Macroeconomic Stability and the Distribution of Growth Rates

Vatcharin Sirimaneetham and Jonathan R.W. Temple

It is often argued that macroeconomic instability can form a binding constraint on economic growth. Drawing on a new index of stability, threshold estimation is used to divide developing economies into two growth regimes, depending on a threshold level of stability. For the more stable group of countries, the output benefits of investment are greater, conditional convergence is faster, and measures of institutional quality have more explanatory power, suggesting that instability forms a binding constraint for the less stable group. Macroeconomic stability is also shown to dominate several other candidates for identifying distinct growth regimes. JEL codes: O23, O40

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

Although macroeconomic stability could be important for growth, the strength of the empirical relationship remains uncertain. One argument is that the observed correlation between stability and growth is mainly due to a few countries with the very worst macroeconomic outcomes. Once a certain

Vatcharin Sirimaneetham is a consultant in the Poverty Reduction and Economic Management Network, East Asia and the Pacific Region, World Bank; his email address is vsirimaneetham@worldbank.org. Jonathan R.W. Temple (corresponding author) is a professor of economics at the University of Bristol; his email address is jon.temple@bristol.ac.uk. The authors are grateful to the journal editor, three anonymous referees, David Ashton, Holger Breinlich, Edmund Cannon, Huw Dixon, Jan Fidrmuc, Max Gillman, Andreas Leukert, and Patrick Minford for related comments or discussion as well as to seminar participants at the University of Bristol, Cardiff University, the 2006 Royal Economic Society Conference at Nottingham, and a Brunel University workshop at the Centre for Economic Development and Institutions on aspects of growth and macroeconomic policy. Temple thanks the Leverhulme Trust for financial support under the Philip Leverhulme Prize Fellowship scheme. A supplemental appendix to this article is available at http://wber.oxfordjournals.org/.

THE WORLD BANK ECONOMIC REVIEW, VOL. 23, NO. 3, pp. 443–479 doi:10.1093/wber/lhp008
Advance Access Publication September 16, 2009

threshold level of stability has been achieved, the marginal benefits of additional stability could be minimal. Another argument, which dates to at least Sala-i-Martin (1991), is that macroeconomic disarray could be a symptom of deeper problems. Recent research, especially after the work of Acemoglu, Johnson, and Robinson (2001), Acemoglu and others (2003), and Easterly and Levine (2003), argues that macroeconomic policies lack explanatory power relative to institutions. But this is far from a consensus, and Henry and Miller's (2009) case study of two Caribbean islands presents a different view.

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

Although the practical analysis of growth policy is often framed in terms of necessary and sufficient conditions, incorporating this idea into empirical models is not straightforward. Another approach—similar in spirit, but more general—frames the problem in terms of binding constraints, as in recent work by Hausmann, Rodrik, and Velasco (2008) and Rodrik (2007). If the marginal effects of policies and other growth determinants are not independent, one or more constraints on growth may be binding, with reforms elsewhere having limited benefits, at least until the key constraints are addressed. This contrasts with the linear regressions usually adopted in the empirical growth literature, which implicitly assume that different growth determinants smoothly substitute for one another.

With all this in mind, this article explores methods designed to close the gap between the vocabulary of policy analysis and the empirical models used to explain growth variation. First, direct comparisons of growth rate distributions are used, with countries divided into groups based on an index of macroeconomic stability. These distributions clearly show that macroeconomic instability is not always a binding constraint. In particular, even when a country ranks low in terms of macroeconomic stability, this is not a sufficient condition for slow growth. But the highest long-run growth rates are confined to countries with stable macroeconomic outcomes.

The article then examines how regressions can accommodate the binding constraints view. Standard regressions are used to quantify the effects of macroeconomic stability over 1970–99, restricting the sample to developing economies. These linear models assume that any adverse effect of instability can be offset by other factors and thus that instability is never a binding constraint. To allow for instability as a binding constraint, threshold estimation, based on Hansen (1996, 2000), is used. The results indicate that the sample can be split into two groups by macroeconomic stability. For the more stable group of countries, the elasticity of steady-state output to the investment rate is greater, conditional convergence is faster, and the standard growth determinants of the Solow model (together with a measure of institutional quality) explain 75–90 percent of the cross-section variation in growth rates, a remarkably high proportion. For the less stable group, instability reduces growth, while the Solow variables have less explanatory power, investment is less effective, and the residual variance is much higher. Fundamentals such as good institutions are not strongly associated with growth unless macroeconomic stability is also in place. These results suggest that instability can indeed form a binding constraint on growth.

The analysis acknowledges an important criticism of past research: that policy outcomes are likely to be endogenous in both an economic and a statistical sense. Rodrik (2005) points out that observed policies are decision variables that must be endogenous to social and economic circumstances. The implication is that macroeconomic stability is not randomly assigned and will almost certainly be correlated with omitted country characteristics—and thus with the error term of the growth regression. When this problem arises in the microeconometric literature, the availability of control variables is often limited, but there may be plausible candidates for instrumental variables through "natural experiments." Growth researchers face almost the mirror image of that situation: there are many possible control variables but few plausible candidates for instruments.

This article uses two approaches. The first follows Barro (1996) in exploiting the observed association between French colonial heritage and macroeconomic stability, linked to the membership of many former French colonies in the CFA franc zone. This implies that French colonial heritage could be a suitable instrument, but it would not be difficult to criticize the necessary exclusion restriction. For example, French colonial heritage is likely to have influenced the legal system, with a variety of effects on development, a debate reviewed in La Porta, Lopez-de-Silanes, and Shleifer (2008).

The analysis therefore emphasizes an alternative approach that considers an unusually wide range of possible control variables, including various indicators of geographic characteristics and institutions. This comprehensive approach increases the chance of identifying controls that influence the extent of stability, in order to lessen the correlation between macroeconomic stability and the error term, even though macroeconomic stability is not randomly assigned. This relates to "selection-on-observables" from the treatment effects literature and is appropriate if the central endogeneity problem is omitted variables rather than simultaneity bias.[1] The approach is based on Bayesian methods for

1. Simultaneity bias is relevant if policy outcomes depend directly on growth outcomes, which may be plausible in the short run but less so over the 30 years considered here. It is more plausible that growth and policy outcomes are jointly influenced by other variables, such as institutions, hence this article's emphasis on the omitted variable problem rather than on simultaneity.

model averaging and thus addresses the model uncertainty problem highlighted by Levine and Renelt (1992). The evidence that stability matters varies with the sample of countries, but in the largest sample considered the estimated benefits of stability are robust across a wide range of specifications.

Finally, the results are used to construct counterfactual distributions of growth rates and steady-state levels of GDP per capita. These distributions indicate what might have happened had all developing economies achieved macroeconomic stability over 1970–99. To the extent that the estimated benefit of stability can be interpreted as a causal effect, the variation in stability exerts a major influence on the distributions of growth rates and steady-state GDP per capita. But it is important to acknowledge some major qualifications. As mentioned, macroeconomic instability may be a symptom of other problems. Instability may arise in the wake of conflict or relatively severe external shocks. The estimates are thus best interpreted as an upper bound on the importance of good macroeconomic management.

The article is organized as follows. Section I briefly reviews the literature on macroeconomic policy and growth and discusses the empirical analysis of binding constraints. Section II describes the new measure of stability. Section III looks at the relationship between stability and growth in a variety of ways, emphasizing threshold estimation. Section IV examines robustness using Bayesian methods. Section V uses the earlier growth regressions to generate counterfactual distributions of growth rates and steady-state levels of income. And section VI presents some implications of the findings.

I. The Literature on Macroeconomic Policy and Growth

Much of the literature on policy and growth has studied trade regimes and, more recently, such factors as entry barriers and regulation. But this article is about macroeconomic stability—not market-led development or the Washington Consensus. As initially summarized by Williamson (1990), the Washington Consensus reflected principles that went well beyond macroeconomic policies and included tax reform, financial and trade policy liberalization, openness to foreign direct investment, privatization, deregulation, and protection of property rights. Rather than investigate these, this article examines whether the Washington Consensus was right to emphasize the benefits of stable macroeconomic outcomes. Attempts to achieve stability can be controversial, especially when reductions in fiscal deficits are proposed. Moreover, it is rarely clear how much stability is "enough."[2]

2. This article does not address the subtler and much more difficult questions that relate to short-run policy activism such as demand management. The results concern macroeconomic outcomes (rather than policies) assessed over the long run and should be interpreted in that light; they do not imply, for example, that budget deficits must always be avoided.

Motivated by these considerations, empirical studies such as Bleaney (1996) and Fischer (1991, 1993) concluded that macroeconomic stability matters for sustained growth. More recent researchers are not so convinced. Macroeconomic policy outcomes have generally improved over time, while many developing countries grew more slowly during the 1980s and 1990s than they had previously. This led to the conclusion that the growth dividend of greater macroeconomic stability has been disappointing, an argument reviewed in Montiel and Servén (2006). The reasons behind the post-1980 growth collapse in developing economies are discussed in Easterly (2001b) and Rodrik (1999) and seem likely to go beyond macroeconomic policy decisions.

Other evidence casts further doubt on the role of stability. Improvements in policy indicators explain relatively few growth accelerations (Hausmann, Pritchett, and Rodrik 2005), and in general policy indicators are far more persistent than growth rates are, suggesting that policy will usually leave the medium-run variation in growth unexplained (Easterly and others 1993). Perhaps most fundamental, empirical studies such as 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. Easterly (2005, p. 1055) concludes that "the long-run effect of policies on development is difficult to discern once you also control for institutions." This highlights a problem in the empirical literature: that economic disarray usually extends across a range of outcomes. It can be hard to disentangle the effects of specific macroeconomic outcomes from one another and from other growth determinants. Perhaps bad macroeconomic outcomes are best seen as symptoms of deeper underlying problems, including institutional weaknesses and exposure to external shocks.

Although some claims for the importance of policy may have been exaggerated, a commonsense view commands wide support: there is likely a threshold level in the quality of macroeconomic management below which growth becomes difficult or impossible. Easterly (2001a) provides a clear and persuasive exposition of this view, indicating that governments may not be able to initiate growth, but they can destroy growth prospects with bad enough macroeconomic policies. He illustrates the consequences of policy errors using several historical examples, showing that the worst policy outcomes—hyperinflation, high black market premiums, and 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 countries with sound macroeconomic policies and slow growth—Bolivia in the 1990s, for example, discussed in Kaufmann, Mastruzzi, and Zavaleta (2003).

The commonsense view dominates recent assessments of the role of policy but is largely absent from the empirical literature. Traditionally, cross-country research on policy and growth uses simple linear models of the form

$$g = \eta + \alpha P + \beta' Z + \varepsilon \tag{1}$$

where g is the growth rate, P indicates the quality of macroeconomic policy, Z is a vector of other growth determinants, η and α are parameters, β is a parameter vector, and ε is an error term. This linear specification assumes that bad policies can be offset by other factors or, put differently, that the variables can smoothly substitute for one another. Yet many informal accounts of growth are phrased in terms of necessary conditions, which cannot be captured by a linear regression of the form in equation (1). There is surprisingly little research that considers necessary conditions in a formal way, with the exceptions of Hausmann, Rodrik, and Velasco (2008) on binding constraints and Hausmann, Pritchett, and Rodrik (2005) on the factors that instigate growth accelerations. A survey by Montiel and Servén (2006) also draws heavily on the binding-constraints perspective, and some additional discussion can be found in Temple (2009).

A simple way to address the problem is to examine the distribution of growth rates. If macroeconomic instability can form a binding constraint, unstable countries should have growth rates that are tightly distributed around a low mean, because instability is a sufficient condition for slow growth. By contrast, for stable countries growth rates should be more widely dispersed around a higher mean. Wide dispersion would arise because stable countries may lack other growth preconditions, leading to variation in performance across these countries driven by variation in other growth determinants (see figure 1, left panel, for hypothetical distributions of growth rates across countries).

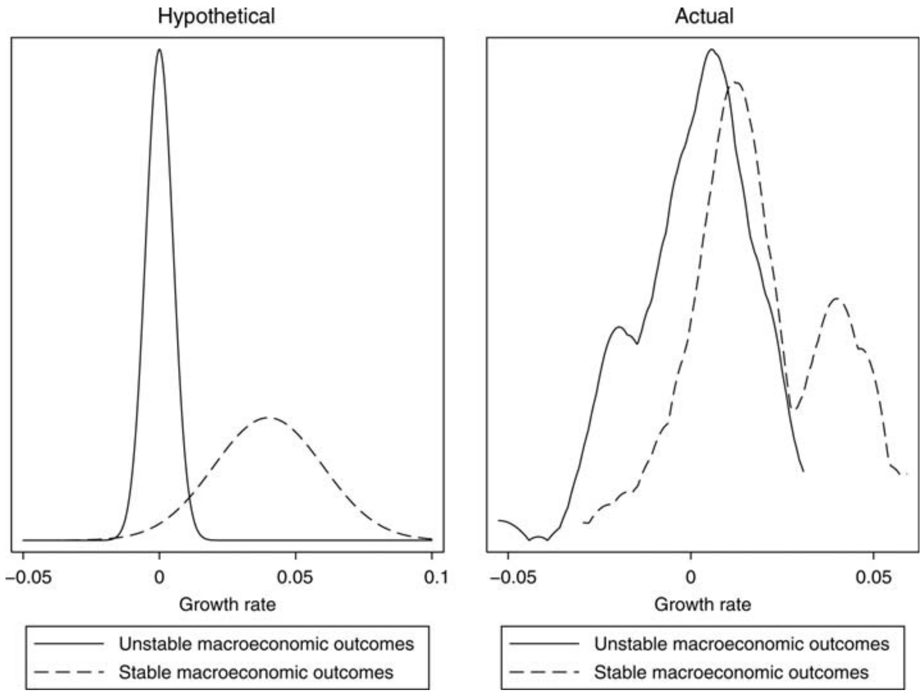
The binding constraints view also has implications for the specification of empirical growth models. One way to capture the idea is a simple nonlinear model with two regimes:

$$\begin{aligned} g &= \eta_1 + \varepsilon_1 & \text{if } P \leq \gamma \\ g &= \eta_2 + \alpha P + \beta' Z + \varepsilon_2 & \text{if } P > \gamma \end{aligned} \tag{2}$$

using similar notation to the previous example. The model implies that if the policy indicator P fails to exceed some threshold value γ, governments effectively destroy any prospect of growth, given low η_1 and a low variance of the error term ε_1 and regardless of other country characteristics. Section III uses Hansen's (1996, 2000) methods to estimate more general versions of equation (2) and shows that macroeconomic stability appears to be a more important threshold variable than other candidates, such as measures of geography and institutions.

The analysis here is based solely on cross-section variation, which has some advantages over a panel analysis for this research question. One drawback of panel data is that short-run deterioration in policy outcomes may be

FIGURE 1. Distributions of Growth Rates



Note: For clarity the distribution for intermediate macroeconomic outcomes is omitted.
Source: Authors' analysis based on data listed in table 1.

associated with a short-term growth slowdown, even if macroeconomic stability and growth are not associated in the long run (Bruno and Easterly 1998). The panel data approach could easily capture these short-run responses rather than genuine long-run effects on potential output. As Pritchett (2000) and Solow (2001) emphasize, models of growth are models of the evolution of potential output, and empirical analyses should be designed with this in mind.

Moreover, cross-section variation may be more informative than panel data about the effects of the ex ante prospects for stability, since a panel data analysis could be driven mainly by the effects of the realizations of outcomes. Given the spans of data currently available, there is a case for using cross-section data to identify the long-run impact of macroeconomic stability. A strong association between stability and growth in the international cross section would shift the burden of proof in the debate, placing new demands on those who argue that macroeconomic stability is largely irrelevant.

TABLE 1. Variables and Definitions

Variable	Description	Sources
ABSLAT	Absolute latitude (distance from the equator)	Hall and Jones (1999)
BD	Burnside-Dollar policy index	Burnside and Dollar (2000)
BMP	Log of (1 + mean black market premium)	Easterly and Sewadeh (2002)
ELR7097	Easterly, Levine, and Roodman update of Burnside-Dollar policy index	Easterly, Levine, and Roodman (2004)
ERATE	Variation of the Dollar real exchange rate measure	Dollar (1992)
EXPRISK	Protection against expropriation risk. Higher values mean lower risk. Mean value 1985–95.	Acemoglu, Johnson, and Robinson (2001)
FR	Log of a measure of natural openness to trade	Frankel and Romer (1999)
GOVKKM	A composite index of overall quality of governance that uses the mean of indexes for voice and accountability, political stability, and government during 1996–2000. Higher values indicate higher quality governance.	Kaufmann, Kraay, and Mastruzzi (2005)
INFLA	Log of (1 + median inflation rate based on GDP deflator)	World Bank (2004)
INVEST	Log of mean investment share in GDP, 1970–99	Heston, Summers, and Aten (2002)
LITERACY	Log of (100 − illiteracy rate of population ages 15 and older in 1970)	World Bank (2004)
MACRO	The first principal component from a classical principal components analysis of BMP, ERATE, INFLA, OVERVALU, and SURPLUS. Higher values indicate better policy outcomes.	See text
MACROOL	A macroeconomic stability index based on a classical principal components analysis that excludes Guyana, Nicaragua, and Sudan.	See text
OVERVALU	Log of mean overvaluation index. Dollar (1992) provides data for 1976–85. Easterly and Sewadeh (2002) update the data to 1999.	Dollar (1992); Easterly and Sewadeh (2002)
POLCON	A measure of the extent of political constraints in policymaking. A higher value implies stronger constraints. The mean value for 1970–99 is used.	Henisz (2000)
POLITY	A measure of the degree of democracy. The POLITY score is the democratic score minus autocratic score on a −10 to 10 scale, where higher values mean higher degree of democracy. The mean value for 1970–99 is used.	Marshall and Jaggers (2002)
POPG	Log of the average annual growth rate of the population ages 15–64 for 1970–99, plus 0.05.	World Bank (2004)
RGDP7099C	Log of real GDP per capita ("rgdpch") in 1999 minus the log of real GDP per capita for 1970. This is divided by 29, to obtain annual growth rates.	Heston, Summers, and Aten (2002)
RGDP7099W	Log of real GDP per worker ("rgdpwok") in 1999 minus that of 1970. This is divided by 29, to obtain annual growth rates.	Heston, Summers, and Aten (2002)

(Continued)

TABLE 1. Continued

Variable	Description	Sources
RGDPPC70	Log of real GDP per capita ("rgdpch") in 1970.	Heston, Summers, and Aten (2002)
RGNEAP	East Asia and Pacific regional dummy variable	Easterly and Sewadeh (2002)
RGNLAC	Latin America and the Caribbean regional dummy variable	Easterly and Sewadeh (2002)
RGNMENA	Middle East and North African regional dummy variable	Easterly and Sewadeh (2002)
RGNSA	South Asian regional dummy variable	Easterly and Sewadeh (2002)
RGNSSA	Sub-Saharan African regional dummy variable	Easterly and Sewadeh (2002)
RMACRO	The first principal component from a robust principal components analysis.	See text
SCHOOL70	Log of average years of schooling at all educational levels of population age 15 and older in 1970.	Barro and Lee (2001)
SURPLUS	Mean central government budget surplus as a share of GDP, 1970–99	World Bank (2004)

Source: Authors' construction.

II. A NEW WAY TO MEASURE MACROECONOMIC STABILITY

This section introduces a new index of macroeconomic stability that combines several indicators, and uses it to measure the average extent of stability over 1970–99. This combination has two advantages. From a statistical point of view, it lessens the outlier problems associated with skewed distributions. And from an economic point of view, it aims to capture an underlying latent variable, the quality of the macroeconomic decision-making process, rather than relying on more specific "symptoms" such as high inflation. Using several proxies for this latent variable reduces measurement error and makes sense if, as suggested by Sala-i-Martin (1991), macroeconomic disarray is associated with undesirable outcomes across a range of indicators. This approach acknowledges the difficulty in identifying the separate effects of fiscal discipline, inflation control, and exchange rate management in small cross-country data sets. Instead, it makes sense to reduce the dimensions of the problem and focus on a single index of policy outcomes. Arguably, there is more hope of answering questions about policy outcomes and growth when the relevant hypotheses are deliberately characterized in broad terms, given the limitations of the available data.

The composite measure is based on fiscal discipline, inflation, and exchange rate management. The preferred index is based on an outlier-robust version of principal components analysis, using Rousseeuw's (1984) minimum covariance determinant method. The empirical analysis discussed later focuses on developing economies with available data, excluding transition economies and countries with small populations (fewer than 250,000 people in 1970). The main indicators were constructed from a sample of 78 countries; data availability means that the growth regressions discussed later use 60–70 observations, while the Bayesian model averaging in section IV uses 72 observations. See table 1 for definitions and sources of variables used in the analysis, and table 4 in section III for a list of countries.

The individual policy indicators are as follows. Fiscal discipline is measured using data on the average central government budget surplus as a share of GDP (*SURPLUS*) over 1970–99.[3] Some countries, notably Guyana and Sudan, have extreme negative values for this variable, reflecting persistently high budget deficits. The principal component analysis, and hence the later results, is robust to excluding these countries or replacing *SURPLUS* with the monotonic but bounded transformation, arctan (*SURPLUS*).[4]

Success in keeping inflation low is captured in the variable *INFLA*. This is the natural log of 1 plus the median inflation rate over 1970–99, computed from the GDP deflator. The median inflation rate is used to capture success in keeping inflation low on average. Relative to the more commonly used of the mean, this measure is less at risk of being dominated by short-lived episodes of hyperinflation.

Exchange rate management is measured in three ways: 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, *BMP* is defined as the natural log of 1 plus the mean value of the black market premium over the period.

Dollar (1992) introduced the variables *OVERVALU* and *ERATE*, whereas Easterly and Sewadeh (2002) extended *OVERVALU* forward and backward. *OVERVALU* is based on evaluating price levels in a common currency, after correcting for the possible effects of factor endowments on the prices of nontradables by using the component of price levels that is orthogonal to GDP per capita and its square, population density, and two regional dummy variables. A price level higher than predicted by these controls indicates that the domestic prices for tradables may be high; thus high values of *OVERVALU* could indicate a combination of real overvaluation and trade restrictions. The precise interpretation of this measure is discussed further in the appendix.

ERATE is a measure of variability in the overvaluation index for 1976–85 (see table A-1 in Dollar 1992) and can be seen as capturing instability in

3. An alternative would be the stock of central government debt relative to GDP, but *SURPLUS* is available for more countries.

4. This transformation is a natural choice, given that the variable is a ratio that 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 1 standard deviation below the mean, compared with 5 standard deviations below in the raw data.

exchange rate management. Given the probable role of inflation in generating movements in the overvaluation index, it may also indicate more general forms of macroeconomic instability (Rodriguez and Rodrik 2000).

Although the analysis sometimes uses the five outcome indicators individually, they are usually aggregated into a composite index. The best known such index in the recent literature is that of Burnside and Dollar (2000), who construct an aggregate measure of policy quality based on three indicators: inflation, the budget surplus, and the Sachs and Warner (1995) indicator of openness to trade.[5] Since Burnside and Dollar's 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 dummy variables, and proxies for political stability. This procedure is less suited to the aims of this study. In their procedure, growth will typically be correlated with the aggregate policy index by construction. But here the aim is to compare distributions of growth rates across countries with good and bad policy outcomes, which requires a composite policy index that does not use information on growth rates.

The five separate variables are aggregated using a principal components analysis. The first step is to check that the correlations between the variables are high enough to justify using principal components: in the extreme case, where the variables were all pairwise uncorrelated, a principal components analysis would not make sense. A likelihood ratio test can be used to examine that "sphericity" case, allowing for sampling variability in the correlations. This test comfortably rejects sphericity at the 1 percent level (for more details, see the supplemental appendix at http://wber.oxfordjournals.org/).

The first principal component is always normalized in such a way that high values indicate macroeconomic stability (table 2). In terms of standardized indicators (all with mean 0 and variance of 1) the first index can be written as

$$MACRO = 0.334 * SURPLUS - 0.447 * INFLA - 0.585 * BMP - 0.347 * OVERVALU - 0.475 * ERATE. \quad (3)$$

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 percent of the total variance in the standardized data. According to this index, the governments that were most successful in achieving macroeconomic stability during 1970–99 were Singapore, Thailand, Malaysia, Panama, and Benin. By contrast, the analysis suggests that

5. Burnside and Dollar (2000) also experiment with government consumption as a share of GDP but find it to be negatively correlated with the budget surplus and insignificant when the budget surplus is included.

TABLE 2. Results of Principal Components Analysis

Variable	Expected sign	MACRO 1st principal component	MACRO 2nd principal component	RMACRO 1st principal component	RMACRO 2nd principal component	MACROOL 1st principal component	MACROOL 2nd principal component
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 (percent)		41.94	20.29	41.27	24.00	37.29	23.10

Note: Values are the correlation between principal components and the corresponding variables. Numbers in bold indicate the highest correlations between a given principal component and corresponding variables. See table 1 for definitions and sources of variables.

Source: Authors' analysis based on data listed in table 1.

Nicaragua, Guyana, Sudan, Uganda, and Zambia were characterized by long-term instability.

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

For this reason, the main focus of this article is an alternative index, based on an outlier-robust principal components analysis. The relatively small dimensions of the problem suggest the use of the minimum covariance determinant method, which identifies 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 from Rousseeuw 1984; see also Rousseeuw and van Driessen 1999). The covariance matrix for these h observations can be used to represent the associations among the variables and to compute the eigenvectors associated with the principal components. The standard choice $h = 0.75n$ will be used, so that the method effectively discards the least representative 25 percent of the cases in estimating the correlations, building in a high degree of robustness.[6]

This approach to estimating correlations can then be used to extract outlier-robust principal components. The correlations between the first two of these new principal components and the individual policy indicators are shown in the *RMACRO* column of table 2. In terms of loadings on the individual variables, the robust index can be written as:

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

where each variable has now been centered using a robust estimate of its location. Relative to the classical principal components analysis, the outlier-robust principal components analysis places less weight on *SURPLUS*, *OVERVALU*, and *ERATE* and more weight on *INFLA* and *BMP*. Although the weights in the two cases may look different, the simple correlation between *MACRO* and *RMACRO* is 0.98, reflecting high correlations between some of

6. The ROBPCA program can be used to implement the minimum covariance determinant approach. 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. Also, using an outlier-robust approach to principal components analysis does not preclude the possibility of extreme (and hence informative) observations in the final index. Rather, the idea is to limit the influence of small numbers of observations on the weighting scheme used in constructing the index.

TABLE 3. Correlations between GDP Growth and Various Policy Indexes

Policy index	RGDP7099C	MACRO	RMACRO	MACROOL	BD	ELR7097
RGDP7099C	1.000					
MACRO	0.471	1.000				
RMACRO	0.420	0.976	1.000			
MACROOL	0.409	0.995	0.991	1.000		
BD	0.673	0.666	0.623	0.585	1.000	
ELR7097	0.590	0.603	0.621	0.645	0.850	1.000

Note: See table 1 for definitions and sources of variables. Sample size varies between 64 and 78 countries, depending on data availability.

Source: Authors' analysis based on data listed in table 1.

the individual components. With the *RMACRO* index, the five best performers are Singapore, Thailand, Panama, Malaysia, and Togo, and the five worst performers are Nicaragua, Uganda, Ghana, Argentina, and the Democratic Republic of Congo.

An alternative approach would be to use the diagnostic plot suggested by Hubert, Rousseeuw, and Branden (2005), which can identify possible outliers that are then excluded from an otherwise standard principal components analysis. This method indicates that Guyana, Nicaragua, and Sudan might be anomalous observations. However, the *MACROOL* column of table 2 shows that this makes little difference. The proportion of the variance explained by the first principal component falls slightly, but the correlations between this component and the different indicators are similar to those reported in the *MACRO* and *RMACRO* columns.

The correlations between *MACRO*, *RMACRO*, the Burnside-Dollar index, and the updated Burnside-Dollar index for 1970–97 from Easterly, Levine, and Roodman (2004) are high (table 3), suggesting that the various indexes may be capturing an underlying latent variable. This is the case even though the Burnside–Dollar and Easterly–Levine–Roodman measures use a different weighting strategy as well as the Sachs-Warner measure of liberal policies, including trade policies. At the same time, the correlations clarify that the results in sections III and IV should not be interpreted too literally. A measure that is notionally of macroeconomic stability may capture other aspects of policy or equilibrium outcomes, especially when instability is a symptom of a dysfunctional policy environment or periods of conflict.

III. IS MACROECONOMIC INSTABILITY THE WEAKEST LINK?

The preferred index, *RMACRO*, is now used to examine how growth varies across countries with good and bad macroeconomic outcomes. Ordering the countries by *RMACRO* and splitting the sample at the 33rd and 66th percentiles yields three groups of countries (table 4). The distributions of growth rates can then be compared across these groups. The growth rate is measured in annual terms, based on GDP per capita (chain weighted) over 1970–99, using data from version 6.1 of the Penn World Table (Heston, Summers, and Aten 2002).

The median growth rate is substantially lower for the relatively unstable group 1 than for groups 2 and 3 (see figure 2, left panel; group 1 is the least stable, group 3 the most stable). There is less support for the idea that macroeconomic instability always destroys long-term growth prospects, because even in group 1, the 75th percentile of the growth rate is 1.4 percent. The patterns are similar (not shown) when growth is measured using GDP per worker rather than GDP per capita and when classifying countries according to *MACRO* rather than *RMACRO*.

Kernel density plots can be used to summarize the same information in a slightly different way.[7] Stable countries have higher growth on average, but instability does not necessarily preclude growth (figure 1, right panel). There is substantial variation in growth across the countries with unstable outcomes, and a significant fraction display positive growth rates over the period. Nevertheless, there are no countries growing at more than 3.5 percent a year in the unstable group, whereas seven countries in the stable group grew at least this rapidly (Cyprus, Indonesia, Republic of Korea, Malaysia, Mauritius, Singapore, and Thailand). Based on this evidence, macroeconomic stability is a necessary condition for sustaining high growth rates over a long period.

An alternative method is to examine the box plots for all five individual indicators, *SURPLUS, INFLA, BMP, OVERVALU*, and *ERATE*. The patterns (not shown) are generally less supportive of the idea that stability promotes growth, suggesting that combining the indicators into an overall index is worthwhile. The evidence that stability matters is strongest when the Dollar (1992) index of exchange rate distortions (*OVERVALU*) and the black market premium (*BMP*) are used to group countries (see figure 2, right panel, for results using the black market premium).

Growth Regressions

This subsection uses growth regressions to examine the relationship between macroeconomic stability and growth. Conventional linear models are used, estimated by ordinary least squares and two-stage least squares, starting with Mankiw, Romer, and Weil's (1992) version of the Solow model. This is arguably the leading structural model in the literature, and it reduces arbitrariness in the choice of specification. The model is estimated using data for 1970–99 rather than for 1960–85 as in Mankiw, Romer, and Weil. Even conditional on the investment rate, population growth, initial income, and regional dummy

7. The samples are relatively small to apply these methods, and the choice of bandwidth becomes important. This is discussed in the supplemental appendix, available at http://wber.oxfordjournals.org/

Table 4. RMACRO Values and Grouping, by Country

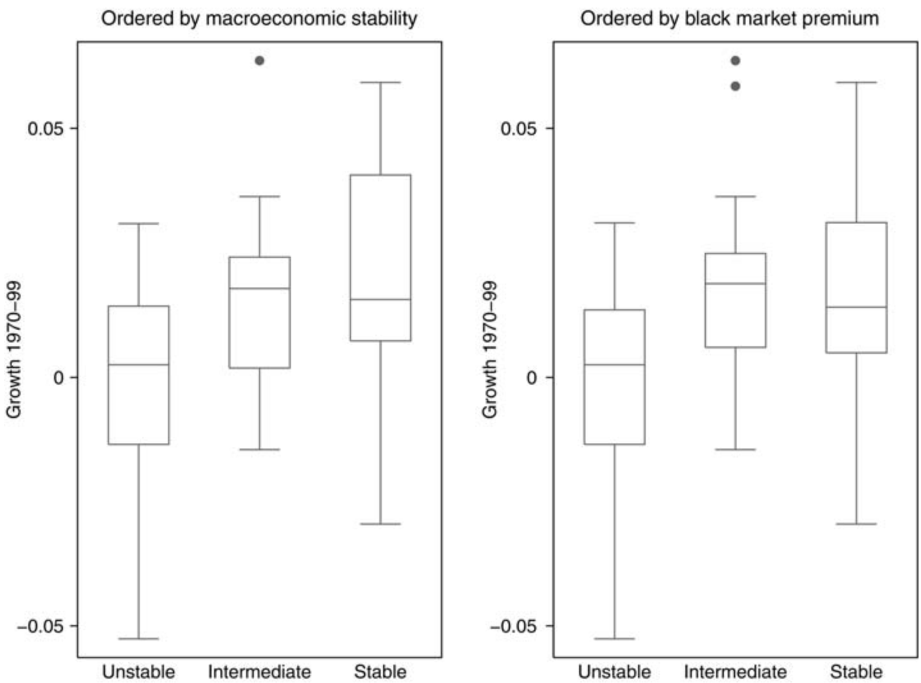
Number	Country	RMACRO	Group	Number	Country	RMACRO	Group
1	Nicaragua	−2.974	1	40	Ethiopia	0.161	2
2	Uganda	−2.009	1	41	Sri Lanka	0.165	2
3	Ghana	−1.680	1	42	Mexico	0.237	2
4	Argentina	−1.669	1	43	Madagascar	0.277	2
5	Congo, Dem. Rep.	−1.610	1	44	Lesotho	0.310	2
6	Guyana	−1.547	1	45	Colombia	0.325	2
7	Iran	−1.504	1	46	Kenya	0.348	2
8	Sudan	−1.476		47	Trinidad and Tobago	0.352	2
9	Sierra Leone	−1.463	1	48	Nepal	0.352	2
10	Somalia	−1.266		49	India	0.364	2
11	Zambia	−1.254	1	50	Botswana	0.371	2
12	Bolivia	−1.185	1	51	Pakistan	0.379	2
13	Brazil	−1.165	1	52	Nigeria	0.560	2
14	Peru	−1.115	1	53	Papua New Guinea	0.566	2
15	El Salvador	−1.061	1	54	Philippines	0.622	3
16	Liberia	−0.911		55	Indonesia	0.669	3
17	Niger	−0.696	1	56	South Korea	0.684	3
18	Algeria	−0.661		57	Tunisia	0.686	3
19	Uruguay	−0.655	1	58	Ecuador	0.763	3
20	Egypt	−0.555	1	59	Mauritius	0.790	3

21	Syria	−0.522	1	60	Congo, Rep.	0.832	3
22	Venezuela	−0.493	1	61	Morocco	0.842	3
23	Jamaica	−0.491	1	62	Mali	0.865	3
24	Yemen	−0.460		63	Chad	0.927	3
25	Zimbabwe	−0.454	1	64	Cameroon	0.954	3
26	Turkey	−0.441	1	65	Gabon	0.983	3
27	Mauritania	−0.399	1	66	Cyprus	0.989	3
28	Costa Rica	−0.371	1	67	Oman	1.124	
29	Paraguay	−0.360	1	68	Central African Rep.	1.126	3
30	Chile	−0.360	2	69	Burkina Faso	1.139	3
31	Malawi	−0.338	2	70	Senegal	1.153	3
32	Haiti	−0.311	2	71	Benin	1.210	3
33	Rwanda	−0.237	2	72	Fiji	1.219	3
34	Israel	−0.137	2	73	Jordan	1.246	3
35	Honduras	−0.123	2	74	Togo	1.371	3
36	Burundi	−0.026	2	75	Malaysia	1.607	3
37	Dominican Rep.	0.003	2	76	Panama	1.652	3
38	Guatemala	0.078	2	77	Thailand	1.742	3
39	Bangladesh	0.102	2	78	Singapore	1.837	3

Note: RMACRO is listed for the 78 countries used in table 2, ordered from worst to best. The 72 countries with a group number are for countries included in figures 1–4. Group indicator refers to the groups underlying the left panel in figure 2. The main 70 country regression sample is based on the same set of countries, minus Gabon and Sierra Leone, for which the literacy indicator was unavailable.

Source: Authors' analysis based on data listed in table 1.

FIGURE 2. Box Plots for Growth Rates



Note: The upper and lower limits of each enclosed box correspond to the 75th and 25th percentiles of the growth rate, while the horizontal line within each box corresponds to the median. "Unstable" refers to group 1 countries in table 4, "intermediate" to group 2 countries, and "stable" to group 3 countries.

Source: Authors' analysis based on data listed in table 1.

variables, a significant partial correlation is found between growth and macroeconomic stability.

The specification relates the log difference in GDP per capita to the log of the investment rate, the log of initial GDP per capita, the log of population growth plus 0.05, and a human capital variable, as in Mankiw, Romer, and Weil (1992). There are two main departures in the current specification. First, regional dummy variables are used to proxy for the initial level of efficiency, as in Temple (1998). Second, the regressions use a measure of the initial level of educational attainment rather than the rate of investment in human capital.[8] This will be the natural log of either the 1970 literacy rate (from World Bank 2004) or average years of schooling in 1970 (from Barro and Lee 2001). In both cases, the data refer to the population ages 15 and older.

Regression 1 excludes the policy indicators (table 5). The Mankiw, Romer, and Weil (1992) regression continues to work well over a different time period;

8. The use of a stock measure rather than a flow can be justified formally as a proxy for the steady-state level of educational attainment, as in equation (12) in Mankiw, Romer, and Weil (1992).

TABLE 5. Macroeconomic Stability and Growth Regressions

Variable	1 OLS	2 OLS	3 OLS	4 OLS	5 OLS	6 2SLS	7 OLS	8 OLS
Regime	All	All	All	All	All	All	1	2
Number of observations	70	70	70	70	60	70	42	28
RMACRO		0.71 (0.30)	0.49 (0.31)	0.64 (0.27)	0.64 (0.29)	1.35 (0.66)	0.70 (0.42)	−1.20 (0.39)
Initial income	−1.10 (0.37)	−0.26 (0.37)	−0.80 (0.37)	−1.04 (0.38)	−1.15 (0.42)	−0.98 (0.33)	−0.83 (0.43)	−1.26 (0.30)
Investment	1.07 (0.32)		1.10 (0.34)	0.83 (0.32)	0.84 (0.48)	0.56 (0.42)	0.45 (0.42)	1.52 (0.41)
Population growth	−0.21 (0.23)		−0.19 (0.22)	−0.12 (0.25)	−0.10 (0.28)	−0.02 (0.23)	−0.15 (0.29)	0.08 (0.15)
LITERACY	0.68 (0.31)			0.88 (0.34)		1.12 (0.32)	0.72 (0.36)	0.41 (0.35)
SCHOOL70					0.79 (0.27)			
GOVKKM							1.06 (0.98)	1.91 (0.44)
Investment elasticity	1.18		1.69	0.96	0.89	0.69	0.66	1.47
R^2	0.51	0.37	0.51	0.57	0.55	n/a	0.47	0.90
Regression standard error	1.56	1.75	1.57	1.47	1.58	1.47	1.38	0.84
Heteroscedasticity								
Breusch-Pagan	0.32	0.02	0.07	0.27	0.18	0.48	0.01	0.09
White	0.66	0.19	0.03	0.64	0.35	0.66	0.07	0.46
Ramsey RESET	0.90	0.58	0.02	0.68	0.24	0.97	0.01	0.84
Anderson-Rubin						0.02		

Note: OLS is ordinary least squares. 2SLS is two-stage least squares. The dependent variable is the annual growth rate over 1970–99 in percentage points. Numbers in parentheses are MacKinnon-White heteroskedasticity-consistent (hc3) standard errors, except for regression 6, for which numbers in parentheses are White heteroskedasticity-consistent standard errors. Constants are included but not reported. Regressions 1–6 include five regional dummy variables, for which the coefficients are not reported. The explanatory variables are standardized to have a standard deviation of 1 in the 70 country sample. Investment elasticity is the elasticity of the steady-state income level to the investment rate. Heteroscedasticity reports p-values associated with two tests for heteroscedasticity. Ramsey RESET (regression equation specification error test) is the p-value associated with this test. Anderson-Rubin is the p-value associated with the Anderson-Rubin test for the significance of the endogenous explanatory variable (RMACRO). See table 1 for definitions and sources of variables.

Source: Authors' analysis based on data listed in table 1.

the explanatory power is similar, although the effect of population growth is imprecisely estimated. The elasticity of steady-state income to the investment rate is 1.18, within the range spanned by Mankiw, Romer, and Weil's estimates. Regression 2 includes only initial income, regional dummy variables, and the new measure of stability, *RMACRO*. The stability measure is significant at the 5 percent level, and the association is strong: if interpreted as a causal effect, a 1 standard deviation improvement in stability would have raised the annual growth rate by 0.71 percentage point over the time period. Regression 3 controls for the effects of investment and population growth, as in Mankiw, Romer and Weil. The effect of *RMACRO* is slightly weaker, as might be expected, but significant at the 12 percent level. The reduction in the size of the coefficient indicates that macroeconomic stability may boost investment, an idea that will be explored later.

Regression 4 includes *LITERACY*, the log of the 1970 literacy rate, which increases the explanatory power of the model. *RMACRO* is once again significant at the 5 percent level. This result is robust to replacing the literacy rate with the log of average years of schooling in 1970, *SCHOOL70*, as in regression 5. This reduces the size of the sample by 10 observations, so regression 4 is the preferred specification in the discussion that follows.

The partial correlations between growth and macroeconomic stability do not appear to be driven by anomalous observations. The results are robust to the deletion of potential outliers, as identified by least absolute deviation regressions.[9] The findings are similarly robust to using single-case diagnostics such as DFFITS and DFBETA, which identify a similar set of outliers to the least absolute deviation method in this case.[10] Added-variable plots (not shown) were also used to identify potential outliers. When the Democratic Republic of Congo and Nicaragua are excluded, the results are slightly less strong, with *RMACRO* significant only at the 8 percent level. Finally, some simple diagnostic tests are supportive: the models do not suffer from omitted nonlinearities (based on the Ramsey RESET test) or heteroskedasticity (based on versions of the Breusch–Pagan and White tests) except in regression 3, which includes investment but not a measure of human capital.

Given the concern that macroeconomic stability is not randomly assigned, an instrumental variable approach might be preferable. One possible route exploits the observed association between French colonial heritage and macroeconomic stability, as in Barro (1996). Many former French colonies maintained a fixed exchange rate with the French franc, and this appears to have been associated with lower inflation rates. The sample contains 15 former French colonies, and for these countries the mean of *RMACRO* is 0.52 and the

9. Outliers were defined by least absolute deviation residuals more than two standard deviations from the mean value.

10. The results are available on request. See Cook and Uchida (2003) for a brief discussion of how DFFITS and DFBETA are computed and used.

standard deviation is 0.72. This compares favorably to a mean of 0.01 and standard deviation of 1.03 for former British colonies and, since *RMACRO* is standardized, to a mean of 0 and a standard deviation of 1 for the sample as a whole.

Regression 6 instruments *RMACRO* using a dummy variable for former French colonies. The significance of *RMACRO* is tested using the Anderson and Rubin (1949) statistic, which is optimal for models that are just-identified (Moreira 2003) and should be robust to weak-instrument problems. The *p*-value associated with the test is 0.02, so *RMACRO* is significant even in the two-stage least squares estimates. The two-stage least squares estimate assigns more weight to macroeconomic stability and less to investment than the ordinary least squares point estimate does. The finding that the two-stage least squares coefficient for *RMACRO* is considerably higher than the ordinary least squares coefficient could be due to measurement error or sampling variability, as Acemoglu, Johnson, and Robinson (2001) and Frankel and Romer (1999) have argued in other contexts. But it could also be due to a failure of the exclusion restriction, so these results should be treated cautiously. The small number of observations reinforces this point. To lessen endogeneity problems arising from the nonrandom assignment of policy, the approach used in the next subsection, namely a comprehensive search through a wide range of control variables and specifications, may be preferable.[11]

In summary, there is an association between macroeconomic stability and growth, even conditional on investment rates. Taking the results at face value, a 1 standard deviation improvement in stability translates into an annual growth rate that is 0.5–0.7 percentage point higher over 30 years. Increasing the annual growth rate by 0.7 percentage point would leave GDP per capita 23 percent higher after 30 years. A later analysis will consider the implications for the location and shape of the distribution of growth rates and the steady-state distribution of GDP per capita.

Threshold Estimation

This subsection uses Hansen's (1996, 2000) methods for sample splitting and threshold estimation to estimate nonlinear models of the following type:

$$\begin{aligned} g &= \eta_1 + \alpha_1 P + \beta_1' Z + \varepsilon_1 \quad \text{if } P \leq \hat{\gamma} \\ g &= \eta_2 + \alpha_2 P + \beta_2' Z + \varepsilon_2 \quad \text{if } P > \hat{\gamma}, \end{aligned} \tag{5}$$

11. Moreover, the applicability of instrumental variable approaches to cross-country growth data may have been exaggerated. When the instrument is correlated with the error term, even weakly, the inconsistency of the instrumental variable estimator can be worse than that of the ordinary least squares estimator, particularly if the instrument is not strongly correlated with the endogenous explanatory variable (see Cameron and Trivedi 2005). There are good reasons to doubt many of the exclusion restrictions adopted in the literature, since most candidates for instruments might be correlated with omitted growth determinants; see Durlauf, Johnson, and Temple (2005, 2009) for more discussion.

where $\hat{\gamma}$ is a threshold estimated jointly with the other parameters in the model and P could be an indicator of macroeconomic outcomes or some other variable, such as a measure of institutional quality. This specification nests the earlier example, equation (2), since the intercept and slope coefficients are allowed to vary across the two regimes. A particular strength of the Hansen approach is that alternative candidates for the threshold variable P can be compared on statistical grounds. Moreover, by comparing the models for different regimes, it is possible to see whether macroeconomic instability forms a binding constraint on growth. If so, instability should limit the benefits of favorable fundamentals, such as geographic and institutional characteristics.

It is possible to test for the existence of a threshold, and hence multiple regimes, using the Hansen (1996) bootstrapped Lagrange multiplier test. Hansen (2000) develops an asymptotic approximation to the least squares estimate of the threshold $\hat{\gamma}$, which allows construction of a (possibly asymmetric) confidence interval. These methods can therefore reveal the extent to which a proposed sample split is estimated with precision and whether the proposed nonlinearity is supported by the data.[12] As in the earlier analysis, the main limitation arises from the possible correlation between macroeconomic stability and the error term, which brings the risk that the assignment of countries across regimes could also be a function of the error term, and so the results should be cautiously interpreted.

Seven possible candidates for the threshold variable P are considered—*RMACRO* and six indicators of either geographic or institutional fundamentals—to determine whether differences in macroeconomic stability give rise to distinct growth regimes or whether fundamentals provide a better way to divide the sample. Two of the fundamentals considered are standard measures of geographic characteristics. The first variable, *FR*, is the log of the Frankel and Romer (1999) measure of natural openness to trade, which is based partly on proximity to large markets. The second variable, *ABSLAT*, is absolute latitude—that is, distance from the equator. In both cases, the data are taken from Hall and Jones (1999).

The remaining four candidates for threshold variables are all measures of institutional quality. These are *GOVKKM*, a composite index of the quality of governance for 1996–2000, from Kaufmann, Kraay, and Mastruzzi (2005); *POLITY*, the extent of democracy, based on the Polity IV database of Marshall and Jaggers (2002) and averaged over 1970–99; *POLCON*, a measure of the extent of political constraints from Henisz (2000) averaged over 1970–99; and *EXPRISK*, the measure of average expropriation risk for 1985–

12. Previous applications of these methods to growth regressions include Hansen (2000) and Papageorgiou (2002). In emphasizing institutions as a potential threshold variable, this article is especially close to the work of Minier (2007) but considers the role of macroeconomic stability in more detail.

95 used in Acemoglu, Johnson, and Robinson (2001). Several of these measures are based partly on observed outcomes rather than constraints. This may lead the benefits of good institutions to be overstated and the benefits of macroeconomic stability to be understated.[13]

For each of the six fundamental variables, a regression is used to relate growth to that variable, the Solow variables, and *RMACRO*. Regional dummy variables are omitted to avoid overfitting problems when the sample is subdivided. Hansen's approach is used to test for a threshold associated with *RMACRO* and alternatively with the fundamental variable.

It is immediately apparent that *RMACRO* dominates all the other candidates as a threshold variable (table 6). In all but one case the null of no threshold is rejected for *RMACRO* at the 10 percent level, while it is not rejected for any of the other six measures of fundamentals. These results suggest that the data are well described by two regimes, where the classification of countries into the two regimes depends on macroeconomic stability rather than geography or institutions. The estimated threshold for *RMACRO* is also reported in table 6, along with its 95 percent confidence interval (which may be asymmetric) and the number of countries in each subsample. Since *RMACRO* is normalized to have a mean of 0 and a standard deviation of 1 in the 70 country sample, it is clear that the threshold is precisely estimated and relatively stable across the various specifications.

An especially interesting result is that when the sample is divided using the estimated threshold, the standard growth variables have much higher explanatory power for the relatively stable countries. For this group, the model typically accounts for 75–90 percent of the variation in growth rates, while the R^2 for the less stable countries is typically 40–50 percent. This is consistent with the binding-constraints view: if macroeconomic stability is achieved, growth is well explained by a standard regression, but the Solow variables (and measures of geographic or institutional fundamentals) have less explanatory power when instability forms a binding constraint on growth, since this limits the benefits of favorable characteristics. The main departure from the earlier hypothesis is that the cross-section residual variance is higher, not lower, for countries that experience macroeconomic instability.[14]

Regressions 7 and 8 show the results for the two groups and are based on a model containing the Kaufmann, Kraay, and Mastruzzi (2005) measure (*GOVKKM*), so the candidate variables for a threshold were *GOVKKM* and *RMACRO*. As in the other cases, the Hansen (1996) test favored macroeconomic stability for splitting the sample. The estimated threshold for *RMACRO*, $\hat{\gamma} = 0.297$, is slightly above the mean and divides the sample into

13. See Glaeser and others (2004) on the general desirability of using measures of constraints or rules rather than measures closely related to outcomes.

14. This is consistent with a competing explanation for the results, namely that measurement errors in the data are more serious for unstable countries.

TABLE 6. Threshold Estimation

Z variable	FR	ABSLAT	GOVKKM	POLITY	POLCON	EXPRISK
RMACRO threshold	0.068	0.167	0.030	0.002	0.013	0.012
Z threshold	0.324	0.320	0.271	0.523	0.600	0.354
γ − RMACRO	0.309	0.180	0.297	−0.185	0.297	0.309
95 percent confidence interval						
Lower	−0.375	−0.300	−0.520	−0.375	−0.520	−1.241
Higher	0.309	0.309	0.714	0.309	0.618	0.324
N [RMACRO $\leq \gamma$]	43	36	42	29	42	36
N [RMACRO $> \gamma$]	27	34	28	40	28	19
R^2 [RMACRO $\leq \gamma$]	0.43	0.60	0.47	0.40	0.42	0.48
R^2 [RMACRO $> \gamma$]	0.80	0.76	0.90	0.77	0.78	0.82
Z p-value [RMACRO $\leq \gamma$]	0.90	0.00	0.29	0.57	0.85	0.49
Z p-value [RMACRO $> \gamma$]	0.29	0.09	0.00	0.07	0.51	0.04

Note: RMACRO threshold is the *p*-value for the Hansen (1996) test of a threshold in RMACRO, in a model that includes RMACRO, the Solow variables, and the *Z* variable. *Z* threshold is the *p*-value for the Hansen (1996) test of a threshold associated with the *Z* variable. The tests indicate that RMACRO can be used to divide the sample into two regimes. The lower rows show the threshold γ for RMACRO estimated using the Hansen (2000) procedure; the 95 percent confidence interval for the threshold (which need not be symmetric); the number of observations in the two regimes on either side of the threshold; the R^2 of the separate growth regressions for the two regimes; and the *p*-value of the *Z* variable for each of the two regimes. The growth regression always has the highest explanatory power in the subsample with greater macroeconomic stability; for an example based on GOVKKM, see regime 1 in regression 7 and regime 2 in regression 8 of table 5. See table 1 for definitions and sources of variables.

Source: Authors' analysis based on data listed in table 1.

42 unstable countries (regression 7) and 28 relatively stable countries (regression 8). Comparing these two sets of results shows that macroeconomic stability is clearly associated with a higher elasticity of steady-state output to the investment rate, faster conditional convergence, and perhaps stronger growth benefits of good institutions. Overall, the explanatory power of the growth regression is much higher and the specification tests more favorable for the stable group. By contrast, a Ramsey RESET test for the less stable group rejects the Solow specification.

Across the six specifications summarized in table 6, a less plausible result is that in the subsamples with relatively stable macroeconomic outcomes *RMACRO* often has a negative sign and is sometimes significant at conventional levels (see the results for regression 8 in table 5). The result that stability has a significantly negative effect in this particular group should be interpreted with caution. It does not arise when the control variable is *ABSLAT*, *POLITY*, or *EXPRISK*. Any significantly negative relationship that emerges may be related to a conditional convergence effect. By construction, all countries in the second regime must have achieved a certain degree of stability, but some may combine instability (relative to other members of the stable group) with strong potential for rapid growth. Simply including initial income as an explanatory variable may not be enough to eliminate such effects. This interpretation of the evidence would be consistent with the idea that once a certain degree of stability has been achieved, the benefits of greater stability may be limited.[15]

Finally, the role of fundamentals (geography and institutions) is considered in more detail. The last two rows of table 6 report the *p*-values associated with these variables for the unstable and stable groups of countries. They show that the posited fundamentals usually lack explanatory power in the less stable countries but often emerge as significant for the more stable countries. Again, this supports an account in terms of binding constraints.

IV. Robustness

This section uses Bayesian methods to examine the robustness of the partial correlation between growth and macroeconomic stability. Levine and Renelt (1992) showed that partial correlations in the empirical growth literature may not be robust to changes in specification. This is a serious problem for growth researchers because the list of candidate predictors is long and it is not easy to rule out particular variables or specifications using prior reasoning. Put differently, there is a model uncertainty problem, and the standard errors in any specific regression will tend to understate the true extent of uncertainty about the parameters. To examine the robustness of the partial correlation, this section uses Bayesian model averaging, as in Brock, Durlauf, and West (2003), Durlauf, Kourtellos, and Tan (2008), Fernandez, Ley, and Steel (2001), Malik and Temple (2009), Raftery, Madigan, and Hoeting (1997), and Sala-i-Martin, Doppelhofer, and Miller (2004).[16]

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

15. The difference in signs for *RMACRO* between the two regimes does not drive the evidence for the existence of a threshold. If the exercise is repeated and *RMACRO* is removed from the models, the *p*-value for a threshold based on *RMACRO* is generally similar to the results in table 6 except for *GOVKKM*, and even there the null of no threshold is still rejected at the 12 percent level.

16. More recently, Crespo Cuaresma and Doppelhofer (2007) and Eicher, Papageorgiou, and Roehn (2007) have developed approaches that allow joint consideration of model uncertainty and sample splits or thresholds. The application of these to macroeconomic stability would be an interesting area for further work, although in samples of the present size, it would be important to allow for outliers.

Consider the case of K possible models, assuming throughout that one of these models generated the observed data D. The models will be denoted by $M_1 \ldots M_K$ 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, $p(\theta_k|M_k)$, to the parameters of each model. Using this structure a Bayesian approach 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 mixing weights are the posterior probabilities that a given model generated the data (see, for example, 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

$$p(\Delta \mid D) = p(\Delta \mid D, M_1)p(M_1 \mid D) + p(\Delta \mid D, M_2)p(M_2 \mid D), \tag{6}$$

where terms in $p(\Delta \mid D, M_k)$ are the conventional posterior distributions obtained under a given model and terms in $p(M_k \mid D)$ are the posterior model probabilities—the probability, given a prior and conditional on having observed D, that model M_k generated the data. This approach requires the evaluation of posterior model probabilities. Briefly, as in Raftery (1995), Raftery, Madigan, and Hoeting (1997), and Sala-i-Martin, Doppelhofer, and Miller (2004), the Bayesian information criterion can be used to approximate the Bayes factors that are needed to compute the posterior model probabilities. This allows a systematic form of model selection and inference to be conducted in a way that acknowledges model uncertainty. For example, to investigate the hypothesis that a slope coefficient β_z is nonzero, the posterior model probabilities are summed for all models in which $\beta_z \neq 0$; this quantity is called a posterior inclusion probability.

As the list of candidate predictors grows, there quickly comes a point where estimation of all the possible models is not feasible, and attention must be restricted to a subset. The approach then follows Raftery, Madigan, and Hoeting (1997) in using a branch-and-bounds search algorithm to identify a subset of models with high posterior probability. For discussion and references, see Malik and Temple (2009) and Sirimaneetham and Temple (2006).

The analysis also draws on the more complex approach of Hoeting, Raftery, and Madigan (1996) because outliers could be a serious problem. In general, any procedure for dealing with model uncertainty or model selection may be influenced by outliers. Even if steps are taken to identify these observations, the final results can easily depend on the order in which model selection and outlier detection is carried out. Hoeting, Raftery, and Madigan suggest a procedure for addressing this issue. First, the full model, containing all the candidate predictors, is estimated by an outlier-robust method due to Rousseeuw (1984), and the standardized residuals are used to identify possible outliers.

Next, model averaging is carried out. As in Hoeting, Raftery, and Madigan, a model is now defined in terms of a set of predictors and a set of observations identified as outliers, where the set of observations identified as outliers include some or all of those identified in the initial stage. (This restriction on the number of candidate outliers is needed to keep the dimensionality of the problem manageable.) Then a Markov chain Monte Carlo model composition approach is used to approximate the posterior model probabilities.

Here, the question of interest is whether *RMACRO* is a robust determinant of growth. The list of candidate predictors will be taken from Sala-i-Martin, Doppelhofer, and Miller (2004), who seek to explain differences in growth rates over 1960–96 for 88 countries (developing and developed). This article instead measures growth over 1970–99 and replaces the Sala-i-Martin, Doppelhofer, and Miller measure of initial GDP for 1960 with a measure for 1970. Despite this change in time period, the same candidate predictors can be used, since the majority of the Sala-i-Martin, Doppelhofer, and Miller explanatory variables were chosen precisely because they are fixed over time or likely to change only slowly. In practice, to keep the application of Bayesian model averaging methods manageable, the analysis that follows focuses on the 31 variables in Sala-i-Martin, Doppelhofer, and Miller (2004, table 2) that have a posterior inclusion probability greater than 4 percent. One of these variables is Dollar's (1992) original index of real exchange rate distortions, measured for 1976–85. This has a low posterior inclusion probability, just 8.2 percent, in the main results of Sala-i-Martin, Doppelhofer, and Miller.

With this set of control variables, the effects of stability can be analyzed at the same time as a wide range of other hypotheses. For example, the Sala-i-Martin, Doppelhofer, and Miller (2004) variables include several measures related to geographic characteristics, including the share of land area in the tropics, the share of population in the tropics, population density, population density in coastal areas in the 1960s, and the prevalence of malaria in the 1960s. Other variables that are included in the Sala-i-Martin, Doppelhofer, and Miller data include regional dummy variables, the relative price of investment goods, life expectancy in 1960, indicators of religion, ethnic diversity, the relative importance of primary exports, and the share of public investment in GDP. The Sala-i-Martin, Doppelhofer, and Miller data span a wide range of the hypotheses investigated in the growth literature, and hence the robustness tests that follow are unusually systematic.[17]

For the purpose of Bayesian model averaging and given the high number of candidate predictors, there are benefits to including as many developing economies in the sample as possible. The measure *RMACRO* is available for

17. One change relative to Sala-i-Martin, Doppelhofer, and Miller is that some explanatory variables are transformed to reduce outlier problems: 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 the analysis, the natural log of these variables is used in place of actual levels.

78 countries but, when merged with the Sala-i-Martin, Doppelhofer, and Miller (2004) data set, the sample is reduced to 63. Country coverage is extended by imputing missing values for a small number of variables in the Sala-i-Martin, Doppelhofer, and Miller data, increasing the number of countries to 72. The decision to impute missing values involves a tradeoff: it introduces measurement error, but it also brings to bear some additional information and lessens the biases that arise when data are missing in nonrandom ways. Here, the number of imputed values in the data matrix for the explanatory variables is just 21, or less than 1 percent of the total number of cells.

The evidence that policy has explanatory power is always much stronger in the 72 country sample than in the 63 country sample, as documented in Sirimaneetham and Temple (2006). The reason for this is clear based on the values of *RMACRO* for the 9 countries that are added to make the 72 country sample. These nine countries include four that are in the bottom decile for *RMACRO* (Guyana, Iran, Nicaragua, and Sierra Leone) and three that are in the top two deciles (Chad, Cyprus, and Fiji). Hence, moving to the larger sample increases the representation of countries at the extreme ends of the distribution of macroeconomic outcomes. This clearly adds identifying variation to the data set. At the same time, considerable faith is needed that policy outcomes and growth are reliably measured for these countries.[18]

The full Bayesian model averaging results are not reported, since the main focus is the posterior inclusion probability associated with *RMACRO*. This is the sum of the posterior model probabilities for all models in which the variable appears. When *RMACRO* is combined with the 31 variables from Sala-i-Martin, Doppelhofer, and Miller (2004), model averaging leads to a posterior inclusion probability of 100 percent, which implies that *RMACRO* appears in every model that is assigned nonzero posterior probability (the Raftery, Madigan, and Hoeting 1997 procedure effectively rounds posterior probabilities down to 0 for the weakest models). The relevant posterior mean—that is, the weighted average of the coefficients on *RMACRO* across all models, where the weights are the posterior model probabilities—is 0.51. This is close to the estimate found in the earlier growth regressions.

When the outlier-robust MC^3 approach of Hoeting, Raftery, and Madigan (1996) is used, the results are weaker, but still supportive. Dollar's (1992) original index of real exchange rate distortions has a high posterior inclusion probability of 99 percent. The evidence for a separate effect of *RMACRO* is weak

18. This is related to a more general debate about the appropriate response to "good" and "bad" leverage points, those observations with extreme values for the explanatory variables; see Dehon, Gassner, and Verardi (2009) and Temple (2000), for example. Here using the 72 country sample comes with the caution that it contains a number of leverage points, affecting inference and the posterior model probabilities.

but becomes much stronger when Dollar's index is excluded. The inclusion probability for *RMACRO* then rises to 69 percent.

Does macroeconomic stability matter, even conditional on institutions? This can be investigated by adding measures of institutional quality to the Bayesian model averaging exercises.[19] These are the same measures used earlier, namely *GOVKKM, POLITY, POLCON*, and *EXPRISK*. Initially, *EXPRISK* is excluded because it reduces the sample of countries substantially. When the other three measures are added to the previous Bayesian model averaging, the posterior inclusion probability of *RMACRO* is 97.4 percent. With the outlier-robust Markov chain Monte Carlo model composition approach, the inclusion probability of *RMACRO* falls to 53 percent. Incidentally, the results strongly support the hypothesis that growth and institutions are highly correlated. The measure *GOVKKM* dominates the others, with an inclusion probability of 100 percent. The inclusion probabilities for the extent of democracy (*POLITY*) and political constraints (*POLCON*) never exceed 35 percent.

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

To summarize, when considering a wide range of candidate growth predictors, the evidence that *RMACRO* matters is sensitive to the inclusion of leverage points. This explains why the results are much stronger for the larger sample based on imputed data. In that sample, there is always a high inclusion probability for either *RMACRO* or Dollar's (1992) index of real exchange rate distortions. Expressed differently, nearly all the best-performing models include at least one of these variables, regardless of how the rest of the specification varies. There is also some evidence that stability matters, even when controlling for institutional quality. This is a demanding test, given that some of these institutional measures are likely to reflect a wide range of outcomes, rather than simply rules and constraints.

V. Counterfactual Distributions

This section examines the role of macroeconomic stability in broader perspective by constructing counterfactual distributions for growth rates and

19. To keep the number of candidate predictors within feasible limits, some of the original Sala-i-Martin, Doppelhofer, and Miller (2004) variables have to be dropped. Those excluded will be the variables with relatively low posterior inclusion probabilities in the main results of Sala-i-Martin, Doppelhofer, and Miller.

steady-state levels of income. These distributions indicate what might have happened had all countries achieved the same level of macroeconomic stability over 1970–99. Regression estimates are used to construct the relevant counterfactuals[20] and to reveal where in the distribution the role of stability may have been especially important, information that is not directly apparent from regression estimates.

In counterfactual distributions, the effects—in terms of changes in the location and shape of the distribution—are rarely uniform throughout the distribution. For example, the changes in growth rates reflected in the shape of the counterfactual distribution depend on the full joint distribution of macroeconomic stability and the growth rate. If all countries with intermediate or better growth rates also had stable outcomes, but countries with low growth did not, then imposing macroeconomic stability throughout the sample would affect only the lower end of the distribution. Changes in the growth rate distribution cannot be summarized simply by a set of regression coefficients, and looking at the whole distribution can add useful information.

The first exercise considers actual and counterfactual distributions of growth rates. The basic goal is to determine what each country's growth rate would have been had all countries achieved the same level of macroeconomic stability over 1970–99. This starts from a growth regression similar to the regressions in section III relating growth to the Solow variables, regional dummy variables, and *RMACRO*. The coefficient on *RMACRO* in this regression is 0.64. The counterfactual growth rate g_i^* is then equal to

$$g_i^* = g_i + 0.64(M^* - RMACRO_i), \tag{7}$$

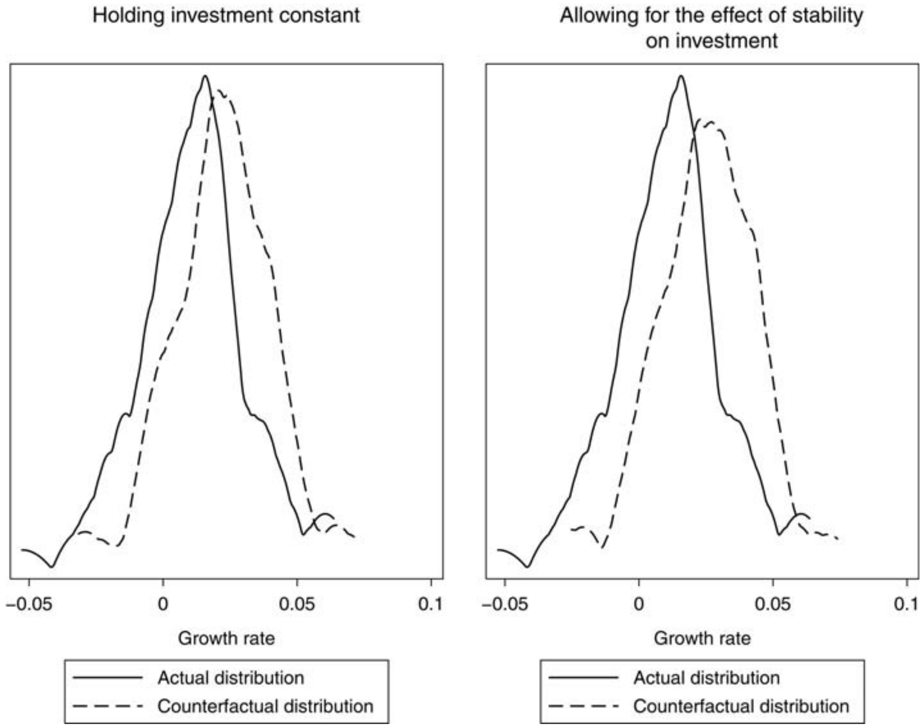
where g_i is the observed growth rate and M^* is the value of *RMACRO* at the 95th percentile in the sample, corresponding to the value for Malaysia.

The distribution of growth rates would have shifted to the right had macroeconomic stability been more widely achieved (figure 3, left panel). But this exercise holds the rate of investment constant and may therefore understate the benefits of macroeconomic stability. This can easily be examined by excluding investment from the growth regression used to construct the counterfactual distribution. The relevant counterfactual distribution now lies slightly further to the right (figure 3, right panel). The benefits of stability continue to be observed throughout the distribution.

The growth regressions include initial income and thus can be seen as modeling the level of the steady-state growth path, as in Mankiw, Romer and Weil (1992). Under the assumption that all countries grow at the same rate in the long-run steady state, the estimated coefficients for 1970–99 can be used to

20. Kernel density estimates of counterfactual distributions are associated in particular with the work of DiNardo, Fortin, and Lemieux (1996) on wage distributions. These methods have also been applied in growth economics by Beaudry and Collard (2006), Beaudry, Collard, and Green (2005), and Desdoigts (1996, 2004).

FIGURE 3. Actual and Counterfactual Distributions of Growth Rates



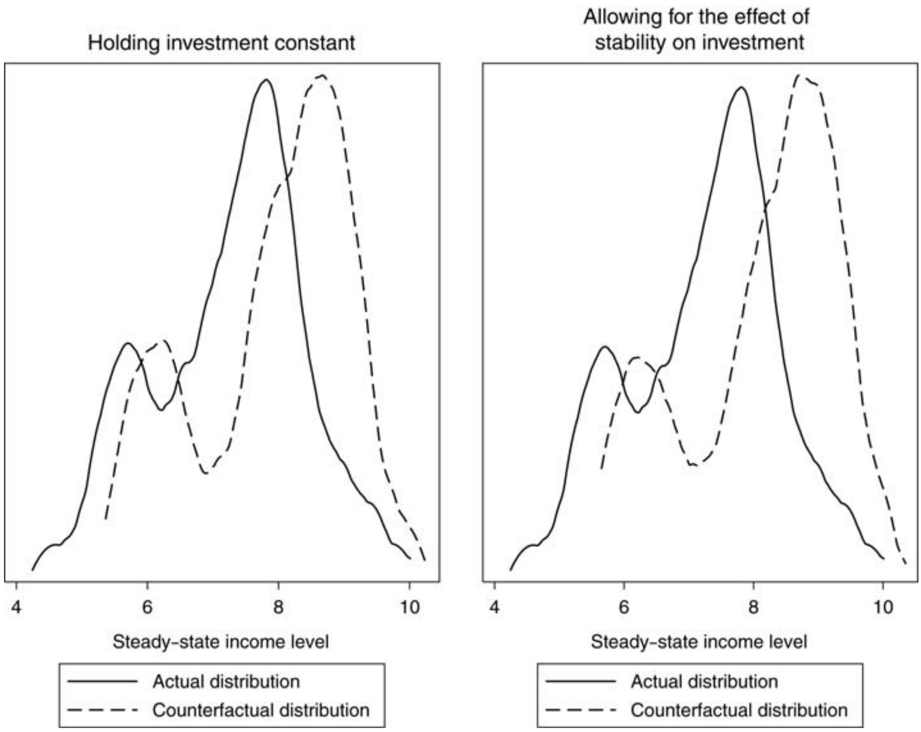
Source: Authors' analysis based on data listed in table 1.

compute the implied steady-state distribution of GDP per capita. Similarly, it is possible to construct a counterfactual distribution that would have been obtained under universal macroeconomic stability.

The actual distributions of the log of GDP per capita are not necessarily expected to have the familiar "twin peaks" pattern of Quah (1996) because the sample is restricted to developing economies. Better macroeconomic outcomes might have moved the distribution of steady-state income levels to the right, and the potential magnitude of this effect is clearly substantial (figure 4, left panel). The analysis is extended by taking into account the effect of *RMACRO* on investment.[21] The counterfactual distribution is slightly further to the right than in the left panel of figure 4, as would be expected if macroeconomic stability were associated with higher investment. Overall, the results indicate that

21. This is based on a simple regression of the log of the investment rate on initial income, initial human capital, regional dummy variables, and *RMACRO*, which is then used to calculate a set of (country-specific) counterfactual investment rates that would have obtained had all countries achieved macroeconomic stability. This is then used in the construction of the counterfactual steady-state distribution (figure 4, right panel).

FIGURE 4. Actual and Counterfactual Distributions of Steady-State Log Income



Source: Authors' analysis based on data listed in table 1.

macroeconomic stability could be a major influence on the steady-state distribution of income levels.

VI. CONCLUSION

This article examined the relationship between macroeconomic stability and growth in developing economies. It introduced a new index of the extent of macroeconomic stability, having aggregated five policy indicators using an outlier-robust version of principal components analysis. With this index, growth is found to be positively associated with macroeconomic stability in a sample of 70 developing economies. If this is interpreted as a causal effect, a 1 standard deviation improvement in the index would raise annual growth by roughly 0.5–0.7 percentage point over 30 years.

Consistent with previous work on this topic, the strength of the evidence depends on the sample of countries. In the largest sample considered, Bayesian methods indicate that the effect is generally robust across a range of specifications. But as the discussion has emphasized throughout, the results are best interpreted as an upper bound on the benefits of good macroeconomic management. Unstable policy outcomes may sometimes reflect deeper institutional weaknesses, exposure to external shocks, or political instability and conflict.

One of the main contributions of this article is to close the gap between the vocabulary of policy analysis and the models used by empirical growth researchers. In particular, threshold estimation can be used to identify distinct growth regimes. Formal tests indicate that the stability index can be used to divide the sample into two groups. In the relatively stable group of countries, investment has a strong effect on output, and the standard growth determinants of the Solow model, together with a measure of institutional quality, can explain 75–90 percent of the cross-section variation in growth rates. In the less stable group instability clearly reduces growth, the Solow variables have less explanatory power, investment is less effective, and the residual variance is much higher. The results also suggest that good institutions are not strongly associated with growth unless macroeconomic stability is also in place. These patterns support the commonsense view that some degree of stability is a necessary condition for rapid growth, even when a separate role for institutions is taken into account. Viewed as a whole, the results indicate that the conclusion of some recent research—that macroeconomic stability is largely irrelevant—may be premature.

Appendix

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

Assuming for now that Dollar's (1992) procedure is effective in modeling nontradables prices, a remaining question is whether differences in tradables prices reflect trade restrictions or exchange rate policies. Exchange rate policies would be more relevant to this article. 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 driven partly by trade costs, which in turn could reflect geographic characteristics. They show that about half the variation in the original Dollar measure can be explained by a combination of the black market exchange rate premium, regional dummy variables, and two geographic indicators—one measuring the ratio of coastal length to land area and the other a dummy variable 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 by variation in trade barriers.

This provides only partial support for the use of *OVERVALU* to measure macroeconomic policy outcomes. This article assumes that the cross-section variation in *OVERVALU* reflects primarily differences in national exchange rate policies. Given that other interpretations are possible, it is worth examining what happens when *OVERVALU* is omitted from the set of indicators developed in section II. Recalculating the principal components based on four indicators rather than five yields the following index:

$$MACROND = 0.332 * SURPLUS - 0.516 * INFLA - 0.615 * BMP - 0.495 * ERATE \quad \text{(A-1)}$$

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

References

Acemoglu, Daron, Simon Johnson, and James Robinson. 2001. "Colonial Origins of Comparative Development: An Empirical Investigation." *American Economic Review* 91(5):1369–401.

Acemoglu, Daron, Simon Johnson, James Robinson, and Yunyong Thaicharoen. 2003. "Institutional Causes, Macroeconomic Symptoms: Volatility, Crises, and Growth." *Journal of Monetary Economics* 50(1):49–123.

Anderson, T.W., and Herman Rubin. 1949. "Estimation of the Parameters of a Single Equation in a Complete System of Stochastic Equations." *Annals of Mathematical Statistics* 20(1):46–63.

Barro, Robert J. 1996. "Inflation and Growth." *Federal Reserve Bank of St. Louis Review* 78(3):153–69.

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

Beaudry, Paul, and Fabrice Collard. 2006. "Globalization, Returns to Accumulation and the World Distribution of Output." *Journal of Monetary Economics* 53(5):879–909.

Beaudry, Paul, Fabrice Collard, and David A. Green. 2005. "Changes in the World Distribution of Output Per Worker, 1960–1998: How a Standard Decomposition Tells an Unorthodox Story." *Review of Economics and Statistics* 87(4):741–53.

Bleaney, Michael. 1996. "Macroeconomic Stability, Investment and Growth in Developing Countries." *Journal of Development Economics* 48(2):461–77.

Brock, William A., Steven N. Durlauf, and Kenneth D. West. 2003. "Policy Evaluation in Uncertain Economic Environments." *Brookings Papers on Economic Activity* 34: 235–322.

Bruno, Michael, and William Easterly. 1998. "Inflation Crises and Long-run Growth." *Journal of Monetary Economics* 41(1):3–26.

Burnside, Craig, and David Dollar. 2000. "Aid, Policies, and Growth." *American Economic Review* 90(4):847–68.

Cameron, A. Colin, and Pravin K. Trivedi. 2005. *Microeconometrics*. Cambridge, UK: Cambridge University Press.

Cook, Paul, and Yuchiro Uchida. 2003. "Privatisation and Economic Growth in Developing Countries." *Journal of Development Studies* 39(6):121–54.

Crespo Cuaresma, Jesus, and Gernot Doppelhofer. 2007. "Nonlinearities in Cross-Country Growth Regressions: A Bayesian Averaging of Thresholds (BAT) Approach." *Journal of Macroeconomics* 29(3):541–54.

Dehon, Catherine, Marjorie Gassner, and Vincenzo Verardi. 2009. "Beware of 'Good' Outliers and Overoptimistic Conclusions." *Oxford Bulletin of Economics and Statistics* 71(3):437–52.

Desdoigts, Alain. 1996. "Smoothing Techniques Applied to a Key Economic Issue: The 'Convergence' Hypothesis." *Computational Statistics* 11(4):481–94.

———. 2004. "Neoclassical Convergence Versus Technological Catch-Up: A Contribution for Reaching a Consensus." *Problems and Perspectives in Management* 3: 15–42.

DiNardo, John, Nicole M. Fortin, and Thomas Lemieux. 1996. "Labor Market Institutions and the Distribution of Wages, 1973–1992: A Semi-Parametric Approach." *Econometrica* 64(5):1001–44.

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.

Durlauf, Steven N., Paul A. Johnson, and Jonathan R.W. Temple. 2005. "Growth Econometrics." In Philippe Aghion, and Steven N. Durlauf, eds., *Handbook of Economic Growth*. Vol. 1A. Amsterdam: North-Holland.

———. 2009. "The Methods of Growth Econometrics." In Terence C. Mills and Kerry Patterson, eds., *Palgrave Handbook of Econometrics*. Vol. 2: Applied Econometrics. London: Palgrave Macmillan.

Durlauf, Steven N., Andros Kourtellos, and Chih Ming Tan. 2008. "Are Any Growth Theories Robust?" *Economic Journal* 118(527):329–46.

Easterly, William. 2001a. *The Elusive Quest for Growth: Economists' Adventures and Misadventures in the Tropics*. Cambridge, Mass.: MIT Press.

———. 2001b. "The Lost Decades: Developing Countries' Stagnation in Spite of Policy Reform 1980–1998." *Journal of Economic Growth* 6(2):135–57.

———. 2005. "National Policies and Economic Growth: A Reappraisal." In Philippe Aghion and Steven N. Durlauf, eds., *Handbook of Economic Growth*. Vol. 1A. Amsterdam: North-Holland.

Easterly, William, and Mirvat Sewadeh. 2002. *Global Development Network Growth Database*. World Bank, Washington, DC.

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

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

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

Eicher, Theo S., Chris Papageorgiou, and Oliver Roehn. 2007. "Unraveling the Fortunes of the Fortunate: An Iterative Bayesian Model Averaging (IBMA) Approach." *Journal of Macroeconomics* 29(3):494–514.

Falvey, Rod, and Norman Gemmell. 1998. "Why Are Prices So Low in Asia?" *World Economy* 21(7):897–911.

———. 1999. "Factor Endowments, Nontradables Prices and Measures of 'Openness'." *Journal of Development Economics* 58(1):101–22.

Fernandez, Carmen, Eduardo Ley, and Mark F.J. Steel. 2001. "Model Uncertainty in Cross-Country Growth Regressions." *Journal of Applied Econometrics* 16(5):563–76.

Fischer, Stanley. 1991. "Growth, Macroeconomics, and Development." In Olivier Jean Blanchard, and Stanley Fischer, eds, *NBER Macroeconomics Annual 1991*. Cambridge, Mass.: MIT Press.

———. 1993. "The Role of Macroeconomic Factors in Growth." *Journal of Monetary Economics* 32(3):485–512.

Frankel, Jeffrey A., and David Romer. 1999. "Does Trade Cause Growth?" *American Economic Review* 89(3):379–99.

Glaeser, Edward L., Rafael La Porta, Florencio Lopez-de-Silanes, and Andrei Shleifer. 2004. "Do Institutions Cause Growth?" *Journal of Economic Growth* 9(3):271–303.

Hall, Robert E., and Charles I. Jones. 1999. "Why Do Some Countries Produce So Much More Output per Worker than Others?" *Quarterly Journal of Economics* 114(1):83–116.

Hansen, Bruce E. 1996. "Inference When a Nuisance Parameter Is Not Identified Under the Null Hypothesis." *Econometrica* 64(2):413–30.

———. 2000. "Sample Splitting and Threshold Estimation." *Econometrica* 68(3):575–603.

Hausmann, Ricardo, Lant Pritchett, and Dani Rodrik. 2005. "Growth Accelerations." *Journal of Economic Growth* 10(4):303–29.

Hausmann, Ricardo, Dani Rodrik, and Andrés Velasco. 2008. "Growth Diagnostics." In Narcís Serra and Joseph E. Stiglitz, eds., *The Washington Consensus Reconsidered: Towards a New Global Governance*. Oxford: Oxford University Press.

Henisz, Witold. 2000. "The Institutional Environment for Economic Growth." *Economics and Politics* 12(1):1–31.

Henry, Peter Blair, and Conrad Miller. 2009. "Institutions Versus Policies: A Tale of Two Islands." *American Economic Review* 99(2):261–67.

Heston, Alan, Robert Summers, and Bettina Aten. 2002. Penn World Table Version 6.1. University of Pennsylvania, Center for International Comparisons, Philadelphia, Penn.

Hoeting, Jennifer, Adrian E. Raftery, and David Madigan. 1996. "A Method for Simultaneous Variable Selection and Outlier Identification in Linear Regression." *Computational Statistics and Data Analysis* 22(3):251–70.

Hubert, Mia, Peter J. Rousseeuw, and Karlien Vanden Branden. 2005. "ROBPCA: A New Approach to Robust Principal Component Analysis." *Technometrics* 47(1):64–79.

Kaufmann, Daniel, Massimo Mastruzzi, and Diego Zavaleta. 2003. "Sustained Macroeconomic Reforms, Tepid Growth: A Governance Puzzle in Bolivia?" In Dani Rodrik, ed., *In Search of Prosperity*. Princeton, NJ: Princeton University Press.

Kaufmann, Daniel, Aart Kraay, and Massimo Mastruzzi. 2005. "Governance Matters IV: Governance Indicators for 1996–2004." World Bank, Washington, DC.

La Porta, Rafael, Florencio Lopez-de-Silanes, and Andrei Shleifer. 2008. "The Economic Consequences of Legal Origins." *Journal of Economic Literature* 46(2):285–332.

Leamer, Edward E. 1978. *Specification Searches: Ad-Hoc Inference with Non-Experimental Data*. New York: John Wiley.

Levine, Ross, and David Renelt. 1992. "A Sensitivity Analysis of Cross-Country Growth Regressions." *American Economic Review* 82(4):942–63.

Malik, Adeel, and Jonathan R.W. Temple. 2009. "The Geography of Output Volatility." *Journal of Development Economics* 90(2):163–178.

Mankiw, N.Gregory, David Romer, and David N. Weil. 1992. "A Contribution to the Empirics of Economic Growth." *Quarterly Journal of Economics* 107(2):407–37.

Marshall, Monty, and Keith Jaggers. 2002. "Polity IV Project: Political Regime Characteristics and Transitions, 1800–2002." University of Maryland, College Park, Center for International Development and Conflict Management.

Minier, Jenny. 2007. "Institutions and Parameter Heterogeneity." *Journal of Macroeconomics* 29(3):595–611.

Montiel, Peter, and Luis Servén. 2006. "Macroeconomic Stability in Developing Countries: How Much Is Enough?" *World Bank Research Observer* 21(2):151–78.

Moreira, Marcelo J. 2003. "A Conditional Likelihood Ratio Test for Structural Models." *Econometrica* 71(4):1027–48.

Papageorgiou, Chris. 2002. "Trade as a Threshold Variable for Multiple Regimes." *Economics Letters* 77(1):85–91.

Pritchett, Lant. 2000. "Understanding Patterns of Economic Growth: Searching for Hills among Plateaus, Mountains, and Plains." *World Bank Economic Review* 14(2):221–50.

Quah, Danny T. 1996. "Twin Peaks: Growth and Convergence in Models of Distribution Dynamics." *Economic Journal* 106(437):1045–55.

Raftery, Adrian E. 1995. "Bayesian Model Selection in Social Research." In Peter V. Marsden, ed., *Sociological Methodology.* Cambridge, UK: Blackwells.

Raftery, Adrian E., David Madigan, and Jennifer E. Hoeting. 1997. "Bayesian Model Averaging for Linear Regression Models." *Journal of the American Statistical Association* 92(437):179–91.

Rodriguez, Francisco, and Dani Rodrik. 2000. "Trade Policy and Economic Growth: A Skeptic's Guide to the Cross-National Evidence." In Ben S. Bernanke and Kenneth Rogoff, eds., *NBER Macroeconomics Annual 2000.* Cambridge, Mass.: MIT Press.

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

———. 2005. '*Why We Learn Nothing from Regressing Economic Growth on Policies*.' Cambridge, Mass: Harvard University.

———. 2007. *One Economics, Many Recipes.* Princeton, NJ: Princeton University Press.

Rousseeuw, Peter J. 1984. "Least Median of Squares Regression." *Journal of the American Statistical Association* 79(388):871–80.

Rousseeuw, Peter J., and Katrien van Driessen. 1999. "A Fast Algorithm for the Minimum Covariance Determinant Estimator." *Technometrics* 41(3):212–23.

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

Sala-i-Martin, Xavier. 1991. "Comments on Fischer (1991), 'Growth, Macroeconomics, and Development.'" In Olivier Jean Blanchard, and Stanely Fischer, eds., *NBER Macroeconomics Annual 1991*. Cambridge, Mass.: MIT Press.

Sala-i-Martin, Xavier, Gernot Doppelhofer, and Ronald I. Miller. 2004. "Determinants of Long-Term Growth: A Bayesian Averaging of Classical Estimates (BACE) Approach." *American Economic Review* 94(4):813–35.

Sirimaneetham, Vatcharin, and Jonathan R.W. Temple. 2006. "Macroeconomic Policy and the Distribution of Growth Rates." Discussion Paper 06/584. University of Bristol, UK.

Solow, R.M. 2001. "Applying Growth Theory across Countries." *World Bank Economic Review* 15(2):283–88.

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

———. 2000. "Growth Regressions and What the Textbooks Don't Tell You." *Bulletin of Economic Research* 52(3):181–205.

———. 2009. "Review of *One Economics, Many Recipes: Globalization, Institutions and Economic Growth* by Dani Rodrik." *Economic Journal* 119(535):F224–F230.

Williamson, John. 1990. "What Washington Means by Policy Reform." In John Williamson, ed., *Latin American Adjustment: How Much Has Happened?* Washington, DC: Institute for International Economics.

World Bank. 2004. *World Development Indicators 2004.* CD-ROM. Washington, DC: World Bank.