

Comments on Bosworth and Collins, “The Empirics of Growth: An Update”
Brookings Panel on Economic Activity, September 4-5, 2003

Jeffrey Frankel, Harvard University

I am very sympathetic with the general approach of this paper and, for that matter, with the specifics as well, which in many ways are a statement of the “state of the art” in empirical growth analysis. I have found little to criticize, though perhaps if I were a drama critic I might wish for a more exciting conclusion.

I agree with the authors that much of the growth literature has set up some artificial all-or-nothing choices – convergence versus divergence, changes vs. levels, factor accumulation vs. Total Factor Productivity, history versus policies, trade versus institutions, and so on.... In each case, the right answer is not “all or nothing,” but “a balance of both.” Let’s call it “balanced growth theory.”

The authors have it right on the convergence debate. I agree that for most purposes, it is better to include initial income on the right hand side along with the other variables – the conditional convergence specification. Given the conditional convergence specification, it does not matter if we put the end-of-sample level of income on the left-hand side, or the change, as in this study. If the data have a strong opinion that the coefficient on initial income should be close to 0 (no convergence) or close to -1 (complete convergence), they will tell us. Usually the truth is in between, a coefficient

like .7 on initial income, implying a 30% rate of conditional convergence [over, say, 30 years].

The calculation I just mentioned assumed that initial income was expressed in log form. The authors have not done this. Instead initial income is expressed in level form, relative to the US level. They tell us at one point that their numbers imply 30% convergence since 1960, which is in line with others' estimates, but I would have preferred to have been able to read this estimate directly from the reported coefficient.

I also agree with the authors that the whole debate about whether growth is determined by capital accumulation (as in the neoclassical growth model) versus Total Factor Productivity has been a bit overdone. New growth theory was phrased as a declared overthrowing of Solow, followed by a supposed neoclassical revival¹, followed by an allegation that the neoclassical revival had gone too far. This is a little too fashion-conscious for me. I recall that the main conclusion of Solow's famous 1958 paper "Technical Change and the Aggregate Production Function" was precisely that all the action was *not* in capital accumulation, but rather (7/8) in the residual. I assume that the aficionados all recognize that the Solow residual "school of thought" is the opposite of the Solow growth model school of thought, but I am guessing that this confuses many of our students.

I do think that the decomposition into the roles of factor accumulation versus TFP is useful. I just don't think we should be surprised or disappointed when Bosworth and Collins find that it is roughly half and half. This is what they find in the simple

decomposition, as well as in the regression exercise of seeing to what extent the channel for various policies and other growth determinants runs through factor accumulation or through TFP. Some of both.

I am particularly pleased to see that the authors find comparably large roles for capital formation and TFP as channels for conditional convergence. It stands to reason that both are important: On the one hand, differences in capital/labor ratios create differences in rates of return which in turn promote equalization through such mechanisms as international capital flows, for the right countries, i.e., those that are open and stable. Meanwhile, on the other hand, such countries can also be expected to catch up to the global productivity frontier through technology transfer and emulation of state-of-the-art techniques and management practices. Many authors mention only one channel of catch-up, to the exclusion of the other; but it seems obvious that both should be important.

Nevertheless, the authors support the finding of Alwyn Young (1995) that in the case of the East Asian NICs, the growth miracles were more factor accumulation and less TFP. [Interestingly, though, Table 1 shows China the other way around.] This is a proposition that does matter, and probably deserves the attention it has gotten. Paul Krugman popularized the finding in his famous or infamous 1994 article “The Myth of the Asian Economic Miracle.” People outside of the economics profession were shocked at what seemed to be Paul’s claim that Asia’s miraculous rise from poverty to prosperity, in the span of a few decades, had been an illusion. At the time, I was slightly amused by the reaction that his article created in the world of international affairs. To me the article

¹ Mankiw, Romer and Weil (1992).

could just as easily have been titled, “The Asian Economic Miracle is Mostly Due to Saving, Education, and Urban Migration,” in which case nobody would have taken much note of it.

The authors are very careful, in their treatment of problems like measurement error, on which I will say little, and endogeneity, on which I will say much more. Accumulation of physical and human capital, and productivity, are of course likely to be endogenous. There is a sentence in the draft I received that reads “Endogeneity of both factor inputs provides caution against using growth accounts to infer a casual interpretation of the growth process.” Presumably the authors meant to warn against a *causal* interpretation; but given the carefulness of the paper, they could as easily be warning others against a *casual* interpretation.

The authors consider their most striking finding to be that there is only “minor evidence of a direct role for government policies. Instead the most important determinants appear to be factors that cannot be changed substantially in the short run. (p.31).” They are referring to the tendency of inflation, budget balance, and trade distortions to lose most of whatever statistical significance they might have had, when controlling for such deeper determinants as life expectancy, geography, and institutions. I am going to concentrate the remainder of my comments in this area.

Their finding ties in well with some other important recent research, as well as with some current trends in the practice of aid and development policy in Washington. The current trend is to say, not that such policies as macroeconomic discipline and

openness are not important, but that countries can't be artificially forced from the outside to agree to such policies, as under typical IMF or World Bank programs. Instead the country needs to "take ownership" of the reforms. If the political economy dictates transfers from rural farmers to urban workers, or if a federalist constitution gives provinces claim to income tax revenue, an agreement on paper with the IMF or World Bank to devalue the currency or reduce the budget deficit may be doomed to fail. This is the argument of a recent paper by Acemoglu, Johnson, Robinson and Thaicharon. They find econometrically that institutions offer more explanatory power than policies, which is consistent with the finding of the Bosworth-Collins paper. They also use the case study of Ghana to illustrate how an IMF-encouraged devaluation, with the aim of raising the real price of traded goods such as cocoa, can soon be offset by the governing elite, because the cocoa marketing board controls the price paid to the small inland farmers for cocoa.

But institutions are not the only candidates as deeper determinants in growth equations. The question is well-framed in a recent paper, Rodrik, Subramanian, and Trebbi (2002). The rendition that follows is similar to theirs. Three big theories of deep determinants seem to have emerged: tropical conditions, openness, and institutions. Each has been captured by some now-standard measures. Although each may be more exogenous than macro policies, each has serious endogeneity problems of its own that must be addressed. (Table 1 illustrates.) Let us consider each in turn..

(1) I would use the phrase *tropical conditions* for what some, including the authors, have taken to calling geography. By now “geography” has (belatedly) made its way deep into the literatures on trade and growth in many different ways. So it is important to clarify here what sort of geography we mean. We are talking about the natural climate, biology, and geology -- especially differences between the *tropics* and temporal zones, such as presence of malaria and other debilitating tropical diseases, agricultural pests, length of the growing season, and other climate effects.²

Table 1: DEEP DETERMINANTS OF GROWTH

Determinant:	1. Tropical conditions	2. Openness	3. Institutions
Measures	Malaria and other diseases; crop pests; length of growing season	Trade/GDP, Tariffs; FDI	Property rights, Rule of Law
Sample endogeneity problems	Suppression by man of malaria or pests	Imported investment or luxury goods; Endogenous tariffs	Regulation systems develop with income; Ratings may be subjective
Exogenous instrumental variables	Distance from equator, tropical area; temperature, rainfall; frost days	Gravity model: including remoteness, landlockedness, linguistic & historical links	European settler mortality rates; extractive industries (plantation crops and mining)

Sources: Acemoglu, Johnson, & Robinson (2001, 2002), Acemoglu, Johnson, Robinson, & Thaicharoen (2002), Easterly & Levine (2002), Engerman & Sokoloff (1997, 2002), Hall & Jones (1999), Rodrik, Subramanian, & Trebbi (2002), and Sachs (2003).

² Diamond (1997), Gallup, Sachs, and Messenger (1998), Hall and Jones (1999), and Sachs (2001).

(2) By *openness*, we mean international integration along several dimensions, but trade is the most important. A common measure is the simple ratio of trade to GDP.

(3) Finally, *Institutions*. Measures of institutional quality are usually indicators of the rule of law, protection of property rights, and the extent of constraints on the executive. [The theory is that weak institutions lead to intermittent dictatorship, and lack of any constraints preventing elites and politicians from plundering the country, and low incentives for investment. Bosworth and Collins use an average of indicators from the International Country Risk Guide covering: law and order, bureaucratic quality, corruption, expropriation risk, and government repudiation of contracts.]

I mentioned that each of the three has serious endogeneity problems. Fortunately, reasonable instruments have been proposed and implemented for each.

The presence of malaria can be partly endogenous: it was stamped out in Panama and Singapore, despite their tropical locations, by superior technology and social organization. The instrumental variables to capture the exogenous component of the tropical geography theory started out fairly

crude, and have been getting progressively better: moving from continental dummies to latitude, and from there to percent of land area in the tropics, to average temperature or number of frost days. The state of the art must be Jeff Sachs' latest (2003) measure of ecological pre-disposition to malaria, since he now has an army of biologists to figure it out for him. But the measure that Bosworth and Collins use, a composite of tropical area and frost days, should be fine.

Trade and trade policies are both endogenous. This is why I proposed an instrumental variable: geographical suitability for trade as predicted by the gravity model.³ It includes such exogenous determinants of trade as remoteness from big trading partners, landlockedness, etc. This instrumental variable idea has been widely accepted.

Institutions can also be endogenous. Many institutions -- such as the structure of financial markets, mechanisms of income redistribution and social safety nets, regulation and tax systems -- tend to evolve in response to the level of income. But the problem is worse. Not only are institutions themselves likely to be endogenous, but the measures we are talking about

³ Frankel and Romer (1999).

are subjective evaluations of institutions. I submit that if you ask international businessmen to rate the quality of Switzerland's fire departments compared to Colombia's fire departments, the Swiss will come out on top even if they don't merit it, because of the halo effect of Switzerland's general reputation. My point is that reported evaluations of institutional quality are likely to be endogenous with respect to national economic performance.

What is a good instrumental variable for institutional quality?

Acemoglu, Johnson, and Robinson (2001) proposed the mortality rate among European settlers (more precisely, among soldiers and clergymen) during the period of initial colonization. This is a better instrument than it sounds. In fact, it is probably the best we have. The theory is that, out of all the lands that Europeans colonized, only those where Europeans actually settled were given good European institutions. This theory is related to the idea of Engerman and Sokoloff (1997, 2002) that lands endowed with extractive industries and plantation crops (mining, sugar, cotton) developed institutions of slavery, inequality, dictatorship, and state control, whereas those climates suited to fishing and small farms (developed institutions based on individualism, democracy, egalitarianism, and capitalism).

Acemoglu et al chose their instrument on the reasoning that initial settler mortality rates determined whether Europeans subsequently settled in large numbers.

Bosworth and Collins find that the institutions variable tends to drive out the significance of policies, even when the Acemoglu instrument is used for institutions. This is essentially the same as the finding of a string of earlier authors, including Acemoglu et al (2002), Easterly and Levine (2002) and Hall and Jones (1999): institutions drives out the effect of policies, and geography matters primarily as a determinant of institutions.⁴ Nobody denies the important role of, e.g., macroeconomic stability; but the claim is that macroeconomic policies are merely the outcome of institutions. The conclusion has been phrased most aggressively by Rodrik et al in their title as “Institutions Rule.” Institutions trump everything else -- the effects of both tropical geography and trade pale in the blinding light of institutions. Bosworth and Collins do find that tropical geography remains significant against this onslaught, thus lining up with Sachs’ retort “Institutions Don’t Rule.”

⁴ Easterly and Levine just group openness together with other policies. Hall and Jones consider latitude a proxy for European institutions, and thus don’t distinguish the independent effect of tropical conditions.

I was a bit disappointed to see that the authors did not find a stronger role for openness to trade, despite what I am sure was their best intentions. I am not overly concerned that they did not find a big role for the Sachs-Warner (1995) measure of trade policy. I've always accepted the argument that the strategy of using trade barriers such as tariffs in a growth regression is not necessarily a solution to the problem of the endogeneity of trade.⁵ To take one problem with it, countries tend to switch away from tariff revenue as a source of public finance as they industrialize. Further, the Sachs-Warner measure of trade distortions is a bit idiosyncratic. Rodriguez and Rodrik (2000) find it is driven overwhelmingly, not by tariffs or quotas, but by the black market premium for foreign exchange, and a measure of state export monopoly, and they argue that these largely reflect policies not related to trade.

Bosworth and Collins report that they also had no luck with our approach. I am prompted to ask three questions, in increasing order of importance. First, their tables report only results where openness itself is measured by the Sachs-Warner method. I hope and assume that when they

⁵ Sala-i-Martin (1991).

used the Frankel-Romer instrument, openness was measured not by the black market premium but by the ratio of trade to GDP? Second, the gravity instrument has been updated and refined since 1995. The computed trade instrument in the original Frankel-Romer paper was driven by only two variables: bilateral distance and landlockedness, estimated for 63 countries. [Size was also an important determinant in the trade equation, but since size appears also in the growth equation as well, it can't be said to be driving the instrument.] In Frankel and Rose (2002), we extended the data set over which the gravity model is estimated to 200 diverse countries, and we added as variables in explaining bilateral trade: common languages, common borders, common colonizers, and FTA membership. Perhaps for this reason, the t-ratio on openness in the growth equation rose, from 2 to 3 in Frankel-Romer, to 3 to 5 in Frankel-Rose. [And the correlation of the gravity-based instrument with actual trade/GDP rose from .62 to .72.]

My third question, and potentially the most important, is whether the authors controlled for size, in particular, population, in the growth equation. As I see it, one simply can't test for the effect of openness without including size. Small countries are poor countries, other things equal. One of the reasons for the success of the US economy is that our 50 states stretch from sea to sea, with an internal market large enough to achieve scale economies

and diverse endowments of natural resources and other factors, and free internal trade. When smaller countries rely on international trade, it is as a second best alternative to the preferred strategy, which is to be large. If a growth equation included the 50 American states as independent observations, their ratios of trade to GDP would be much higher than the ratio for the federal union; if the equation neglected to include size as another variable, it thus would erroneously predict higher state levels of income per capita than the national average. The opportunity to trade with one's fellow citizens is even more important than the opportunity to trade with people on the other side of the border. Indeed, some authors who neglect to include size estimate a significant negative coefficient on openness, because the inverse correlation between size and openness is so strong. This is my most important comment: put size in the growth regression, and see what you get for trade then.

Any of these instrumental variables – tropical geography, gravity, or settler mortality -- could in theory also be endogenous, notwithstanding that we have already pushed back some distance in the direction of exogeneity, from policies to social structure, and then to history/geography. The quality of instrumental variables is largely in the eye of the beholder, and I am

repeatedly surprised how some beholders see some things. The proper test of the ex ante plausibility of the author's claim that a variable is a good candidate for an instrumental variable in the sense that it is econometrically exogenous is not whether one can think of a story whereby it *might* be correlated with other independent variables, but rather how convoluted and implausible-sounding the story has to be.

At a conference seven or eight years, a discussant at a conference took issue with our claim that the geographical variables specified in the gravity model made a good instrument for a country's trade. We said Botswana predictably has a relatively low ratio of trade to GDP because of its remote location, landlockedness, and high ratio of land to population. I think the discussant's point was that such geographic variables, though predetermined, might be correlated with the error term. He a story roughly along the lines that a country with large land area was more likely to have higher military spending, in turn resulting in slower economic growth. Over the years, I have often used this as a methodological example: I tell my students that you know you have a relatively good instrument according to how convoluted is the story that the discussant has to tell about its potential endogeneity.

When I have related this incident, I have omitted the name of the discussant so as not to embarrass an accomplished macroeconomist, on account of what I assumed was a day when he hadn't had sufficient time to prepare real comments. I discovered recently that he had been telling a similarly one-sided story, about how foolish these guys Frankel and Romer were for thinking that national borders were necessarily exogenous just because they were predetermined, as a way of illustrating the pitfalls of choosing instrumental variables.⁶

The point is not that the story about military spending, or many other possible variations, could not be true. One can of course tell such a story about virtually any regression growth equation, or almost any regression equation at all for that matter. The question is how plausible or likely the story sounds to the paper's readership. But the lesson is that these judgments can be in the eye of the beholder.

With that caveat, I will state my subjective ratings for the three sets of instruments in use for the three big categories of fundamental growth factors [tropical diseases and pests, trade, and institutions]. I still believe that the

⁶ The discussant was Steve Durlauf: Brock and Durlauf (2001).

trade predictions of the gravity model are a relatively good instrument for a country's openness to trade. The instruments available for tropical diseases and pests are even better.

The big challenge is institutions. I don't wish by any means to denigrate the importance of institutions. And, as I said, the settler mortality instrument is probably the best we have. I think everyone should use it, until something better comes along, and I regret that I have never used it. But I view it as not as good as the instruments for trade and tropics. For one thing, it is only available for former colonies. And there is another problem that I regard as more important. What are the big questions we are trying to answer? We already knew, long ago, that Australia, Canada, New Zealand and the United States, belonged with Europe in the list of countries that had industrialized. The big question is why they did and the third world countries didn't.⁷ Is it policies, institutions, culture...? In that light, to be told that the areas where Europeans settled did well is not exactly news. It just repeats the big data point we already had. It doesn't help us all that much choose among policies, institutions, and cultures.

⁷ There were exceptions to the rule: the failure of Argentina during the 20th century and the success of Japan, the failure of Eastern Europe during the last third of the century and the success of the East Asian NICs. But not everyone agrees over what lessons to draw from these cases.

For my final point, I would like to move on from the topic of econometric exogeneity in historical data (“what is a good instrument?”), to the different question of conceptual exogeneity in analysis of alternative future policies (“what would be the effect of a hypothetical reform?”) What if economists are called to answer questions -- like what are the best currency arrangements - in a circumstance such as post-war Iraq. Are fundamental changes in policies, structure, and institutions possible politically and socially? Or is all pre-determined by history and geography? Are statistics from past history a guide to consequences of future policy changes?

Needless to say, one can never be sure that a statistical or econometric pattern that characterizes the data in the past will continue to hold in the future . Particularly if we are talking about the future consequences of a deliberate change in policy, there are all kinds of cases where we can even predict a shift in behavior if we think about it carefully. The Lucas critique of changes in the monetary policy regime is one very famous example, but there are many more. Nevertheless, if we can *never* use past experience to predict consequences of some innovation in policy, then we might as well give up and go home. Much like the choice of exogenous instruments

within the sample period of historical data, there is little substitute for thoughtful deliberation and judicious choices in extrapolating to lessons for future policy changes. Of Rodrik's critiques of Frankel-Romer, the one that was hardest to prove wrong was that, even if the gravity variable is exogenous, there is no guarantee that changes in openness due to deliberate tariff policy will have the identical effects on growth as does variation in openness due to geographical determinants of transport costs. The point is potentially valid in theory. When I think of all the arguments arrayed on both sides of the debate over trade, however, I don't believe that a debate between globalizers and anti-globalizers over the benefits of increases in trade due to reduced transport costs would be very different from a debate over increases in trade due to reduced government barriers. Evidence on one question is relevant for the other question.

The example of settler mortality rates highlights how deep-rooted institutions can be, and how infrequently and slowly they change in general. But notwithstanding historical influences, institutions can change, and sometimes quickly. Most institutional change happens at a time of national upheaval, such as the end of a war or the birth of an independent country. [We have all been reflecting recently on how successfully Japan and Germany were remade after the end of World War II. The breakup of the

colonial empires in the 1950s through 1980s offered another opportunity that some countries seized much better than others. In the early 1990s, the ruins of the Soviet Union left an opportunity for building new institutions in many transition economies that, though it appeared frustratingly slow and erratic at the time, ten years later has begun to look better. Finally, today, such new countries as East Timor and Macedonia, or such locations of upheaval as Afghanistan and Iraq, are open to guidance on institutional design coming from the international community, more than were the nations that became independent with the original breakup of the big colonial empires several decades ago.]

The point is that even if macro or trade policies have on average been prisoner to slowly-changing institutions and their historical or geographical determinants over a particular sample period, that is not necessarily a reason to despair of the possibility of genuine policy changes in the future, or of seeking to guide such changes by the light of our discipline. True, it is useful to keep one's eyes open, and to realize that well-intentioned policies may turn out instead to be the slaves of defunct 19th century colonizers. But we defunct academic scribblers must do our parts as well.

References

- Acemoglu, Daron, Simon Johnson, James Robinson, 2001, "Colonial Origins of Comparative Development: An Empirical Investigation," American Economic Review 91, 1369-1401.
- Acemoglu, Daron, Simon Johnson, James Robinson, and Yonyong Thaicharoen, 2002, "Institutional Causes, Macroeconomic Symptoms: Volatility, Crises and Growth," NBER WP 9124, August
- Barro, Robert, 1991, "Economic Growth in a Cross Section of Countries," *Quarterly J. Ec.*, CVI, May, 407-444.
- Brock, William, and Steven Durlauf, 2001, "Growth Economics and Reality," NBER WP no. 8041; in *World Bank Economic Review*.
- Diamond, Jared, 1997, Guns, Germs and Steel, WW Norton and Co.: New York.
- Easterly, William, and Ross Levine, 2002, "Tropics, Germs, and Endowments," NBER Working Paper No. 9106; Carnegie-Rochester Conference Series on Public Policy.
- Frankel, Jeffrey, and David Romer, "Does Trade Cause Growth?" *American Economic Review*, 89, no. 3 (June 1999), 379-399.
- Frankel, Jeffrey, and Andrew Rose, 2002, "An Estimate of the Effect of Common Currencies on Trade and Income," *Quarterly Journal of Economics*, vol. CXVII, no. 2, May 2002, 437-466.
- Gallup, John, Jeffrey Sachs, and Andrew Messenger, 1998, "Geography and Economic Development," NBER Working Paper No. 6849, Dec.]

Hall, Robert, and Chad Jones, 1999, "Why Do Some Countries Produce So Much More Output per Worker than Others?" Quarterly Journal of Economics 114, 1, Feb. 83-116.

Krugman, Paul, 1994, "The Myth of Asia's Miracle," Foreign Affairs 73, no. 6, Nov./Dec., 62-78.

Mankiw, N. Gregory, David Romer and David Weil, 1992, "A Contribution to the Empirics of Economic Growth," Quarterly Journal of Economics 107, 407-437.

Rodríguez, Francisco, and Dani Rodrik, 2001, "Trade Policy and Economic Growth: A Skeptic's Guide to the Cross-National Evidence," *NBER Macroeconomics Annual 2001* (Cambridge, MA: MIT Press).

Rodrik, Dani, Arvind Subramanian, and Francesco Trebbi, "Institutions Rule: The Primacy of Institutions over Geography and Integration in Economic Development," CID Working Paper No. 97, October 2002.

Sachs, Jeffrey, 2001, "Tropical Underdevelopment," NBER Working Paper No. 8119, Feb.

Sachs, Jeffrey, 2003, "Institutions Don't Rule: Direct Effects of Geography on Per Capita Income," NBER Working Paper No. 9490, February.

Sachs, Jeffrey, and Andrew Warner, 1995, "Economic Reform and the Process of Global Integration," Brookings Papers on Economic Activity 1, pp.1-95.

Sala-i-Martin, Xavier, 1991, "Comment," NBER Macroeconomics Annual 6, 368-378.

Solow, Robert, 1957, "Technical Change and the Aggregate Production Function" Review of Economics and Statistics, August 1957, 312-330

Young, Alwyn, 1995, "The Tyranny of Numbers: Confronting the Statistical Reality of the East Asian Growth Experience," Quarterly Journal of Economics 110, no. 3, August, 641-680.