, R. E., and C. I. Jones (1999): "Why do some countries produce so much more output per worker than others?", *Quarterly Journal of Economics,* February, 83-116.

[30] Hausmann, R., Pritchett, L. and Rodrik, D. (2005). "Growth accelerations," *Journal of Economic Growth*, 10(4), 303-329.

[31] Hausmann, R., Rodrik, R. and Velasco, A. (2005). Growth diagnostics. Manuscript, Harvard University.

[32] Henisz, Witold (2000), "The institutional environment for economic growth", *Economics and Politics*, 12(1), 1-31.

[33] 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.

[34] Hoeting, J., Raftery, A. E. and Madigan, D. (1996). "A method for simultaneous variable selection and outlier identification in linear regression," *Computational Statistics and Data Analysis*, 22, 251-270.

[35] Hubert, Mia, Rousseeuw, Peter J. and Vanden Branden, Karlien (2005). "ROBPCA: a new approach to robust principal component analysis," *Technometrics*, 47, 64-79.

[36] Kaufmann, Daniel, Aart Kraay and Massimo Mastruzzi (2005). "Governance matters IV: Governance indicators for 1996-2004," Manuscript, the World Bank.

[37] Leamer, E. E. (1978): S*pecification Searches. Ad-Hoc Inference with Non-Experimental Data*, John Wiley, New York.

[38] Levine, R., and D. Renelt (1992): "A sensitivity analysis of cross-country growth regressions", *American Economic Review,* 82(4), 942-63.

[39] Madigan, D.M. and A. E. Raftery (1994): "Model selection and accounting for model uncertainty in graphical models using Occam's Window", *Journal of the American Statistical Association*, 89, 1335-1346.

[40] Madigan, D. and York, J. (1995): "Bayesian Graphical Models for Discrete Data", *International Statistical Review,* 63, 215 232.

[41] Malik, A. and Temple, J. R. W. (2005). The geography of output volatility. CEPR discussion paper no. 5516.

[42] Mankiw, N. Gregory, David Romer and David N.Weil (1992). "A contribution to the empirics of economic growth", *Quarterly Journal of Economics*, 107, 407-37.

[43] Marshall, Monty and Keith Jaggers (2002), Polity IV project: Political regime characteristics and transitions, 1800-2002, Center for International Development and Conflict Management, University of Maryland, College Park.

[44] Miller, A. (2002). *Subset selection in regression.* Chapman and Hall, London (second edition).

[45] Quah, D. T. (1996). "Twin peaks: Growth and convergence in models of distribution dynamics," *Economic Journal*, 106(437), 1045-1055.

[46] Raftery, A.E. (1995): "Bayesian model selection in social research", In *Sociological Methodology* (Peter V. Marsden, ed.), Blackwells, Cambridge, pp. 111-196.

[47] Raftery, A. E., D. Madigan, and J. E. Hoeting (1997): "Bayesian model averaging for linear regression models," *Journal of American Statistical Association*, 92, 179-191.

[48] Rodriguez, F. and Rodrik, D. (2000). Trade Policy and Economic Growth: A Skeptic's Guide to the Cross-National Evidence. in Bernanke, B. S. and Rogoff, K. (eds.) *NBER Macroeconomics Annual* 2000. MIT Press, 261-325.

[49] Rodrik, Dani (1999). "Where Did All the Growth Go? External Shocks, Social Conflict, and Growth Collapses," *Journal of Economic Growth*, 4(4): 385-412.

[50] Rodrik, Dani (2002). "After Neoliberalism, what?", Unpublished paper.

[51] Rodrik, Dani (2005). "Why we learn nothing from regressing economic growth on policies," Manuscript, Harvard.

[52] Rousseeuw, Peter J. (1984). "Least median of squares regression," *Journal of the American Statistical Association*, 79, 871-880.

[53] Rousseeuw, P. J. and van Driessen, K. (1999). A fast algorithm for the minimum covariance determinant estimator. *Technometrics*, 41(3), 212-223.

[54] Sachs, Jeffrey and Andrew S. Warner (1995), "Economic reform and the process of global integration", *Brookings Papers on Economic Activity* 1, 1-118.

[55] Sala-i-Martin, X. (1991). Comments on Fischer (1991), "Growth, Macroeconomics, and Development," in Blanchard, Olivier Jean and Fischer, Stanley (1991). *NBER Macroeconomics Annual* 1991. Cambridge and London: MIT Press.

[56] Sala-i-Martin, X., Doppelhofer, G. and R. I. Miller (2004). "Determinants of long-term growth: a Bayesian averaging of classical estimates (BACE) approach," *American Economic Review*, 813-835.

[57] Schwarz, G. (1978): "Estimating the dimension of a model", *Annals of Statistics*, 6, 461-464.

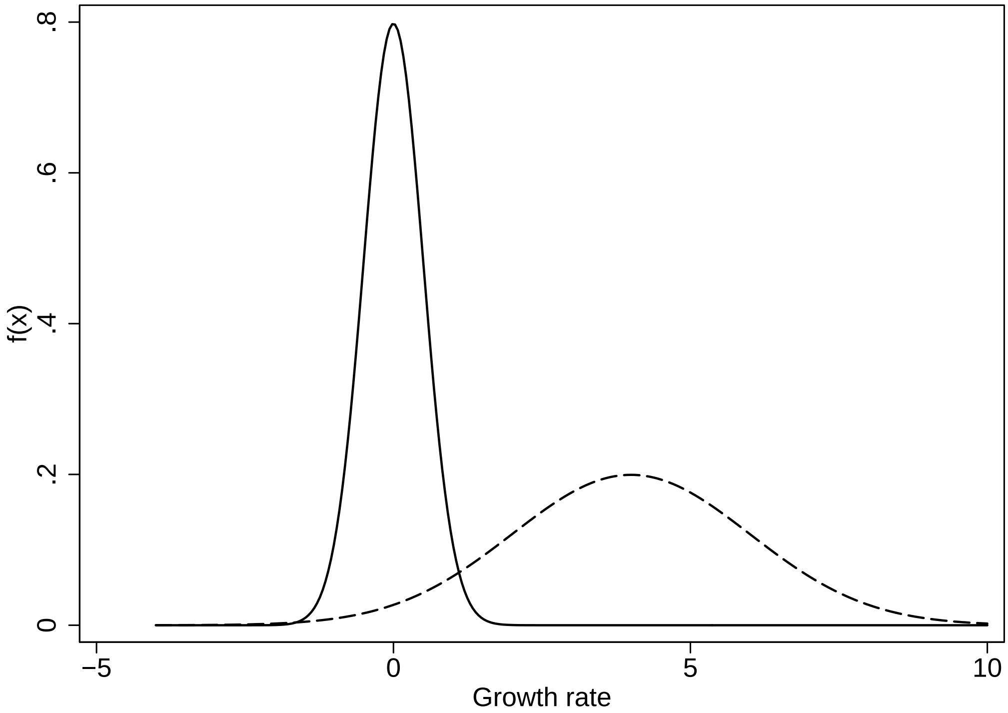
[58] Temple, Jonathan R. W. (1998). "Equipment Investment and the Solow Model," *Oxford Economic Papers*, 50(1): 39-62.

[59] Temple, Jonathan R. W. (2000). "Growth regressions and what the textbooks don't tell you," *Bulletin of Economic Research*, 52(3), 181-205.

[60] Williamson, John (1990). "What Washington means by policy reform" in John Williamson (ed.) *Latin American Adjustment: How Much Has Happened?* Institute for International Economics, Washington D.C.

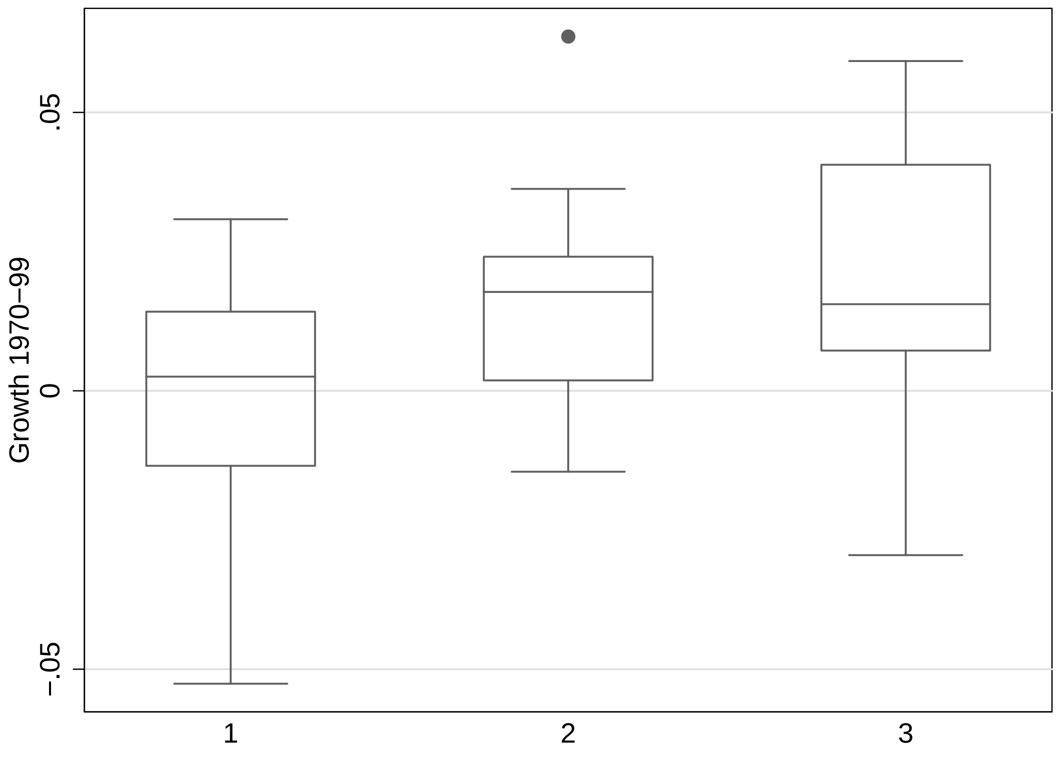
[61] World Bank (2004). *World Development Indicators 2004*, World Bank. CD-ROM version.

Figure 1 – Hypothetical distributions of growth rates



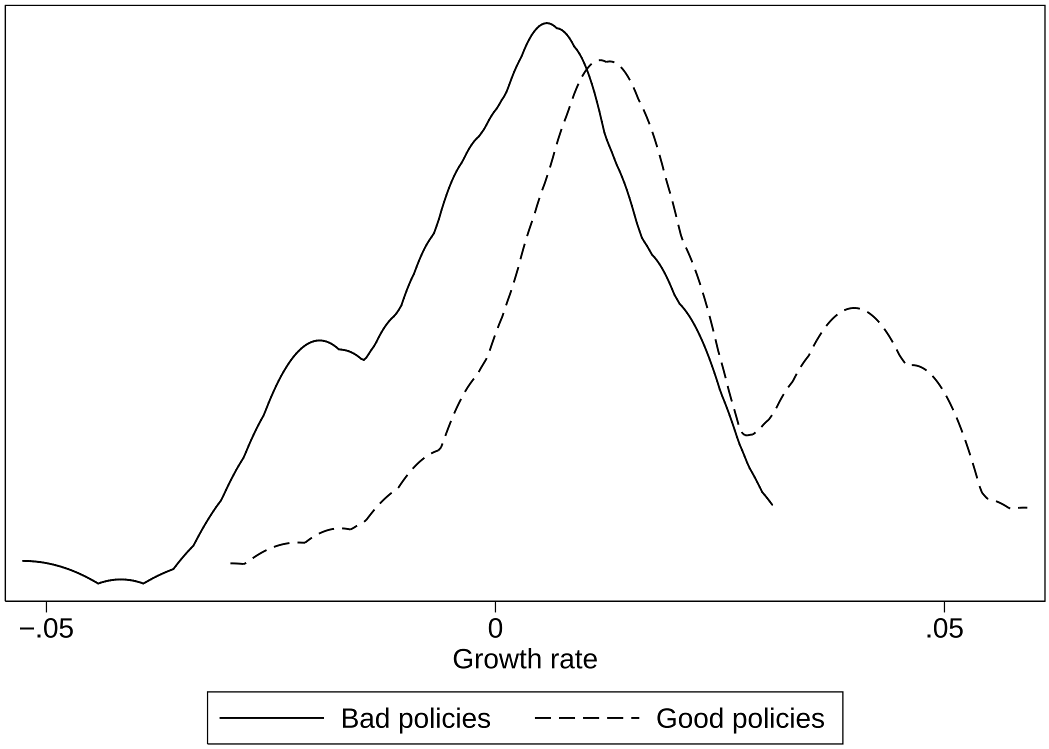
The solid line shows a hypothetical distribution for countries with bad policies, the dashed line for countries with good policies.

Figure 2 – Box-plots for growth rates, where countries are grouped by policy



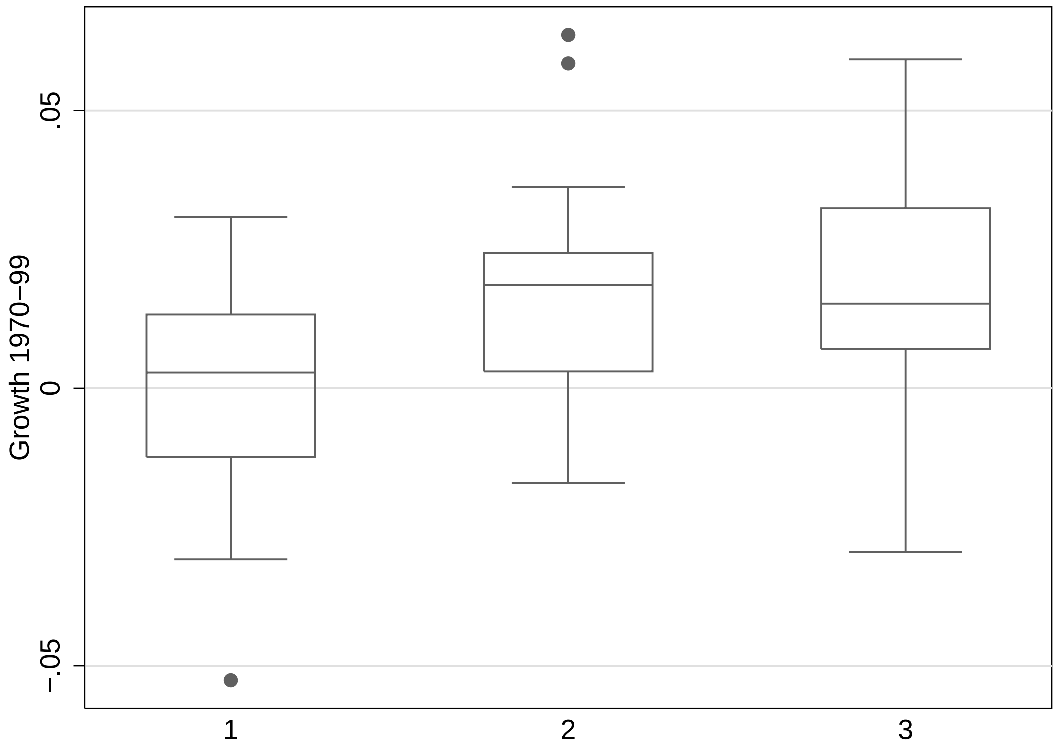
This graph shows three Tukey box-plots for three groups, from bad policy (group 1) to good policy (group 3). The horizontal line in the box indicates the median growth rate for that group. The policy classification is based on RMACRO.

Figure 3 – Kernel density plots



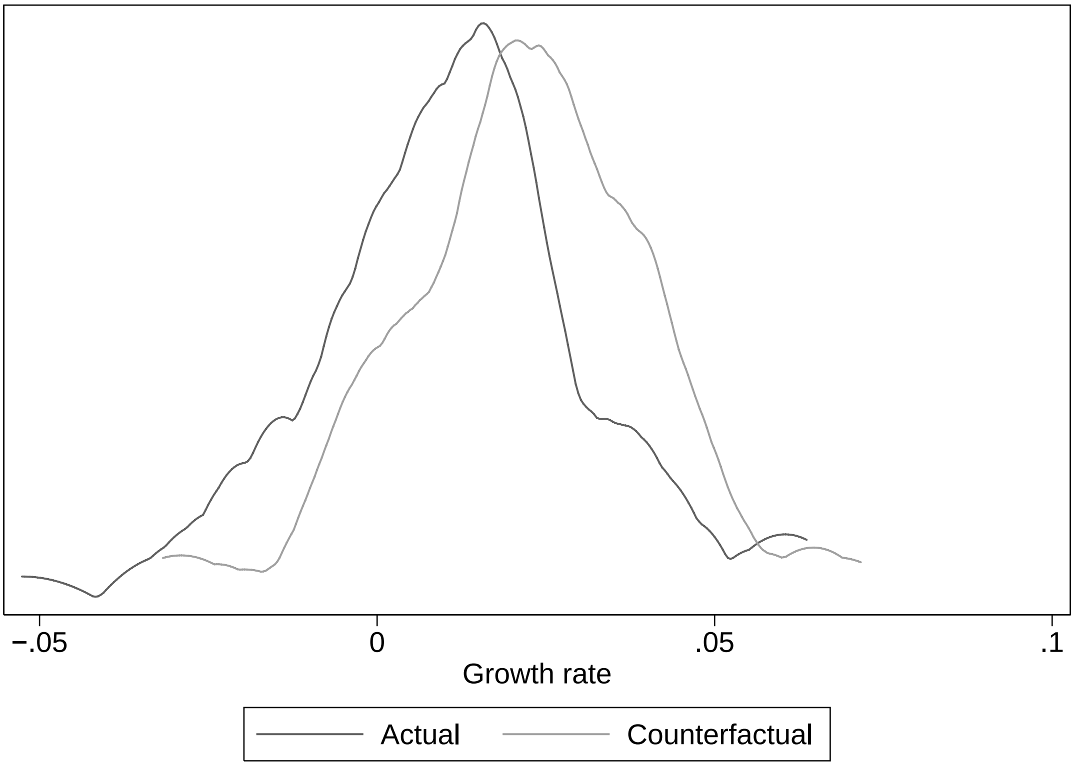
The solid line shows the distribution for countries with bad policies, the dashed line for countries with good policies. The policy classification is based on RMACRO.

Figure 4 – Box-plots for growth rates, groups based on BMP



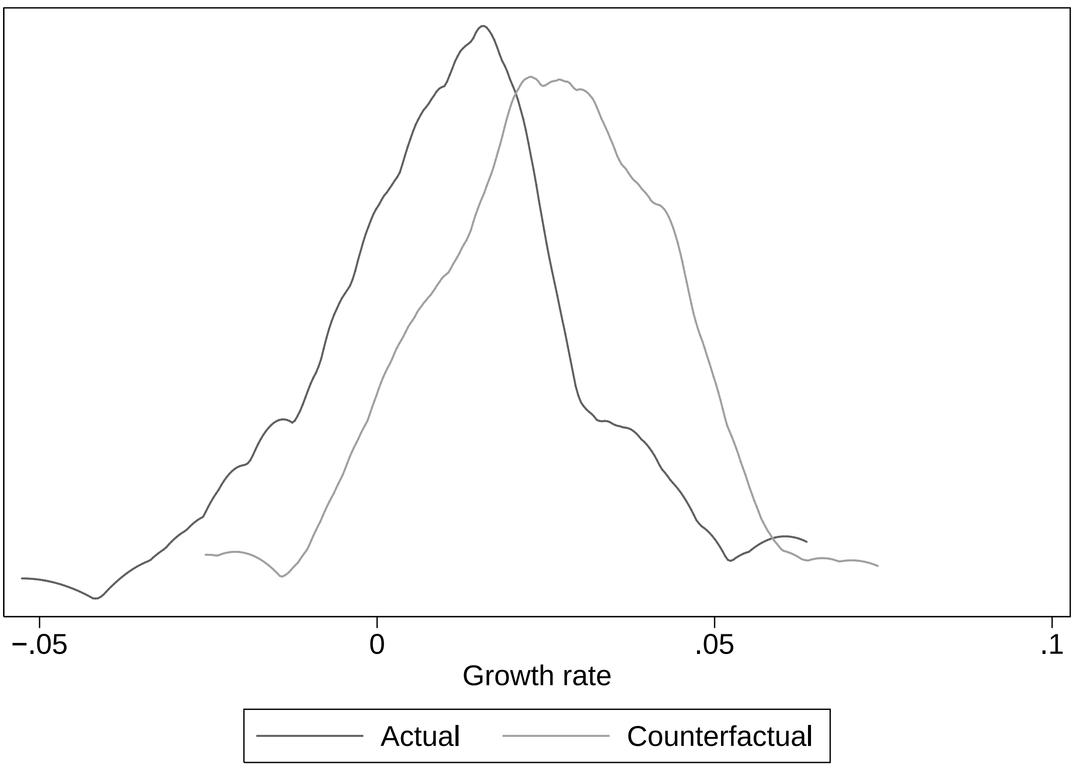
This graph shows three Tukey box-plots for three groups, from bad policy (group 1) to good policy (group 3). The horizontal line in the box indicates the median growth rate for that group. The policy classification is based on the black market premium, BMP.

Figure 5 – Actual and counterfactual distribution of growth rates



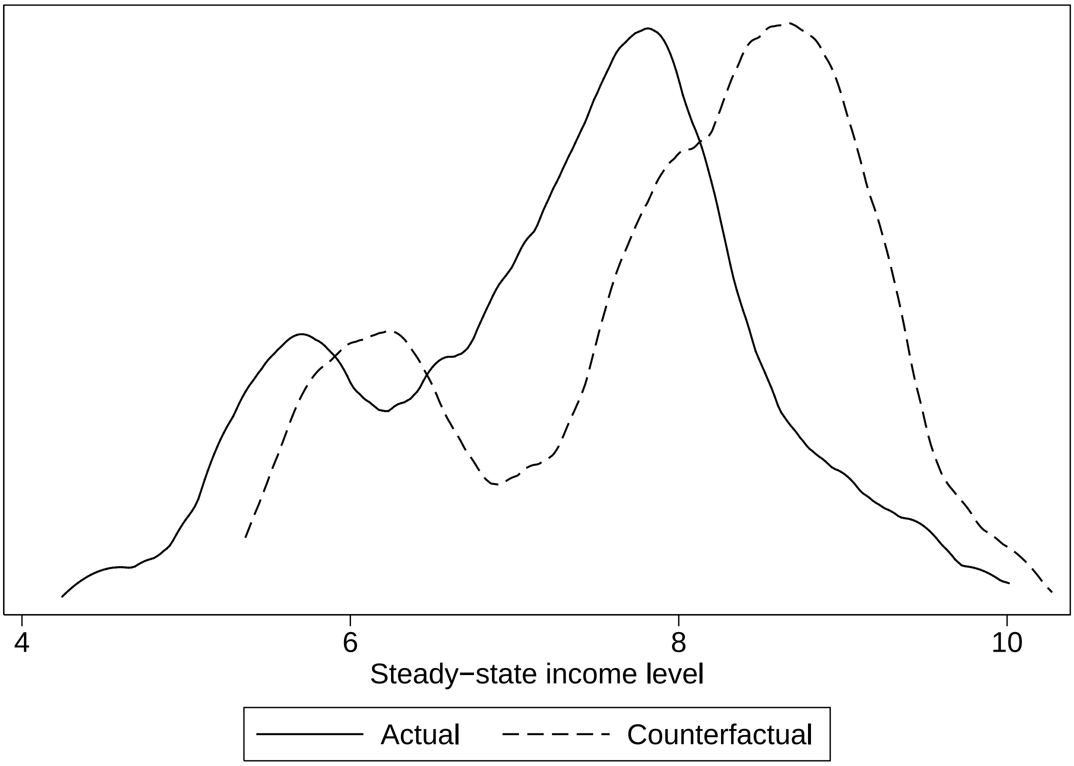
The solid line shows the actual distribution, the dashed line the counterfactual.

Figure 6 – Actual and counterfactual distribution of growth rates (including investment effects)



The solid line shows the actual distribution, the dashed line the counterfactual.

Figure 7 –Steady-state income levels (in logs)



The solid line shows the actual distribution, the dashed line the counterfactual.

Figure 8 – Steady-state distribution (in logs) with investment effects

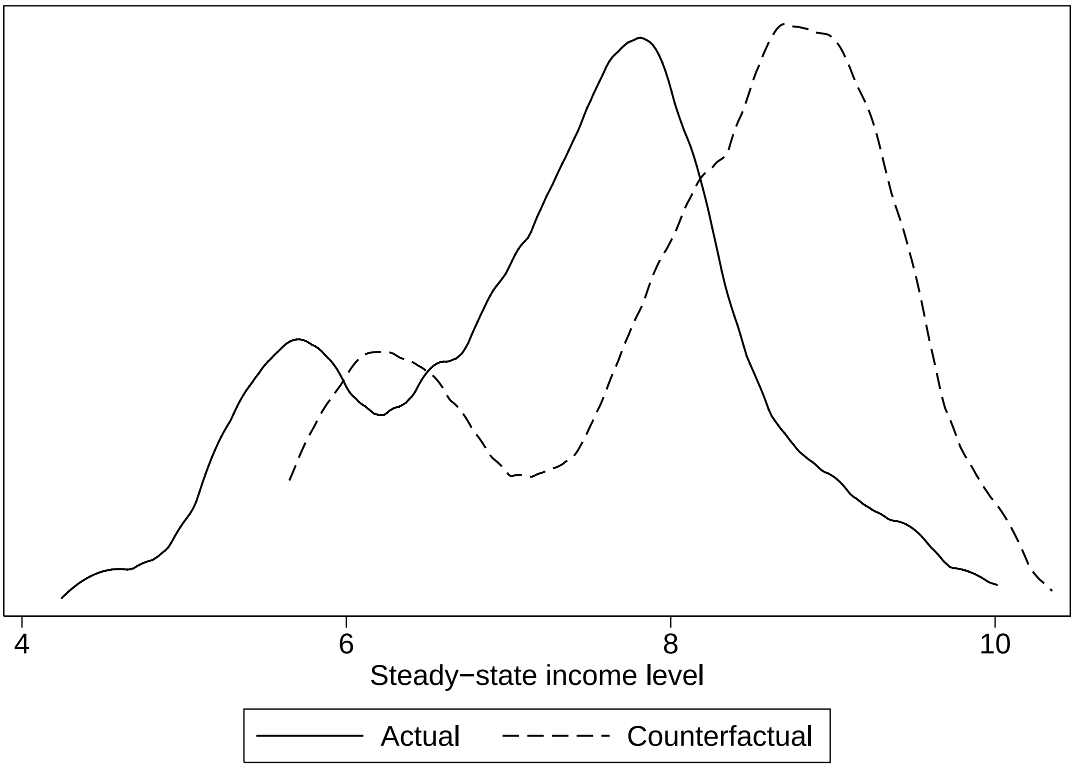


Table 1: Principal Component Analysis

Variable	Expected sign	(1) MACRO 1st PC	(1) MACRO 2nd PC	(2) RMACRO 1st PC	(2) RMACRO 2nd PC	(3) MACROOL 1st PC	(3) MACROOL 2nd PC
SURPLUS	+	0.484	**0.579**	**0.340**	0.297	0.276	**0.768**
INFLA	-	**-0.647**	0.437	**-0.744**	0.172	**-0.727**	0.161
BMP	-	**-0.848**	0.184	**-0.888**	-0.034	**-0.843**	0.120
OVERVALU	-	-0.503	**-0.633**	-0.395	**-0.951**	-0.327	**-0.654**
ERATE	-	**-0.688**	0.232	**-0.653**	-0.164	**-0.665**	0.311
Number of countries		78		78		75	
% 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 (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: Simple correlations among GDP growth
and various aggregated macroeconomic policy indices

	RGDP7099C	MACRO	RMACRO	MACROOL	BD	ELR7097
RGDP7099C	1.0000					
MACRO	0.4715	1.0000				
RMACRO	0.4194	0.9759	1.0000			
MACROOL	0.4087	0.9952	0.9913	1.0000		
BD	0.6618	0.6658	0.6226	0.5848	1.0000	
ELR7097	0.5570	0.6031	0.6208	0.6450	0.8498	1.0000

Table 3 – Growth Regressions

Column	(1)	(2)	(3)	(4)	(5)
Observations	70	70	70	70	60
RMACRO		0.71	0.49	0.64	0.64
		(0.30)	(0.31)	(0.27)	(0.29)
Initial income	-1.10	-0.26	-0.80	-1.04	-1.15
	(0.37)	(0.37)	(0.37)	(0.38)	(0.42)
Population growth	-0.21		-0.19	-0.12	-0.10
	(0.23)		(0.22)	(0.25)	(0.28)
Investment	1.07		1.10	0.83	0.84
	(0.32)		(0.34)	(0.32)	(0.48)
Literacy	0.68			0.88	
	(0.31)			(0.34)	
Average schooling					0.79
					(0.27)
R^2	0.51	0.37	0.51	0.57	0.55
Regression s.e.	1.56	1.75	1.57	1.47	1.58
Heteroscedasticity					
Breusch-Pagan	0.32	0.02	0.07	0.27	0.18
White	0.66	0.19	0.03	0.64	0.35
Ramsey RESET	0.90	0.58	0.02	0.68	0.24

MacKinnon-White heteroskedasticity-consistent (hc3) standard errors reported in parentheses. 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, and so the coefficients represent the effect of a one-standard-deviation change on the annual growth rate. All regressions include five regional dummies, for East Asia and the Pacific; Middle East and North Africa; South Asia; Sub-Saharan Africa; and Latin America and the Caribbean. Constant and coefficients on regional dummies not reported. ‘Heteroscedasticity’ reports p-values associated with two tests for heteroscedasticity. ‘Ramsey RESET’ is the p-value associated with a RESET test.

Appendix Table 1: List of countries

Latin America and Caribbean	East Asia and the Pacific
Argentina	Fiji
Bolivia	Indonesia
Brazil	Malaysia
Chile	Papua New Guinea
Colombia	Philippines
Costa Rica	Singapore
Dominican Republic	South Korea
Ecuador	Thailand
El Salvador	
Guatemala	**Sub-Saharan Africa**
Guyana	Benin
Haiti	Botswana
Honduras	Burkina Faso
Jamaica	Burundi
Mexico	Cameroon
Nicaragua	Central African Republic
Panama	Chad
Paraguay	Congo, Democratic Republic
Peru	Congo, Republic
Trinidad and Tobago	Ethiopia
Uruguay	Gabon
Venezuela	Ghana
	Kenya
Middle East and North Africa	Lesotho
Algeria	Liberia
Cyprus	Madagascar
Egypt	Malawi
Iran	Mali
Israel	Mauritania
Jordan	Mauritius
Morocco	Niger
Oman	Nigeria
Syria	Rwanda
Tunisia	Senegal
Turkey	Sierra Leone
Yemen	Somalia
	Sudan
South Asia	Togo
Bangladesh	Uganda
India	Zambia
Nepal	Zimbabwe
Pakistan	
Sri Lanka	

Appendix Table 2: Variables and definitions

Variable	Description	Sources
INVEST	Natural logarithm of investment share in GDP, 1970-99	Penn World Table 6.1
POPG	Natural logarithm of average annual growth rate of population aged 15-64, 1970-99, plus 0.05.	World Bank (2004a)
SCHOOL70	Natural logarithm of average years of schooling at all educational levels of population aged over 15 in 1970.	Barro and Lee (2000)
LITERACY	Natural log of (100-illiteracy rate of population aged over 15 in 1970)	World Bank (2004a)
RGDPPC70	Natural logarithm of real GDP per capita (rgdpch) in 1970.	Penn World Table 6.1
RGDP7099C	Natural logarithm of real GDP per capita (rgdpch) in 1999 minus same variable for 1970. This is divided by 29, to obtain annual growth rates.	Penn World Table 6.1
RGDP7099W	Natural log of real GDP per worker (rgdpwok) in 1999 minus that of 1970. This is divided by 29, to obtain annual growth rates.	Penn World Table 6.1
Regional dummies	Five regions: East Asia and the Pacific (RGNEAP), Middle East and North Africa (RGNMENA), South Asia (RGNSA), Sub-Saharan Africa (RGNSSA), and Latin America and Caribbean (RGNLAC)	Easterly and Sewadeh (2002)
MACRO	The first principal component from a classical principal components analysis of SURPLUS, INFLA, BMP, OVERVALU and ERATE. Higher values mean better policy outcomes.	see text
RMACRO	As above, but from a robust principal components analysis.	see text
SURPLUS	Mean central government budget surplus as a share of GDP, 1970-99	World Bank (2004a)
DEBT	Natural log of mean central government debt over GDP, 1970-99	World Bank (2004a)
INFLA	Natural log of (1+inflation rate based on median GDP deflator)	World Bank (2004a)
REALI	Mean lending rate adjusted by GDP deflator.	World Bank (2004a)
BMP	Natural log of (1+mean black market premium)	Easterly and Sewadeh (2002)
OVERVALU	Natural log of mean overvaluation index. Dollar (1992) provides data for 1976-85. Easterly and Sewadeh (2002) update the data to 1999.	Dollar (1992) and Easterly and Sewadeh (2002)
ERATE	Variation of the Dollar real exchange rate measure around its mean.	Dollar (1992)
POLITY	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.	Marshall and Jaggers (2000)
POLCON	Extent of political constraints in policy-making process. Higher value imply stronger constraints. Mean value 1970-1999.	Henisz (2000)
EXPRISK	Protection against expropriation risk. Higher values mean lower risk. Mean value 1985-1995.	Acemoglu, Johnson and Robinson (2001)
GOVKKZ	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.	Kaufmann, Kraay and Mastruzzi (2005)

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.

Appendix Table 3: Descriptive statistics

Variable	Observation	Mean	Std.Dev.	Min	Max
RGDP7099C	77	0.0116	0.0210	-0.0526	0.0636
RGDP7099W	76	0.0098	0.0207	-0.0564	0.0678
MACRO	78	0.0000	1.0000	-3.2762	1.9651
RMACRO	78	0.0000	1.0000	-2.9744	1.8367
MACROOL	75	0.0000	1.0000	-2.4166	2.0575
SURPLUS	88	-4.1989	5.4175	-26.9389	16.2312
DEBT	71	3.7845	0.7485	2.0096	5.6481
INFLA	90	2.4243	0.7430	0.9092	4.6012
REALI	85	7.1370	13.0313	-38.9671	71.0281
BMP	89	2.9829	1.6887	-0.0363	7.8161
OVERVALU	82	4.7428	0.3322	4.0934	5.9453
ERATE	81	0.1620	0.0987	0.0400	0.5000
SCHOOL70	67	0.8184	0.7848	-1.6190	2.0910
LITERACY	86	3.6792	0.6941	1.7487	4.5559
INVEST	86	2.4652	0.5540	0.7966	3.8144
RGDPPC70	79	6.4498	0.6971	5.1888	7.9270
POPG	90	-2.5715	0.0901	-2.8942	-2.3915
RGNEAP	90	0.1333	0.3418	0.0000	1.0000
RGNECA	90	0.0222	0.1482	0.0000	1.0000
RGNMENA	90	0.1333	0.3418	0.0000	1.0000
RGNSA	90	0.0556	0.2303	0.0000	1.0000
RGNSSA	90	0.4000	0.4926	0.0000	1.0000
RGNLAC	90	0.2444	0.4322	0.0000	1.0000
BDDATA	73	1.2277	0.9196	-1.0254	3.6353
ELRBDC7097	75	1.5146	0.8546	-0.7699	3.3234
ELR7097	75	1.4320	0.7910	-0.9231	3.2636

Appendix Table 4 – Imputation of missing cases in the SDM data

SDM control variables	Variables used in predictive regression	Number of imputed observations
Log investment price, 1960-64	ABSLATIT, EAST, LAAM, SAFRICA	1
Fraction of area in tropics	ABSLATIT, EAST, LAAM, SAFRICA	3
Log population density (coast), 1960s	LDENS60, LANDAREA, LANDLOCK, ABSLATIT, EAST, LAAM, SAFRICA	3
Malaria prevalence, 1960s	ABSLATIT, EAST, LAAM, SAFRICA	5
Ethnolinguistic fractionalization	ABSLATIT, EAST, LAAM, SAFRICA	1
Government share of GDP, 1960-64	GVR61, ABSLATIT, EAST, LAAM, SAFRICA	1
Fraction of population in tropics	ABSLATIT, EAST, LAAM, SAFRICA	3
Export share of primary exports in 1970	ABSLATIT, EAST, LAAM, SAFRICA	1
Government investment relative to GDP, 1970-74	RGDPPC70, GVR61, ABSLATIT, EAST, LAAM, SAFRICA	3

Missing data for the above variables in our BMA exercises are replaced with fitted values from simple regressions. The table shows the variables in the regressions and the number of imputed observations. The explanatory variables in each regression are the stated combination of absolute latitude (ABSLATIT), an East Asian dummy (EAST), a Latin American dummy (LAAM), a sub-Saharan Africa dummy (SAFRICA), the log of population density in 1960 (LDENS60), land area (LANDAREA), a dummy for landlocked countries (LANDLOCK), the ratio of government consumption to GDP in 1961 (GVR61), and real GDP per capita in 1970 (RGDPPC70).