

Growth Econometrics

Steven N. Durlauf, Paul A. Johnson and Jonathan R. W. Temple

Final Draft: October 22, 2004

Durlauf thanks the University of Wisconsin and John D. and Catherine T. MacArthur Foundation for financial support. Johnson thanks the Department of Economics, University of Wisconsin for its hospitality in Fall 2003, during which part of this chapter was written. Temple thanks the Leverhulme Trust for financial support under the Philip Leverhulme Prize Fellowship scheme. Ritesh Banerjee, Ethan Cohen-Cole, Giacomo Rondina and Lisa Wong have provided excellent research assistance. Finally, we thank Gordon Anderson, William Brock and Stephen Bond for useful discussions.

Growth Econometrics

Abstract

This paper provides a survey and synthesis of econometric tools that have been employed to study economic growth. While these tools range across a variety of statistical methods, they are united in the common goals of first, identifying interesting contemporaneous patterns in growth data and second, drawing inferences on long-run economic outcomes from cross-section and temporal variation in growth. We describe the main stylized facts that have motivated the development of growth econometrics, the major statistical tools that have been employed to provide structural explanations for these facts, and the primary statistical issues that arise in the study of growth data. An important aspect of the survey is attention to the limits that exist in drawing conclusions from growth data, limits that reflect model uncertainty and the general weakness of available data relative to the sorts of questions for which they are employed.

Steven N. Durlauf
Department of Economics
University of Wisconsin
1180 Observatory Drive
Madison, WI 53706-1393
United States

Paul A. Johnson
Department of Economics
Vassar College
124 Raymond Avenue
Poughkeepsie, NY 12064-0708
United States

Jonathan R. W. Temple
Department of Economics
University of Bristol
8 Woodland Road
Bristol BS8 1TN
United Kingdom

The totality of our so-called knowledge or beliefs, from the most causal matters of geography and history to the profoundest laws of atomic physics...is a man-made fabric which impinges on experience only along the edges...total science is like a field of force whose boundary conditions are experience...A conflict with experience on the periphery occasions readjustments in the interior of the field. Reevaluation of some statements entails reevaluation of others, because of their logical interconnections...But the total field is so underdetermined by its boundary conditions, experience, that there is much latitude of choice as to what statements to reevaluate in the light of any single contrary experience.

W. V. O. Quine¹

I. Introduction

The empirical study of economic growth occupies a position that is notably uneasy. Understanding the wealth of nations is one of the oldest and most important research agendas in the entire discipline. At the same time, it is also one of the areas in which genuine progress seems hardest to achieve. The contributions of individual papers can often appear slender. Even when the study of growth is viewed in terms of a collective endeavor, the various papers cannot easily be distilled into a consensus that would meet standards of evidence routinely applied in other fields of economics.

A traditional defense of empirical growth research would be in terms of expected payoffs. Each time an empirical growth paper is written, the probability of gaining genuine understanding may be low, but the payoff to that understanding is potentially vast. But even this argument relies on being able to discriminate between the status of different pieces of evidence – the good, the bad and the ugly – and this process of discrimination carries many difficulties of its own.

Rodriguez and Rodrik (2001) begin their skeptical critique of evidence on trade policy and growth with an apt quote from Mark Twain: “It isn’t what we don’t know that kills us. It’s what we know that ain’t so.” This point applies with especial force in the identification of empirically salient growth determinants. As illustrated in Appendix 2 of this chapter, approximately as many growth determinants have been proposed as there are countries for which data are available. It is hard to believe that all these determinants are

¹“Two Dogmas of Empiricism,” *Philosophical Review*, 1951.

central, yet the embarrassment of riches also makes it hard to identify the subset that truly matters.

There are other respects in which it is difficult to reconcile alternative empirical studies, including the functional form posited for the growth process. An important distinction between the neoclassical growth model of Solow (1956) and Swan (1956) and many of the models that have been produced in the endogenous growth theory literature launched by Romer (1986) and Lucas (1988) is that the latter can require the specification of a nonlinear data generating process. But researchers have not yet agreed on the empirical specification of growth nonlinearities, or the methods that should be used to distinguish neoclassical and endogenous growth models empirically.

These and other difficulties inherent in the empirical study of growth have prompted the field to evolve continuously, and to adopt a wide range of methods. We argue that a sufficiently rich set of statistical tools for the study of growth have been developed and applied that they collectively define an area of growth econometrics. This chapter is designed to provide an overview of the current state of this field. The chapter will both survey the body of econometric and statistical methods that have been brought to bear on growth questions and provide some assessments of the value of these tools.

Much of growth econometrics reflects the specialized questions that naturally arise in growth contexts. For example, statistical tools are often used to draw inferences about long-run outcomes from contemporary behaviors. This is most clearly seen in the context of debates over economic convergence; as discussed below, many of the differences between neoclassical and endogenous growth perspectives may be reduced to questions concerning the long-run effects of initial conditions. The available growth data typically span at most 140 years (and many fewer if one wants to work with a data set that nontrivially spans countries outside Western Europe and the United States) and the use of these data to examine hypotheses about long-run behavior can be a difficult undertaking. Such exercises lead to complicated questions concerning how one can identify the steady-state behavior of a stochastic process from observations along its transition path.

As we have already mentioned, another major and difficult set of growth questions involves the identification of empirically salient determinants of growth when

the range of potential factors is large relative to the number of observations. Model uncertainty is in fact a fundamental problem facing growth researchers. Individual researchers, seeking to communicate the extent of support for particular growth determinants, typically emphasize a single model (or small set of models) and then carry out inference as if that model had generated the data. Standard inference procedures based on a single model, and which are conditional on the truth of that model, can grossly overstate the precision of inferences about a given phenomenon. Such procedures ignore the uncertainty that surrounds the validity of the model. Given that there are usually other models that have strong claims on our attention, the standard errors can understate the true degree of uncertainty about the parameters, and the choice of which models to report can appear arbitrary. The need to properly account for model uncertainty naturally leads to Bayesian or pseudo-Bayesian approaches to data analysis.²

Yet another set of questions involves the characterization of interesting patterns in a data set comprised of objects as complex and heterogeneous as countries. Assumptions about parameter constancy across units of observation seem particularly unappealing for cross-country data. On the other hand, much of the interest in growth economics stems precisely from the objective of understanding the distribution of outcomes across countries. The search for data patterns has led to a far greater use of classification and pattern recognition methods, for example, than appears in other areas of economics. Here and elsewhere, growth econometrics has imported a range of methods from statistics, rather than simply relying on the tools of mainstream econometrics.

Whichever techniques are applied, the weakness of the available data represents a major constraint on the potential of empirical growth research. Perhaps the main obstacle to understanding growth is the small number of countries in the world. This is a problem for the obvious reason (a fundamental lack of variation or information) but also because it limits the extent to which researchers can address problems such as measurement error and parameter heterogeneity. Sometimes the problem is stark: imagine trying to infer the consequences of democracy for growth in poorer countries. Because the twentieth century provided relatively few examples of stable, multi-party democracies among the

²See Draper (1995) for a general discussion of model uncertainty and Brock, Durlauf, and West (2003) for discussion of its implications for growth econometrics.

poorer nations of the world, statistical evidence can make only a limited contribution to this debate, unless one is willing to make exchangeability assumptions about nations that would seem not to be credible.³

With a larger group of countries to work with, many of the difficulties that face growth researchers could be addressed in ways that are now standard in the microeconometrics literature. For example, the well known concerns expressed by Harberger (1987), Solow (1994) and many others about assuming a common linear model for a set of very different countries could, in principle, be addressed by estimating more general models that allow for heterogeneity. This can be done using interaction terms, nonlinearities or semiparametric methods, so that the marginal effect of a given explanatory variable can differ across countries or over time. The problem is that these solutions will require large samples if the conclusions are to be robust. Similarly, some methods for addressing other problems, such as measurement error, are only useful in samples larger than those available to growth researchers. This helps to explain the need for new statistical methods for growth contexts, and why growth econometrics has evolved in such a pragmatic and eclectic fashion.

One common response to the lack of cross-country variation has been to draw on variation in growth and other variables over time, primarily using panel data methods. Many empirical growth papers are now based on the estimation of dynamic panel data models with fixed effects. Our survey will discuss not only the relevant technical issues, but also some issues of interpretation that are raised by these studies, and especially their treatment of fixed effects as nuisance parameters. We also discuss the merits of alternatives. These include the before-and-after studies of specific events, such as stock market liberalizations or democratizations, which form an increasingly popular method for examining certain hypotheses. The correspondence between these studies and the microeconomic literature on treatment effects helps to clarify the strengths and limitations of the event-study approach, and of cross-country evidence more generally.

³See Temple (2000b) and Brock and Durlauf (2001a) for a conceptual discussion of this issue.

Despite the many difficulties that arise in empirical growth research, we believe some progress has been made. Researchers have uncovered stylized facts that growth theories should endeavor to explain, and developed methods to investigate the links between these stylized facts and substantive economic arguments. We would also argue that an important contribution of growth econometrics has been the clarification of the limits that exist in employing statistical methods to address growth questions. One implication of these limits is that narrative and historical approaches (Landes (1998) and Mokyr (1992) are standard and valuable examples) have a lasting role to play in empirical growth analysis. This is unsurprising given the importance that many authors ascribe to political, social and cultural factors in growth, factors that often do not readily lend themselves to statistical analysis.⁴ For these reasons, Willard Quine's classic statement of the underdetermination of theories by data, cited at the beginning of this chapter, seems especially relevant to the study of growth.

The chapter is organized as follows. Section II describes a set of stylized facts concerning economic growth. These facts constitute the objects that formal statistical analysis has attempted to explain. Section III describes the relationship between theoretical growth models and econometric frameworks for growth, with a primary focus on cross-country growth regressions. Section IV discusses the convergence hypothesis. Section V describes methods for identifying growth determinants, and a range of questions concerning model specification and evaluation are addressed. Section VI discusses econometric issues that arise according to whether one is using cross-section, time series or panel data, and also examines the issue of endogeneity in some depth. Section VII evaluates the implications of different data and error properties for growth analysis. Section VIII concludes with some thoughts on the progress made thus far, and possible directions for future research.

⁴Narrative approaches can, of course, be subjected to criticisms every bit as severe as apply to quantitative studies. Similarly, efforts to study qualitative growth ideas using formal tools can go awry; see Durlauf (2002) for criticism of efforts to explain growth and development using the idea of social capital.

II. Stylized facts

In this section we describe some of the major features of cross-country growth data. Our goal is to identify some of the salient cross-section and intertemporal patterns that have motivated the development of growth econometrics. Section II.i makes some general observations on growth in the very long-run. Section II.ii discusses the main data set used to study growth since 1960. Section II.iii describes general facts about differences in output per worker across countries. Section II.iv extends this discussion by focusing on growth miracles and disasters. Basic facts concerning convergence are reported in Section II.v. In Section II.vi we describe the general slowdown in growth over the last two decades. Section II.vii extends this discussion by considering the question of predictability of growth rates over time. Section II.viii identifies growth differences across levels of development and across geographic regions. In Section II.ix, we characterize some aspects of stagnation and volatility. Section II.x draws some general conclusions about the basic growth facts.

i. a long-run view

Taking a long view of economic history, a central fact concerning aggregate economic activity across countries is the massive divergence in living standards that has occurred over the last several centuries. A snapshot of the world in 1700 would show all countries to be poor, if their living standards were assessed in today's terms. Over the course of the 18th and 19th centuries, growth rates increased slightly in the UK and other countries in Western Europe. Annual growth rates appear to have remained low, by modern standards, even in the midst of the Industrial Revolution; but because this growth was sustained over time, GDP per capita steadily rose. The outcome was that the UK, some other countries in Western Europe, and then the USA gradually advanced further ahead of the rest of the world.

What was happening elsewhere? As Pritchett (1997) argues, even in the absence of national accounts data, we can be almost certain that rapid productivity growth was never sustained in the poorer regions of the world. The argument proceeds by

extrapolating backwards from their current levels of GDP per capita, using a fast growth rate. This quickly implies earlier levels of income that would be too low to support human life.

ii. data after 1960

Today's overall inequality across countries is partly the legacy of rapid growth in a small group of Western economies, and its absence elsewhere. But there have been important deviations from this general pattern. Since the 1960s, some developing countries have grown at rates that are unprecedented, at least based on the experiences of the advanced economies of Europe and North America. The tiger economies of East Asia have seen GDP per worker grow at around 5% a year, or even faster, for the best part of forty years. A country that grows at such rates over forty years will see GDP per worker rise more than sevenfold, as in the case of Hong Kong, Singapore, South Korea and Taiwan.

In the rest of this section, we describe these patterns in more detail. As with most of the empirical growth literature, we will focus on the period after 1960, the point at which national accounts data start to become available for a larger group of countries.⁵ Our calculations use version 6.1 of the Penn World Table (PWT) due to Heston, Summers, and Aten (2002). They have constructed measures of real GDP that adjust for international differences in price levels, and are therefore more comparable across space than measures based on market exchange rates.⁶

For the purposes of our analysis, the “world” will consist of 102 countries, those with data available in PWT 6.1 and with populations of at least 350,000 in the year 1960. These 102 countries account for a large share of the world's population. The most important missing countries are economies in Eastern Europe that were centrally planned for much of the period. Because of its enormous population, collectivist China is included in the sample, but is a country for which output measurement is especially difficult. In a

⁵Another reason for this starting point is that many colonies did not gain independence until the 1960s.

⁶For more discussion of the PWT data, and further references, see Temple (1999).

small number of cases, data for GDP per worker for 2000 are extrapolated from preceding years using growth rates for the early and mid-1990s. The Appendix gives more details of the sample, and the extrapolation procedure.

Throughout, we use data on GDP per worker. Most formal growth models are based on production functions, and their implications relate more closely to GDP per worker than GDP per capita. Jones (1997) provides another justification for this choice. When there is an unmeasured non-market sector, such as subsistence agriculture, GDP per worker could be a more accurate index of average productivity than GDP per capita.

The paths of GDP per worker and GDP per capita will diverge when there are changes in the ratio of workers to population, which is one form of participation rate. There has been an upwards trend in these participation rates where such rates were originally low, while at the upper end of the distribution participation has been stable.⁷ For a sample of 90 countries with available data, the median participation rate rose from 41% to 45% between 1960 and 2000. There was a sharp increase at the 25th percentile (from 33% to 40%) but very little change at the 75th percentile. This pattern suggests that growth in GDP per capita has usually been close to growth in GDP per worker, except for the countries that started with low participation rates.

There is an important point to bear in mind, when interpreting our later tables and graphs, and those found elsewhere in the literature. Our unit of observation is the country. In one sense this is clearly an arbitrary way to divide the world's population, but one that can have systematic effects on perceptions of stylized facts. We can illustrate this with a specific example. Sub-Saharan Africa has many countries that have small populations, while India and China combined account for about 40% of the world's population. In a decade where India and China did relatively well, such as the 1990s, a country-based analysis will understate the overall improvement in living standards. In contrast, in a decade where Africa did relatively well, such as the 1960s, the overall growth record would appear less strong if assessed on a population-weighted basis. The point that countries differ greatly in terms of population size is important when interpreting tables, graphs and regressions that use the country as the unit of observation.

⁷The figures we use for participation rates are those implicit in the Penn World Table, 6.1.

iii. differences in levels of GDP per worker

Initially, we document the international disparities in GDP per worker. We first look at data for countries with large populations. Table 1 lists a set of countries that together account for 4.3 billion people. Of the countries with large populations, the main omissions are Germany, because of the difficulty posed by reunification, and economies that were centrally planned, including Russia.

Table 1: International Disparities in GDP per Worker

Country	Population(m, 2000)	R1960	R2000
USA	275	1	1
United Kingdom	60	.69	.69
Argentina	37	.62	.40
France	60	.60	.76
Italy	58	.55	.84
South Africa	43	.47	.34
Mexico	97	.44	.38
Spain	40	.40	.68
Iran	64	.30	.30
Colombia	42	.27	.18
Japan	127	.25	.60
Brazil	170	.24	.30
Turkey	67	.17	.24
Philippines	76	.17	.13
Egypt	64	.17	.21
Korea, Republic of	47	.15	.57
Bangladesh	131	.10	.10
Nigeria	127	.08	.02
Indonesia	210	.08	.14
Thailand	61	.07	.20
Pakistan	138	.07	.11
India	1016	.06	.10
China	1259	.04	.10
Ethiopia	64	.04	.02
Mean		.29	.35
Median		.21	.27

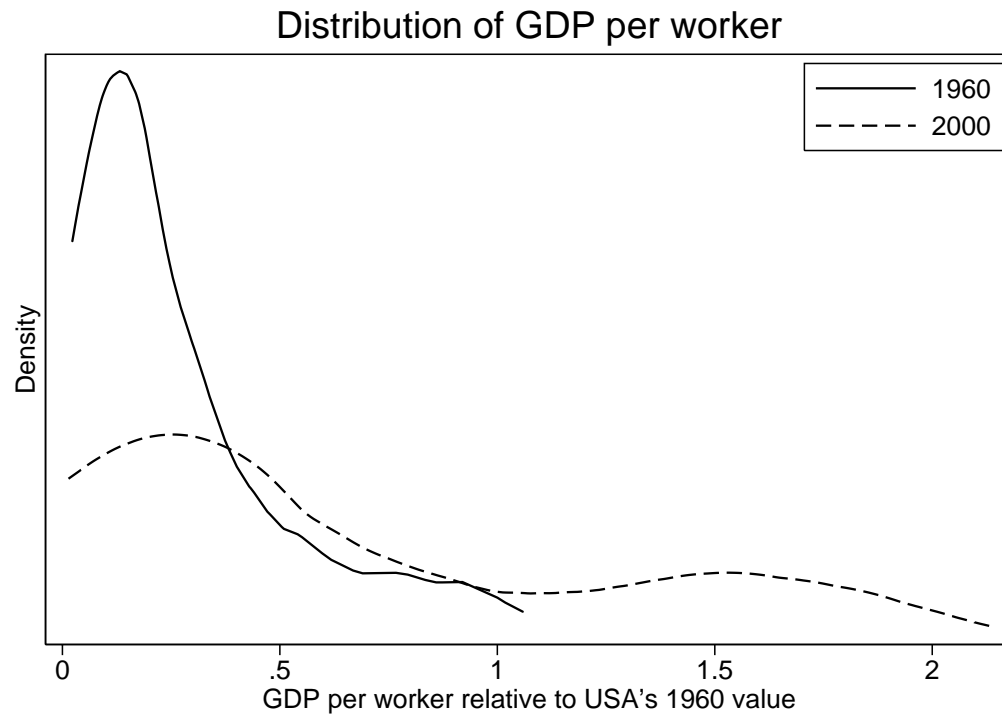
Notes:

- R is GDP per worker as a fraction of that in the USA.

The table shows GDP per worker, relative to the USA, for 1960 and 2000. The countries are ranked in descending order in terms of their 1960 position. Some clear patterns emerge: the major economies of Western Europe have maintained their position relative to the USA (as in the case of the UK) or substantially improved it (France, Italy, Spain). Among the poorer nations, there are some countries that have improved their relative position dramatically (Japan, Republic of Korea, Thailand) and others that have performed badly (Argentina, Nigeria). If we look at the mean and median of relative

GDP per worker, there has been a moderate increase, suggesting a slight tendency for reduced dispersion. But these statistics disguise a wide variety of experience, and we will discuss the issue of convergence in more detail below.

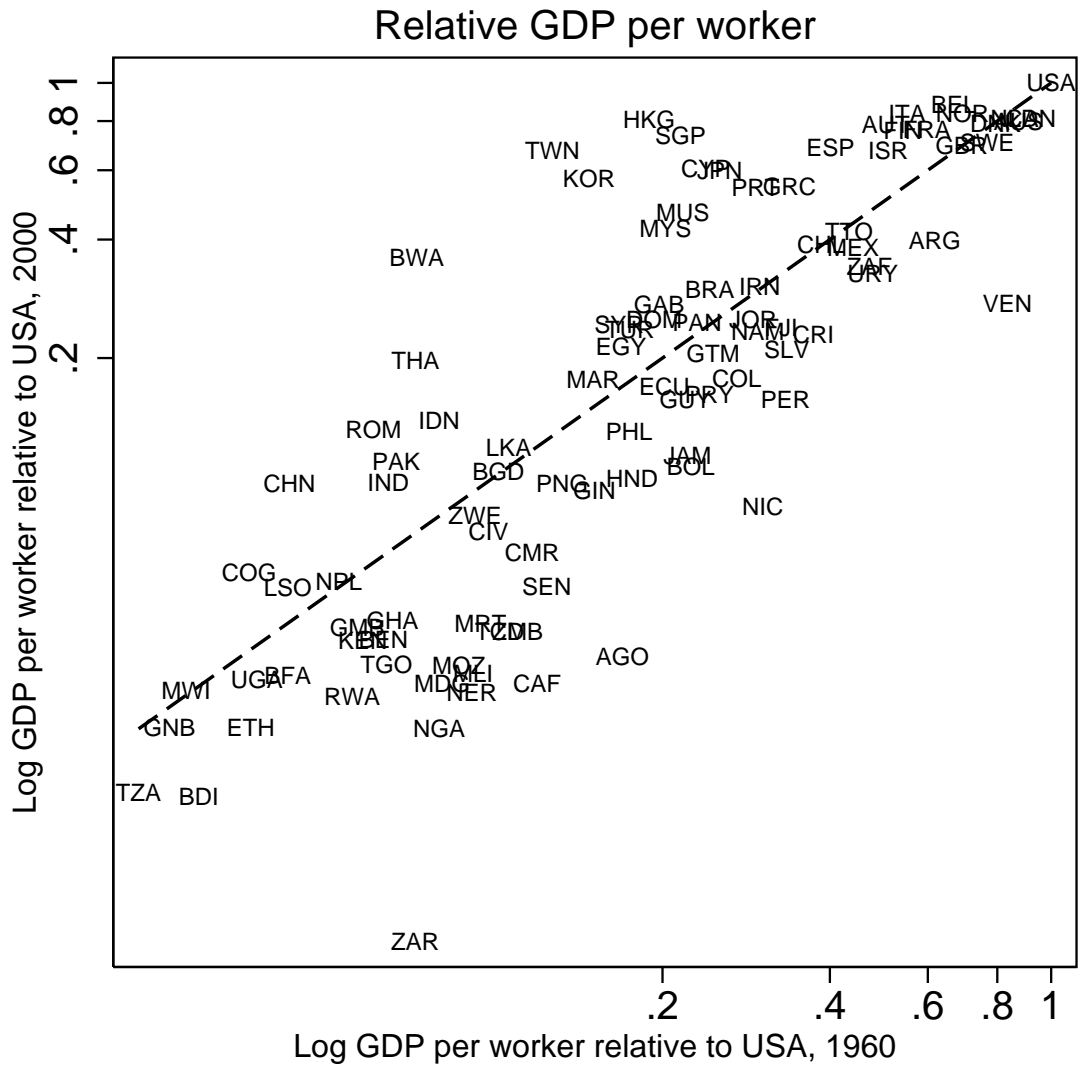
Figure 1: Cross-Country Density of Output per Worker



We now consider the shape of the international distribution of GDP per worker, using the USA's 1960 value as the benchmark. Figure 1 shows a kernel density plot of the distribution of GDP per worker in 1960 and 2000, relative to the benchmark. The rightwards movement reflects the growth that took place over this period. Also noticeable is a thinning in the middle of the distribution, the "Twin Peaks" phenomenon identified in a series of papers by Quah (1993a,b,1996a,b,c,1997).

Is the position in the league table of GDP per worker in 1960 a good predictor of that in 2000? The answer is a qualified yes: the Spearman rank correlation is 0.84. This pattern is shown in more detail in Figure 2, which plots the log of GDP per worker relative to the USA in 2000, against that in 1960. In this and later figures, one or two outlying observations are omitted to facilitate graphing.

Figure 2: Output Per Worker: 1960 versus 2000



The high rank correlation is not a new phenomenon. Easterly et al (1993) report that, for 28 countries for which Maddison (1989) has data, the rank correlation of GDP per capita in 1988 with that in 1870 is 0.82.

iv. growth miracles and disasters

Despite some stability in relative positions, it is easy to pick out countries that have done exceptionally well and others that have done badly. There is an enormous range in observed growth rates, to an extent that has not previously been observed in world history. To show this, we rank the countries by their annual growth rate between 1960 and 2000, and present a list of the fifteen best performers (Table 2) and the fifteen worst (Table 3). To show the dramatic effects of sustaining a high growth rate over forty years, we also show the ratio of GDP per worker in 2000 to that in 1960.

These tables of growth miracles and disasters show a regional pattern that is familiar to anyone who has studied recent economic growth. The best performing countries are mainly located in East Asia and Southeast Asia. These countries have sustained exceptionally high growth rates; for example, GDP per worker has grown by a factor of 11 in the case of Taiwan. If we now turn to the growth disasters, we can see many instances of “negative growth”, and these are predominantly countries in sub-Saharan Africa. Later in this section, we will compare Africa’s performance with that of other regions in more detail.⁸

⁸Easterly and Levine (1997) and Collier and Gunning (1999a,b) examine various explanations for slow growth in Africa.

Table 2: Fifteen Growth Miracles, 1960-2000

Country	Growth 1960-2000	Factor increase
Taiwan	6.25	11.3
Botswana	6.07	10.6
Hong Kong	5.67	9.09
Korea, Republic of	5.41	8.24
Singapore	5.09	7.29
Thailand	4.50	5.83
Cyprus	4.30	5.39
Japan	4.13	5.04
Ireland	4.10	5.00
China	3.99	4.77
Romania	3.91	4.63
Mauritius	3.88	4.58
Malaysia	3.82	4.48
Portugal	3.48	3.93
Indonesia	3.34	3.72

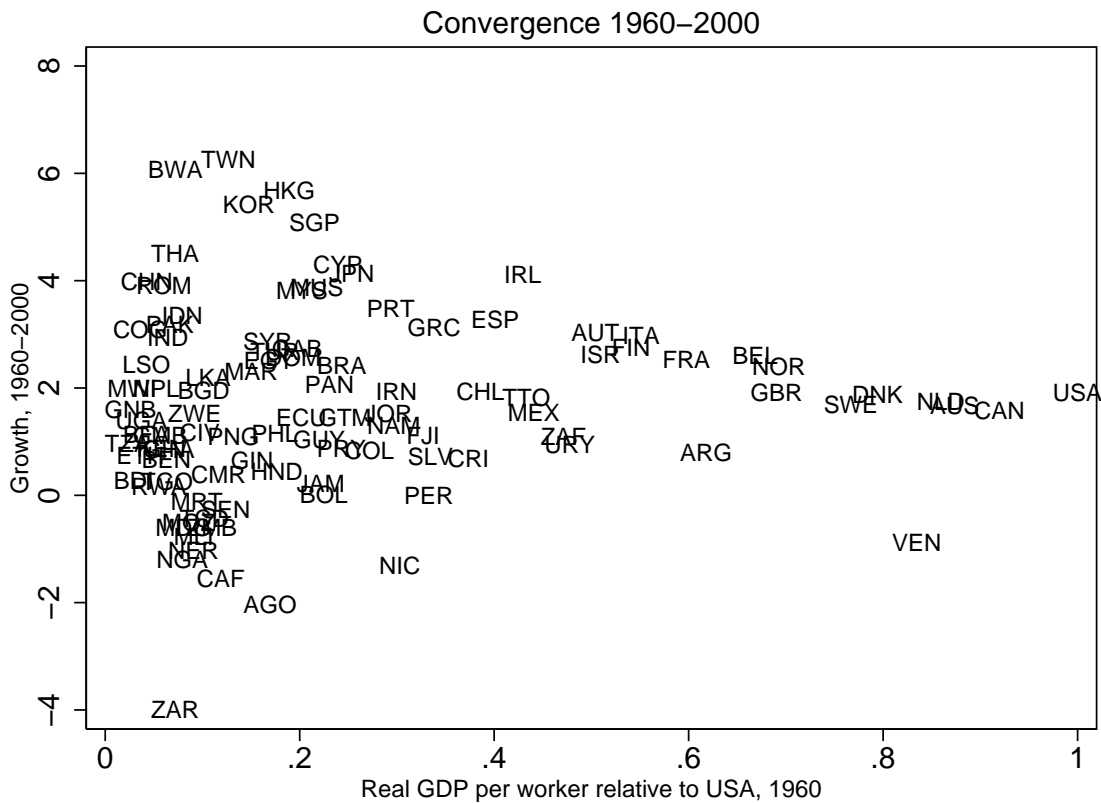
Table 3: Fifteen Growth Disasters, 1960-2000

Country	Growth 1960-2000	Ratio
Peru	0.00	1.00
Mauritania	-0.11	0.96
Senegal	-0.26	0.90
Chad	-0.43	0.84
Mozambique	-0.50	0.82
Madagascar	-0.60	0.79
Zambia	-0.61	0.78
Mali	-0.77	0.74
Venezuela	-0.88	0.70
Niger	-1.03	0.66
Nigeria	-1.21	0.62
Nicaragua	-1.30	0.59
Central African Republic	-1.56	0.53
Angola	-2.04	0.44
Congo, Democratic Rep.	-4.00	0.20

v. convergence?

An alternative way of showing the diversity of experience is to plot the growth rate over 1960-2000 against the 1960 level of real GDP per worker, relative to the USA. This is shown in Figure 3. The most obvious lesson to be drawn from this figure is the diversity of growth rates, especially at low levels of development. The figure does not provide much support for the idea that countries are converging to a common level of income, since that would require evidence of a downward sloping relationship between growth and initial income. Neither does it support the widespread idea that poorer countries have always grown slowly.

Figure 3: Growth Versus Initial Income: 1960-2000



vi. the growth slowdown

Next, we present similar figures for two sub-periods, 1960-1980 and 1980-2000. These plots, shown as Figures 4 and 5, reveal another important pattern. For many developing countries, growth was significantly lower in the second period, with many countries seeing a decline in real GDP per worker after 1980. We can see this more clearly by looking at the international distribution of growth rates for the two sub-periods. Figure 6 shows kernel density estimates, and reveals a clear pattern: the mass of the distribution has shifted leftwards (slower growth) while at the same time the variance has increased (greater dispersion in growth rates).

Figure 4: Growth Versus Initial Income 1960-1980

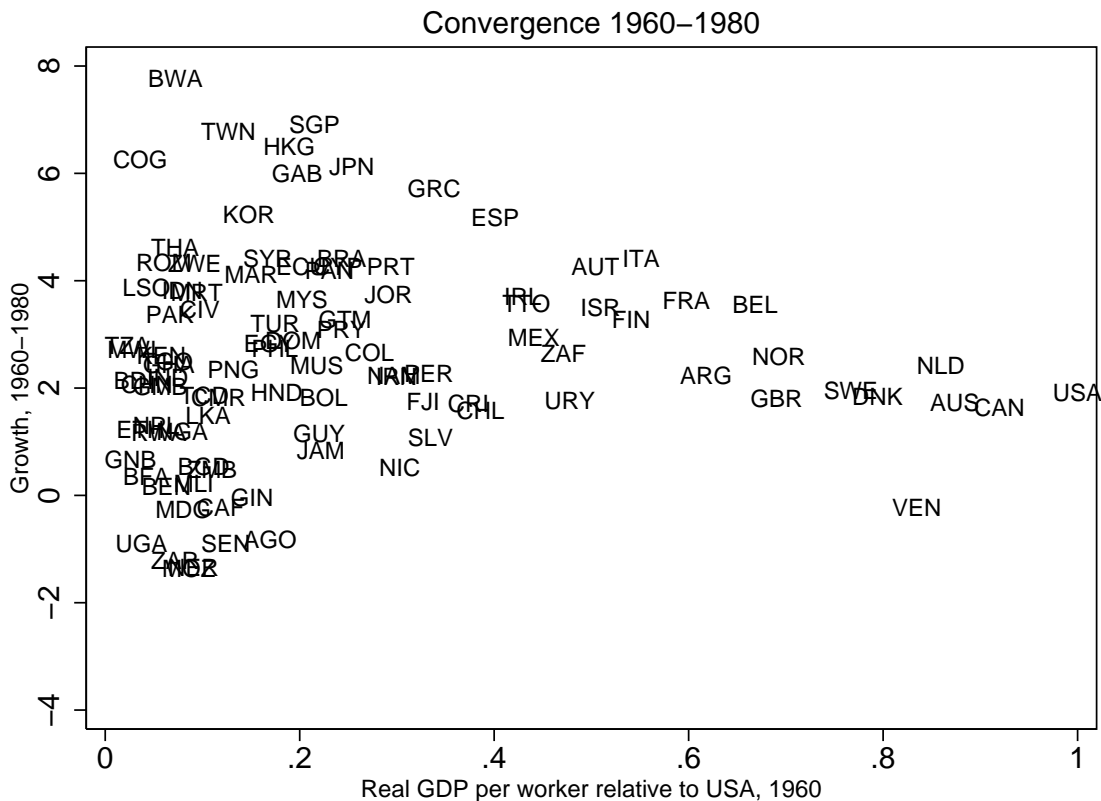


Figure 5: Growth Versus Initial Income: 1980-2000

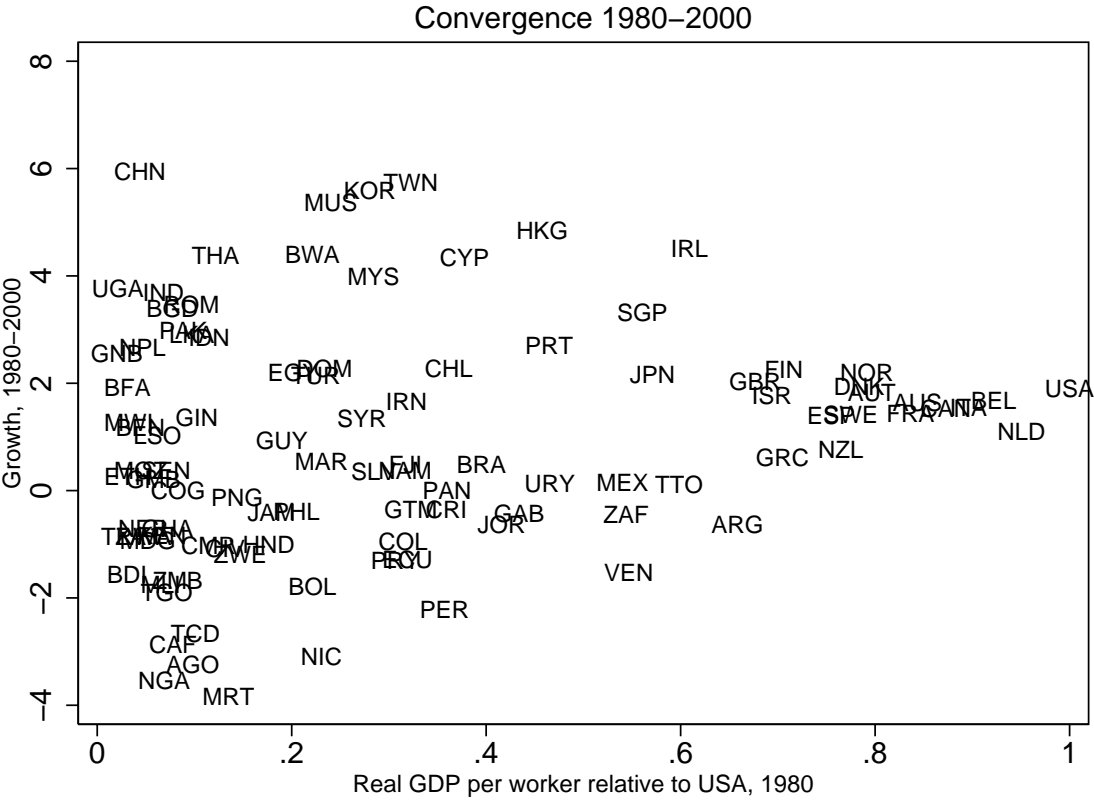


Figure 6: Density of Growth Rates across Countries

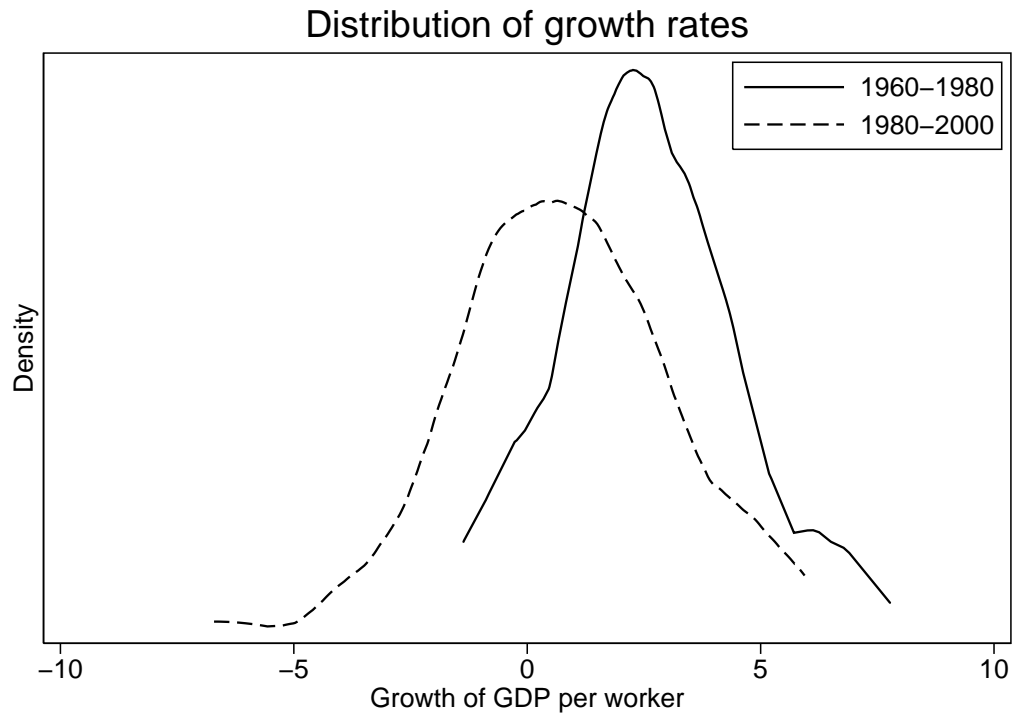
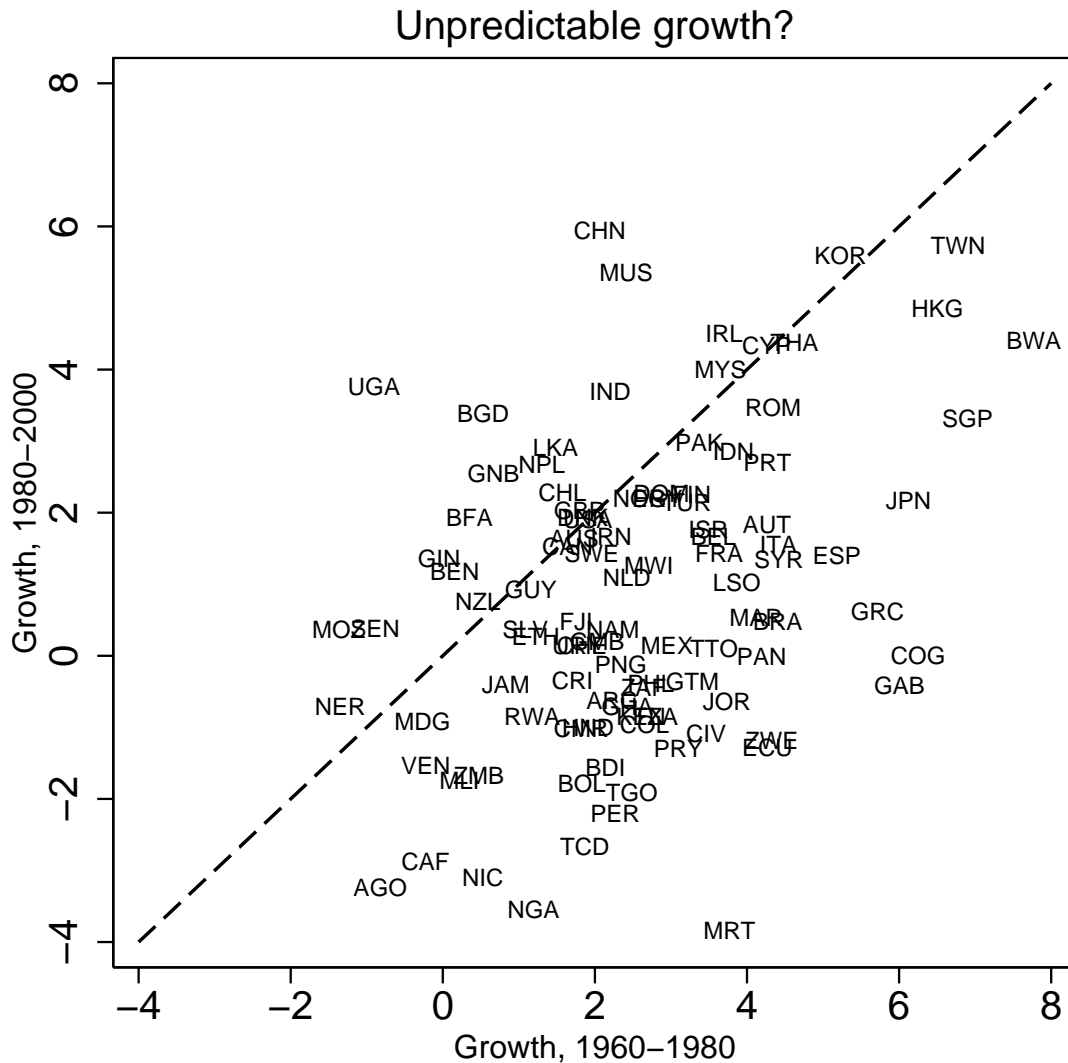


Figure 7: Growth Rates in 1960-1980 versus 1980-2000



A different way to highlight the growth slowdown is to plot the growth rate in 1980-2000 against that in 1960-1980 as is done in Figure 7, which also includes a 45 degree line. Countries above the line have seen growth increase, whereas countries below have seen growth decline. There are clearly more countries in which growth has declined over time, with the crucial exceptions of China and India, which have seen a dramatic improvement. To reveal the same pattern, Table 4 lists the countries in various categories, classified by growth rates in 1960-80 and in 1980-2000.

Table 4: Growth in 1960-1980 and 1980-2000

	$G2 \leq 0$	$0 < G2 \leq 1.5$	$1.5 < G2 \leq 3$	$G2 > 3$
$G1 \leq 0$	Angola, Central African Republic, DR Congo, Madagascar, Niger, Venezuela	Guinea, Mozambique, Senegal		Uganda
$0 < G1 \leq 1.5$	Jamaica, Mali, Nicaragua, Nigeria, Rwanda, Zambia	Benin, El Salvador, Ethiopia, Guyana, New Zealand	Burkina Faso, Guinea-Bissau, Nepal, Sri Lanka	Bangladesh
$1.5 < G1 \leq 3$	Argentina, Bolivia, Burundi, Cameroon, Chad, Colombia, Costa Rica, Ghana, Honduras, Kenya, Papua New Guinea, Peru, Philippines, South Africa, Tanzania, Togo	Fiji, Gambia, Malawi, Mexico, Namibia, Netherlands, Sweden, Switzerland, Uruguay	Australia, Canada, Denmark, Chile, Dominican Rep., Egypt, Iran, Norway, UK, USA	China, India, Mauritius
$G1 > 3$	Ecuador, Gabon, Guatemala, Ivory Coast, Jordan, Mauritania, Panama, Paraguay, Zimbabwe	Brazil, Rep. Congo, France, Greece, Lesotho, Morocco, Spain, Syria, Trinidad and Tobago	Austria, Belgium, Finland, Indonesia, Israel, Italy, Japan, Pakistan, Portugal, Turkey	Botswana, Cyprus, Hong Kong, Ireland, Korea, Malaysia, Romania, Singapore, Taiwan, Thailand

Notes:

- The above table classifies countries according to their annual growth rates over 1960-80 (G1) and over 1980-2000 (G2).

vii. does past growth predict future growth?

Another lesson to be drawn from Figure 7 and Table 4 is that relative performance has been unstable. The correlation between growth in 1960-1980 and that in 1980-2000 is just 0.40, so past growth is not a particularly useful predictor of future growth.⁹ For the whole sample, the correlations across decades are also weak (Table 5). It is less well known that the cross-decade correlation has tended to increase over time, as is clear from Table 5's below diagonal elements for the whole sample. This is tentative evidence that national economies are gradually sorting themselves into a pattern of distinct winners and losers.

Table 5: Growth Rate Correlations Across Decades

	1960-1970	1970-1980	1980-1990	1990-2000
Whole sample				
Growth 1960-1970	1.00			
Growth 1970-1980	0.16	1.00		
Growth 1980-1990	0.28	0.31	1.00	
Growth 1990-2000	0.11	0.33	0.44	1.00
Rich country group				
Growth 1960-1970	1.00			
Growth 1970-1980	0.73	1.00		
Growth 1980-1990	0.06	0.40	1.00	
Growth 1990-2000	-0.07	0.37	0.61	1.00

Notes:

- Whole sample is 102 countries. Rich country group is 19 countries.

⁹Easterly et al (1993) emphasized this point, and suggested that the lack of persistence in growth rates indicates the importance of good luck.

viii. Growth differences by development level and geographic region

Can we say anything more about the characteristics of the winners and losers? First, we investigate the relationship between growth and initial development levels in more detail. We rank the sample of 102 countries by initial income in 1960, and then look at the distribution of growth rates for subgroups. In Table 6, for various ranges of initial income relative to the USA, we show the growth rate at the 25th percentile, the median, and the 75th percentile. If we take the 22 countries which began somewhere between 5% and 10% of GDP per worker in the USA, the annual growth rate at the 25th percentile is negative, but is 2.9% at the 75th percentile. This diversity of experience extends throughout the distribution of relative incomes, but is less pronounced for the richest group.

Table 6: Growth, 1960-2000, by Initial Relative Income

Percentile	N	25 th	Median	75 th
All	102	0.7	1.6	2.7
Relative income:				
R≤0.05	10	1.0	1.5	2.4
R>0.05 & R≤0.10	22	-0.5	0.9	2.9
R>0.10 & R≤0.25	33	0.4	1.9	2.7
R>0.25 & R≤0.50	19	0.8	1.5	3.1
R>0.50	18	1.6	1.9	2.6

Notes:

- This table shows the 25th, 50th and 75th percentiles of the distribution of growth rates for countries at various levels of development in 1960.
- "R" is GDP per worker in 1960 relative to the US level.

Table 7: Growth, 1960-2000, by Country Groups

Group	N	25 th	Median	75 th
Sub-Saharan Africa	36	-0.5	0.7	1.3
South and Central America	21	0.4	0.9	1.5
East and Southeast Asia	10	3.8	4.3	5.4
South Asia	7	1.9	2.2	2.9
Industrialized countries	19	1.7	2.4	3.0

Notes:

- This table shows the 25th, 50th and 75th percentiles of the distribution of growth rates for various groups of countries.

Table 7 shows the quartiles of growth rates for countries in different regions.¹⁰ Once again, sub-Saharan Africa is revealed as a weak performer. Within sub-Saharan Africa, even the country at the 75th percentile shows growth of just 1.3%. Performance is slightly better for South and Central America, but still not strong. Against this background, the record of East and Southeast Asia looks all the more remarkable.

¹⁰These country groupings are not exhaustive; for example Fiji and Papua New Guinea do not appear in any of these groups. Analysis of the group of industrialized countries is subject to the sample selection issue highlighted by DeLong (1988).

In further work (not shown) we have constructed versions of Tables 6 and 7 for 1960-1980 and 1980-2000. These reinforce the patterns already discussed: dispersion of growth rates at all levels of development, major differences across regional groups, and a collapse in growth rates after 1980. Even for the developed countries, growth rates were noticeably lower after 1980 than before, reflecting the well-known productivity slowdown and the reduced potential for catch-up by previously fast-growing countries, such as France, Italy and Japan.

ix. stagnation and output volatility

Some countries did not record fast growth even in the boom of the 1960s. Some have simply stagnated or declined, never sustaining a high or even moderate growth rate for the length of time needed to raise output appreciably. In our sample, there are nine countries that have never exceeded their 1960 level of GDP per worker by more than 30%. Even more striking, a quarter of the countries (26 of 102) never exceeded their 1960 level by more than 60%. To put this in context, a country that grew at an average rate of 2% a year over a forty-year period would see GDP per worker rise by around 120%. Easterly (1994) drew attention to the international prevalence of stagnation, and the failure of some poorer countries to break out of low levels of development.

There are other ways in which the behavior of the poorer countries looks very different to that of rich countries. As emphasized by Pritchett (2000a), it is not uncommon for output to undergo a major collapse in less developed countries (LDCs). To show this, we calculate the largest percentage drop in output over three years recorded for each country, using data from 1960 to the latest available year. The precise statistic we calculate is:

$$100 * \left(1 - \min \left(\frac{Y_{1963}}{Y_{1960}}, \frac{Y_{1964}}{Y_{1961}}, \dots, \frac{Y_{2000}}{Y_{1997}} \right) \right)$$

The largest ten output falls are shown in Table 8, which shows how dramatic an output collapse can be. Several of these output collapses are associated with periods of

intense civil war, as in the cases of Rwanda, Angola and the Democratic Republic of the Congo. But the phenomenon of output collapse is a great deal more widespread than may be explained by events of this type. Of the 102 countries in our sample, 50 showed at least one three-year output collapse of 15% or more. 65 countries experienced a three-year output collapse of 10% or more. In contrast, between 1960 and 2000, the largest three-year output collapse in the USA was 5.4%, and in the UK 3.6%, both recorded in 1979-82. A corollary of these patterns is that time series modeling of LDC output, whether on a country-by-country basis or using panel data, has to be approached with care. It is not clear that the dynamics of output in the wake of a major collapse would look anything like the dynamics at other times.

Table 8: Output Collapses

Country	Largest 3-year drop	Dates
Chad	50%	1980-83
Rwanda	47%	1991-94
Angola	46%	1973-76
Romania	37%	1977-80
Dem. Rep. Congo	36%	1992-95
Mauritania	34%	1985-88
Tanzania	34%	1987-90
Mali	34%	1985-88
Cameroon	33%	1987-90
Nigeria	32%	1997-00

Notes:

- This table shows the ten countries with the largest output collapses over a three-year period, using data on GDP per worker between 1960 and the latest available year.

We conclude our consideration of stylized facts by briefly reporting some evidence on long-run output volatility. Table 9 reports figures on the standard deviation of annual growth rates between 1960 and 2000. Industrialized countries are relatively stable, while sub-Saharan Africa is by far the most volatile region, followed by South and Central America. Volatility is not uniformly higher in developing countries, however: using the standard deviation of annual growth rates, South Africa is less volatile than the USA, Sri Lanka less volatile than Canada, and Pakistan less volatile than Switzerland.

Table 9: Volatility, 1960-2000, by Regions

Group	N	25 th	Median	75 th
Sub-Saharan Africa	36	5.5	7.4	9.3
South and Central America	21	3.9	4.8	5.4
East and Southeast Asia	10	3.8	4.1	4.7
South Asia	7	3.0	3.3	5.2
Industrialized countries	19	2.3	2.9	3.5

Notes:

- This table shows the 25th, 50th and 75th percentiles of the distribution of the standard deviation of annual growth rates, using data from the earliest available year until the latest available, between 1960 and 2000.

x. a summary of the stylized facts

The stylized facts we consider can be summarized as follows:

1. Over the forty-year period as a whole, most countries have grown richer, but vast income disparities remain. For all but the richest group, growth rates have differed to an unprecedented extent, regardless of the initial level of development.
2. Although past growth is a surprisingly weak predictor of future growth, it is slowly becoming more accurate over time, and so distinct winners and losers are beginning to emerge. The strongest performers are located in East and Southeast Asia, which have sustained growth rates at unprecedented levels. The weakest performers are predominantly located in sub-Saharan Africa, where some countries have barely grown at all, or even become poorer. The record in South and Central America is also distinctly mixed. In these regions, output volatility is high, and dramatic output collapses are not uncommon.
3. For many countries, growth rates were lower in 1980-2000 than in 1960-1980, and this growth slowdown is observed throughout most of the income distribution. Moreover, the dispersion of growth rates has increased. A more optimistic reading would also

emphasize the growth take-off that has taken place in China and India, home to two-fifths of the world's population and a greater proportion of the world's poor.

Even this brief overview of the stylized facts reveals that there is much of interest to be investigated and understood. The field of growth econometrics has emerged through efforts to interpret and understand these facts in terms of simple statistical models, and in the light of predictions made by particular theoretical structures. In either case, the complexity of the growth process and the paucity of the available data combine to suggest that scientific standards of proof are unattainable. Perhaps the best this literature can hope for is to constrain what can legitimately be claimed.

Researchers such as Levine and Renelt (1991) and Wacziarg (2002) have argued that, seen in this more modest light, growth econometrics can provide a signpost to interesting patterns and partial correlations, and even rule out some versions of the world that might otherwise seem plausible. Seen in terms of establishing stylized facts, empirical studies help to broaden the demands made of future theories, and can act as a discipline on quantitative investigations using calibrated models. In the remainder of this chapter, we will discuss in more detail the uses and limits of statistical evidence. We first examine how empirical growth studies are related to theoretical models, and then return in more depth to the study of convergence.

III. Cross-country growth regressions: from theory to empirics

The stylized facts of economic growth have led to two major themes in the development of formal econometric analyses of growth. The first theme revolves around the question of convergence: are contemporary differences in aggregate economies transient over sufficiently long time horizons? The second theme concerns the identification of growth determinants: which factors seem to explain observed differences in growth? These questions are closely related in that each requires the specification of a statistical model of cross-country growth differences from which the effects on growth of various factors, including initial conditions, may be identified. In this section, we

describe how statistical models of cross-country growth differences have been derived from theoretical growth models.

Section III.i provides a general theoretical framework for understanding growth dynamics. The framework is explicitly neoclassical and represents the basis for most empirical growth work; even those studies that have attempted to produce evidence in favor of endogenous or other alternative growth theories have generally used the neoclassical model as a baseline from which to explore deviations. Section III.ii examines the relationship between this theoretical model of growth dynamics and the specification of a growth regression. This transition from theory to econometrics produces the canonical cross-country growth regression.

i. growth dynamics: basic ideas

For economy i at time t , let $Y_{i,t}$ denote output, $L_{i,t}$ the labor force assumed to obey $L_{i,t} = L_{i,0}e^{n_i t}$ where the population growth rate n_i is constant, and $A_{i,t}$ the efficiency level of each worker with $A_{i,t} = A_{i,0}e^{g_i t}$ where g_i is the (constant) rate of (labor augmenting) technological progress. We will work with two main per capita notions: output per efficiency unit of labor input, $y_{i,t}^E = \frac{Y_{i,t}}{A_{i,t}L_{i,t}}$ and output per labor unit $y_{i,t} = \frac{Y_{i,t}}{L_{i,t}}$.

As is well known, the generic one-sector growth model, in either its Solow-Swan or Ramsey-Cass-Koopmans variant, implies, to a first-order approximation, that

$$\log y_{i,t}^E = (1 - e^{-\lambda_i t}) \log y_{i,\infty}^E + e^{-\lambda_i t} \log y_{i,0}^E \quad (1)$$

where $y_{i,\infty}^E$ is the steady-state value of $y_{i,t}^E$ and $\lim_{t \rightarrow \infty} y_{i,t}^E = y_{i,\infty}^E$. The parameter λ_i (which must be positive) measures the rate of convergence of $y_{i,t}^E$ to its steady-state value

and depends on the other parameters of the model. Given $\lambda_i > 0$, the value of $y_{i,\infty}^E$ is independent of $y_{i,0}^E$ so that, in this sense, initial conditions do not matter in the long-run.¹¹

Eq. (1) expresses growth dynamics in terms of the unobservable $y_{i,t}^E$. In order to describe dynamics in terms of the observable variable $y_{i,t}$ we can write equation (1) as

$$\log y_{i,t} - g_i t - \log A_{i,0} = (1 - e^{-\lambda_i t}) \log y_{i,\infty}^E + e^{-\lambda_i t} (\log y_{i,0} - \log A_{i,0}) \quad (2)$$

so that

$$\log y_{i,t} = g_i t + (1 - e^{-\lambda_i t}) \log y_{i,\infty}^E + (1 - e^{-\lambda_i t}) \log A_{i,0} + e^{-\lambda_i t} \log y_{i,0} \quad (3)$$

In parallel to equation (1), one can easily see that

$$\lim_{t \rightarrow \infty} (y_{i,t} - y_{i,\infty}^E A_{i,0} e^{g_i t}) = 0 \quad (4)$$

so that the initial value of output per worker has no implications for its long-run value.

This description of the dynamics of output provides the basis for describing the dynamics of growth. Let

$$\gamma_i = t^{-1} (\log y_{i,t} - \log y_{i,0}) \quad (5)$$

denote the growth rate of output per worker between 0 and t . Subtracting $\log y_{i,0}$ from both sides of (3) and dividing by t yields

$$\gamma_i = g_i + \beta_i (\log y_{i,0} - \log y_{i,\infty}^E - \log A_{i,0}) \quad (6)$$

¹¹Implicit in our discussion is the assumption that $y_{i,0}^E > 0$ which eliminates the trivial equilibrium $y_{i,t}^E = 0 \forall t$.

where

$$\beta_i = -t^{-1} (1 - e^{-\lambda_i t}) \quad (7)$$

The β_i parameter will prove to play a key role in empirical growth analysis.

Equation (6) thus decomposes the growth rate in country i into two distinct components. The first component, g_i , measures growth due to technological progress, whereas the second component $\beta_i (\log y_{i,0} - \log y_{i,\infty}^E - \log A_{i,0})$ measures growth due to the gap between initial output per worker and the steady-state value, both measured in terms of efficiency units of labor. This second source of growth is what is meant by “catching up” in the literature. As $t \rightarrow \infty$ the importance of the catch-up term, which reflects the role of initial conditions, diminishes to zero.

Under the additional assumptions that the rates of technological progress, and the λ_i parameters are constant across countries, i.e. $g_i = g$, and $\lambda_i = \lambda \forall i$, (6) may be rewritten as

$$\gamma_i = g - \beta \log y_{i,\infty}^E - \beta \log A_{i,0} + \beta \log y_{i,0} \quad (8)$$

The important empirical implication of equation (8) is that, in a cross-section of countries, we should observe a negative relationship between average rates of growth and initial levels of output over any time period – countries that start out below their balanced growth path must grow relatively quickly if they are to catch up with other countries that have the same levels of steady-state output per effective worker and initial efficiency. This is closely related to the hypothesis of conditional convergence, which is often understood to mean that countries converge to parallel growth paths, the levels of which are assumed to be a function of a small set of variables.¹² Note, however, that a negative coefficient on initial income in a cross-country growth regression does not automatically imply conditional convergence in this sense, because countries might instead simply be moving toward their own different steady-state growth paths.

¹²We provide formal definitions of convergence in Section IV.i.

ii. cross-country growth regressions

Equation (8) provides the motivation for the standard cross-country growth regression that is the foundation of the empirical growth literature. Typically, these regression specifications start with (8) and append a random error term ν_i so that

$$\gamma_i = g - \beta \log y_{i,\infty}^E - \beta \log A_{i,0} + \beta \log y_{i,0} + \nu_i \quad (9)$$

Implementation of (9) requires the development of empirical analogs for $\log y_{i,\infty}^E$ and $\log A_{i,0}$. Mankiw, Romer, and Weil (1992) in a pioneering analysis, show how to do this in a way that produces a growth regression model that is linear in observable variables. In their analysis, aggregate output is assumed to obey a three-factor Cobb-Douglas production function

$$Y_{i,t} = K_{i,t}^\alpha H_{i,t}^\phi (A_{i,t} L_{i,t})^{1-\alpha-\phi} \quad (10)$$

where $K_{i,t}$ denotes physical capital and $H_{i,t}$ denotes human capital. Physical and human capital are assumed to follow the continuous time accumulation equations

$$\dot{K}_{i,t} = s_{K,i} Y_{i,t} - \delta K_{i,t} \quad (11)$$

and

$$\dot{H}_{i,t} = s_{H,i} Y_{i,t} - \delta H_{i,t} \quad (12)$$

respectively, where δ denotes the depreciation rate, $s_{K,i}$ is the saving rate for physical capital and $s_{H,i}$ is the saving rate for human capital and dots above variables denote time

derivatives. Note that the saving rates are both assumed to be time invariant. These accumulation equations, combined with the parameter constancy assumptions used to justify eq. (8) imply that the steady-state value of output per effective worker is

$$y_{i,\infty}^E = \left(\frac{s_{K,i}^\alpha s_{H,i}^\phi}{(n_i + g + \delta)^{\alpha+\phi}} \right)^{\frac{1}{1-\alpha-\phi}} \quad (13)$$

producing a cross-country growth regression of the form

$$\begin{aligned} \gamma_i = g + \\ \beta \log y_{i,0} + \beta \frac{\alpha + \phi}{1 - \alpha - \phi} \log(n_i + g + \delta) - \beta \frac{\alpha}{1 - \alpha - \phi} \log s_{K,i} - \beta \frac{\phi}{1 - \alpha - \phi} \log s_{H,i} - \beta \log A_{i,0} + \nu_i \end{aligned} \quad (14)$$

Mankiw, Romer and Weil assume that $A_{i,0}$ is unobservable and that $g + \delta$ is known. These assumptions mean that (14) is linear in the logs of various observable variables and therefore amenable to standard regression analysis.

Mankiw, Romer, and Weil argue that $A_{i,0}$ should be interpreted as reflecting not just technology, which they assume to be constant across countries, but country-specific influences on growth such as resource endowments, climate and institutions. They assume these differences vary randomly in the sense that

$$\log A_{i,0} = \log A + e_i \quad (15)$$

where e_i is a country-specific shock distributed independently of n_i , $s_{K,i}$, and $s_{H,i}$ ¹³.

Substituting this into (14) and defining $\varepsilon_i = \nu_i - \beta e_i$, we have the regression relationship

¹³This independence assumption is justified, in turn, on the basis that 1) n_i , $s_{K,i}$, and $s_{H,i}$ are exogenous in the neoclassical model with isoelastic preferences and 2) the estimated parameter values are consistent with those predicted by the model.

$$\gamma_i = g - \beta \log A + \beta \log y_{i,0} + \beta \frac{\alpha + \phi}{1 - \alpha - \phi} \log(n_i + g + \delta) - \beta \frac{\alpha}{1 - \alpha - \phi} \log s_{K,i} - \beta \frac{\phi}{1 - \alpha - \phi} \log s_{H,i} + \varepsilon_i \quad (16)$$

Using data from a group of 98 countries over the period 1960 to 1985, Mankiw Romer and Weil produce regression estimates of $\hat{\beta} = -.299$, $\hat{\alpha} = .48$ and $\hat{\phi} = .23$.^{14 15} Mankiw, Romer, and Weil are unable to reject the overidentifying restrictions present in (16). While this result is echoed in studies such as Knight, Loayza, and Villenueva (1993), other authors, Caselli, Equivel, and Lefort (1996), for example, are able to reject the restrictions.

Many cross-country regression studies have attempted to extend Mankiw, Romer and Weil by adding additional control variables Z_i to the regression suggested by (16). Relative to Mankiw, Romer and Weil, such studies may be understood as allowing for predictable heterogeneity in the steady-state growth term g_i and initial technology term $A_{i,0}$ that are assumed constant across i in (16). Formally, the $g_i - \beta \log A_{i,0}$ terms in (6) are replaced with $g - \beta \log A + \pi Z_i - \beta e_i$ rather than with $g - \beta \log A - \beta e_i$ which produced (16). (As far as we know, empirical work universally ignores the fact that $\log(n_i + g + \delta)$ should also be replaced with $\log(n_i + g_i + \delta)$.) This produces the cross country growth regression

$$\gamma_i = g - \beta \log A + \beta \log y_{i,0} + \beta \frac{\alpha + \phi}{1 - \alpha - \phi} \log(n_i + g + \delta) - \beta \frac{\alpha}{1 - \alpha - \phi} \log s_{K,i} - \beta \frac{\phi}{1 - \alpha - \phi} \log s_{H,i} + \pi Z_i + \varepsilon_i \quad (17)$$

¹⁴Based on data from the US and other economies, Mankiw, Romer, and Weil set $g + \delta = .05$ prior to estimation.

¹⁵Using $\lambda = -t^{-1} \log(1 - t\beta)$, the implied estimate of λ is .0142. The relationship $\lambda_i = (1 - \alpha - \phi)(n_i + g + \delta)$ was not imposed by Mankiw, Romer and Weil, who instead treat λ as a constant to be estimated. Durlauf and Johnson (1995, Table II, note b) show that estimating this model when λ varies with n in the way implied by the theory produces only very small changes in parameter estimates.

The regression described by (17) does not identify whether the controls Z_i are correlated with steady-state growth g_i or the initial technology term $A_{i,0}$. For this reason, a believer in a common steady-state growth rate will not be dissuaded by the finding that particular choices of Z_i help predict growth beyond the Solow regressors. Nevertheless, it seems plausible that the controls Z_i may sometimes function as proxies for predicting differences in efficiency growth g_i rather than in the initial technology $A_{i,0}$. As argued in Temple (1999), even if all countries have the same total factor productivity growth (TFP) in the long run, over a twenty- or thirty-year sample the assumption of equal TFP growth is highly implausible, so the variables in Z_i can explain these differences. That being said, the attribution of the predictive content of Z_i to initial technology versus steady state growth will entirely depend on a researcher's prior beliefs. It is possible that proper accounting of the $\log(n_i + g_i + \delta)$ term would allow for some progress in identifying g_i versus $A_{i,0}$ effects since g_i effects would imply a nonlinear relationship between Z_i and overall growth γ_i ; however this nonlinearity may be too subtle to uncover given the relatively small data sets available to growth researchers.

The canonical cross-country growth regression may be understood as a version of (17) when the cross-coefficient restrictions embedded in (17) are ignored (which is usually the case in empirical work). A generic representation of the regression is

$$\gamma_i = \beta \log y_{i,0} + \psi X_i + \pi Z_i + \varepsilon_i \quad (18)$$

where X_i contains a constant, $\log(n_i + g + \delta)$, $\log s_{K,i}$ and $\log s_{H,i}$. The variables spanned by $\log y_{i,0}$ and X_i thus represent those growth determinants that are suggested by the Solow growth model whereas Z_i represents those growth determinants that lie outside Solow's original theory.¹⁶ The distinction between the Solow variables and Z_i is

¹⁶We distinguish $\log y_{i,0}$ from the other Solow variables because of the role it plays in analysis of convergence; see Section IV for detailed discussion.

important in understanding the empirical literature. While the Solow variables usually appear in different empirical studies, reflecting the treatment of the Solow model as a baseline for growth analysis, choices concerning which Z_i variables to include vary greatly.

Equation (18) represents the baseline for much of growth econometrics. These regressions are sometimes known as Barro regressions, given Barro's extensive use of such regressions to study alternative growth determinants starting with Barro (1991). This regression model has been the workhorse of empirical growth research.¹⁷ In modern empirical analyses, the equation has been generalized in a number of dimensions. Some of these extensions reflect the application of (18) to time series and panel data settings. Other generalizations have introduced nonlinearities and parameter heterogeneity. We will discuss these variants below.

iii. interpreting errors in growth regressions

Our development of the relationship between cross-country growth regressions and neoclassical growth theories illustrates the standard practice of adding regression errors in an ad hoc fashion. Put differently, researchers usually derive a deterministic growth relationship and append an error in order to capture whatever aspects of the growth process are omitted from the model that has been developed. One problem with this practice is that some types of errors have important implications for the asymptotics of estimators. Binder and Pesaran (1999) conduct an exhaustive study of this question, one important conclusion of which is that if one generalizes the assumption of a constant rate of technical change so that technical change follows a random walk, this induces nonstationarity in many levels series, raising attendant unit root questions.

¹⁷Such regressions appear to have been employed earlier by Grier and Tullock (1989) and Kormendi and Meguire (1985). The reason these latter two studies seem to have received less attention than warranted by their originality is, we suspect, due to their appearance before endogenous growth theory emerged as a primary area of macroeconomic research, in turn placing great interest on the empirical evaluation of growth theories. To be clear, Barro's development is original to him and his linking of cross-country growth regressions to alternative growth theories was unique.

Beyond issues of asymptotics, the ad hoc treatment of regression errors leaves unanswered the question of what sorts of implicit substantive economic assumptions are made by a researcher who does this. Brock and Durlauf (2001a) address this issue using the concept of exchangeability. Basically, their argument is that in a regression such as (18), a researcher typically thinks of the errors ε_i as interchangeable across observations: different patterns of realized errors are equally likely to occur if the realizations are permuted across countries. In other words, the information available to a researcher about the countries is not informative about the error terms.

Exchangeability is a mathematical formalization of this idea and is defined as follows. For each observation i , there exists an associated information set F_i available to the researcher. In the growth context, F_i may include knowledge of a country's history or culture as well as any "economic" variables that are known. A definition of exchangeability (formally, F -conditional exchangeability) is

$$\mu(\varepsilon_1 = a_1, \dots, \varepsilon_N = a_N | F_1 \dots F_N) = \mu(\varepsilon_{\rho(1)} = a_1, \dots, \varepsilon_{\rho(N)} = a_N | F_1 \dots F_N) \quad (19)$$

where $\rho(\cdot)$ is an operator that permutes the N indices.

Many criticisms of growth regressions amount to arguments that exchangeability has been violated. For example, omitted regressors induce exchangeability violations as these regressors are elements of F . Parameter heterogeneity also leads to nonexchangeability. For these cases, the failure of nonexchangeability calls into question the interpretation of the regression. This is not always the case; heteroskedasticity in errors violates exchangeability but does not induce interpretation problems for coefficients.

Brock and Durlauf argue that exchangeability produces a link between substantive social science knowledge and error structure, i.e. this knowledge may be used to evaluate the plausibility of exchangeability. They suggest that a good empirical practice would for researchers to question whether the errors in a model are exchangeable, and if not, determine whether the violation invalidates the purposes for which the regression is being used. This cannot be done in an algorithmic fashion, but as is the case with empirical

work quite generally, requires judgments by the analyst. See Draper et al (1993) for further discussion of the role of exchangeability in empirical work.

IV. The convergence hypothesis

Much of the empirical growth literature has focused on the convergence hypothesis. Although questions of convergence predate them, recent widespread interest in the convergence hypothesis originates from Abramovitz (1986) and Baumol (1986). This interest and the availability of the requisite data for a broad cross-section of countries, due to Summers and Heston (1988,1991), spawned an enormous literature testing the convergence hypothesis in one or more of its various guises.¹⁸

In this section, we explore the convergence hypothesis. In Section IV.i we consider the specification of notions of convergence as related to the relationship between initial conditions and long-run outcomes. Section IV.ii explores the main technique that has been employed in studying long-run dependence, β -convergence. Section IV.iii considers alternative notions of convergence that focus less on the persistence of initial conditions and instead on whether the cross-section dispersion of incomes is decreasing across time. This section explores both σ -convergence and more general notions and recent methods that fall under the heading of distributional dynamics. It also considers how distributional notions of convergence may be related to definitions found in Section IV.i. Section IV.iv develops time series approaches to convergence. Section IV.v moves beyond the question of whether convergence is present to consider analyses that have attempted to identify the sources of convergence when it appears to be present.

i. convergence and initial conditions

The effect of initial conditions on long-run outcomes arguably represents the primary empirical question that has been explored by growth economists. The claim that

¹⁸See Durlauf (1996) and the subsequent papers in the July 1996 *Economic Journal*, Durlauf and Quah (1999), Islam (2003) and Barro and Sala-i-Martin (2004) for surveys of aspects of the convergence literature.

the effects of initial conditions eventually disappear is the heuristic basis for what is known as the convergence hypothesis. The goal of this literature is to answer two questions concerning per capita income differences across countries (or other economic units, such as regions). First, are the observed cross-country differences in per capita incomes temporary or permanent? Second, if they are permanent, does that permanence reflect structural heterogeneity or the role of initial conditions in determining long-run outcomes? If the differences in per capita incomes are temporary, unconditional convergence (to a common long-run level) is occurring. If the differences are permanent solely because of cross-country structural heterogeneity, conditional convergence is occurring. If initial conditions determine, in part at least, long-run outcomes, and countries with similar initial conditions exhibit similar long-run outcomes, then one can speak of convergence clubs.¹⁹

We first consider how to formalize the idea that initial conditions matter. While the discussion focuses on $\log y_{i,t}$, the log level of per capita output in country i at time t ; these definitions can in principle be applied to other variables such as real wages, life expectancy, etc. Our use of $\log y_{i,t}$ rather than $y_{i,t}$ reflects the general interest in the growth literature in relative versus absolute inequality, i.e. one is usually more interested in whether the ratio of income between two countries exhibits persistence than an absolute difference, particularly since sustained economic growth will imply that a constant levels difference is of asymptotically negligible size when relative income is considered.

We associate with $\log y_{i,t}$ initial conditions, $\rho_{i,0}$. These initial conditions do not matter in the long-run if

$$\lim_{t \rightarrow \infty} \mu(\log y_{i,t} | \rho_{i,0}) \text{ does not depend on } \rho_{i,0} \quad (20)$$

¹⁹This taxonomy is due to Galor (1996) who discusses the relationship between it and the theoretical growth literature, giving several examples of models in which initial conditions matter for long-run outcomes.

where $\mu(\cdot)$ is a probability measure. To see how this definition connects with empirical growth work, empirical studies of convergence are often focused on whether long-run per capita output depends on initial stocks of human and physical capital.

Economic interest in convergence stems from the question of whether certain initial conditions lead to persistent differences in per capita output between countries (or other economic units). One can thus use (20) to define convergence between two economies. Let $\| \cdot \|$ denote a metric for computing the distance between probability measures.²⁰ Then countries i and j exhibit convergence if

$$\lim_{t \rightarrow \infty} \left\| \mu(\log y_{i,t} | \rho_{i,0}) - \mu(\log y_{j,t} | \rho_{j,0}) \right\| = 0 \quad (21)$$

Growth economists are generally interested in average income levels; eq. (21) implies that countries i and j exhibit convergence in average income levels in the sense that

$$\lim_{t \rightarrow \infty} E(\log y_{i,t} - \log y_{j,t} | \rho_{i,0}, \rho_{j,0}) = 0. \quad (22)$$

To the extent one is interested in whether countries exhibit common steady-state growth rates, one can modify (22) to require that the limiting expected difference between $\log y_{i,t}$ and $\log y_{j,t}$ is bounded. One way of doing this is due to Pesaran (2004a) and is discussed below.

These notions of convergence can be relaxed. Bernard and Durlauf (1996) suggest a form of partial convergence that relates to whether contemporaneous income differences are expected to diminish. If $\log y_{i,0} > \log y_{j,0}$, their definition amounts to asking whether

$$E(\log y_{i,t} - \log y_{j,t} | \rho_{i,0}, \rho_{j,0}) < \log y_{i,0} - \log y_{j,0} \quad (23)$$

²⁰There is no unique or single generally agreed upon metric for measuring deviations between probability measures.

A number of modifications of these definitions have been proposed. Hall, Robertson, and Wickens (1997) suggest appending a requirement that the variance of output differences diminish to 0 over time, i.e.

$$\lim_{t \rightarrow \infty} E\left(\left(\log y_{i,t} - \log y_{j,t}\right)^2 \mid \rho_{i,0}, \rho_{j,0}\right) = 0 \quad (24)$$

so that convergence requires output for a pair of countries to behave similarly in the long-run. In our view, this is an excessively strong requirement since it does not allow one to regard the output series as stochastic in the long-run. Eq. (24) would imply that convergence does not occur if countries are perpetually subjected to distinct business cycle shocks. However, Hall, Robertson and Wickens (1997) do identify a weakness of definition (22), namely the failure to control for long-run deviations whose current direction is not predictable. To see this, suppose that $\log y_{i,t} - \log y_{j,t}$ is a random walk with current value 0. In this case, definition (22) would be fulfilled, although output deviations between countries i and j will become arbitrarily large at some future date.

In recent work, Pesaran (2004a) has proposed a convergence definition that focuses specifically on the likelihood of large long-run deviations. Specifically, Pesaran defines convergence as

$$\lim_{t \rightarrow \infty} \text{Prob}\left(\left(\log y_{i,t} - \log y_{j,t}\right)^2 < C^2 \mid \rho_{i,0}, \rho_{j,0}\right) > \pi \quad (25)$$

where C denotes a deviation magnitude and π is a tolerance probability. The idea of this definition is to focus convergence analysis on output deviations that are economically important and to allow for some flexibility with respect to the probability with which they occur.

These convergence definitions do not allow for the distinction between the long-run effects of initial conditions and the long-run effects of structural heterogeneity. From the perspective of growth theory, this is a serious limitation. For example, the distinctions between endogenous and neoclassical growth theories focus on the long-run

effects of cross-country differences initial human and physical capital stocks; in contrast, cross-country differences in preferences can have long-term effects under either theory. Hence, in empirical work, it is important to be able to distinguish between initial conditions $\rho_{i,0}$ and structural characteristics $\theta_{i,0}$. Steady state effects of initial conditions imply the existence of convergence clubs whereas steady-state effects of structural characteristics do not. In order to allow for this, one can modify (21) so that

$$\lim_{t \rightarrow \infty} \left\| \mu(\log y_{i,t} | \rho_{i,0}, \theta_{i,0}) - \mu(\log y_{j,t} | \rho_{j,0}, \theta_{j,0}) \right\| = 0 \text{ if } \theta_{i,0} = \theta_{j,0} \quad (26)$$

implies that countries i and j exhibit convergence. The notions of convergence in expected value (eq. (22)) may be modified in this way as well,

$$\lim_{t \rightarrow \infty} E(\log y_{i,t} - \log y_{j,t} | \rho_{i,0}, \theta_{i,0}, \rho_{j,0}, \theta_{j,0}) = 0 \text{ if } \theta_{i,0} = \theta_{j,0} \quad (27)$$

as can partial convergence in expected value (eq. (23)) and the other convergence concepts discussed above.

In practice, the distinction between initial conditions and structural heterogeneity generally amounts to treating stocks of initial human and physical capital as the former and other variables as the latter. As such, both the Solow variables X and the control variables Z that appear in cross-country growth regression cf. (18) are usually interpreted as capturing structural heterogeneity. This practice may be criticized if these variables are themselves endogenously determined by initial conditions, a point that will arise below.

The translation of these ideas into restrictions on growth regressions has led to a range of statistical definitions of convergence which we now examine. Before doing so, we emphasize that none of these statistical definitions is necessarily of intrinsic interest per se; rather each concept is useful only to the extent it elucidates economically interesting notions of convergence such as eq. (20). The failure to distinguish between convergence as an economic concept and convergence as a statistical concept has led to a good deal of confusion in the growth literature.

ii. β -convergence

Statistical analyses of convergence have largely focused on the properties of β in regressions of the form (18). β -convergence, defined as $\beta < 0$ is easy to evaluate because it relies on the properties of a linear regression coefficient. It is also easy to interpret in the context of the Solow growth model, since the finding is consistent with the dynamics of the model. The economic intuition for this is simple. If two countries have common steady-state determinants and are converging to a common balanced growth path, the country that begins with a relatively low level of initial income per capita has a lower capital-labor ratio and hence a higher marginal product of capital; a given rate of investment then translates into relatively fast growth for the poorer country. In turn, β -convergence is commonly interpreted as evidence against endogenous growth models of the type studied by Romer and Lucas, since a number of these models specifically predict that high initial income countries will grow faster than low initial income countries, once differences in saving rates and population growth rates have been accounted for. However, not all endogenous growth models imply an absence of β -convergence and therefore caution must be exercised in drawing inferences about the nature of the growth process from the results of β -convergence tests.²¹

There now exists a large body of studies of β -convergence, studies that are differentiated by country set, time period and choice of control variables. When controls are absent, $\beta < 0$ is known as unconditional β -convergence: conditional β -convergence is said to hold if $\beta < 0$ when controls are present. Interest in unconditional β -convergence, while not predicted by the Solow growth model except when countries have common steady-state output levels, derives from interest in the hypothesis that all countries are converging to the same growth path, which is critical in understanding the

²¹Jones and Manuelli (1990) and Kelly (1992) are early examples of endogenous growth models compatible with β -convergence. Each model produces steady state growth without exogenous technical change yet each implies relatively fast growth for initially capital poor economies.

extent to which current international inequality will persist into the far future.²² Typically, the unconditional β -convergence hypothesis is supported when applied to data from relatively homogeneous groups of economic units such as the states of the US, the OECD, or the regions of Europe; in contrast there is generally no correlation between initial income and growth for data taken from more heterogeneous groups such as a broad sample of countries of the world.²³

Many cross-section studies employing the β -convergence approach find estimated convergence rates of about 2% per year.²⁴ This result is found in data from such diverse entities as the countries of the world (after the addition of conditioning variables), the OECD countries, the US states, the Swedish counties, the Japanese prefectures, the regions of Europe, the Canadian provinces, and the Australian states, among others; it is also found in data sets that range over time periods from the 1860's through the 1990's.²⁵ Some writings go so far as to give this value a status analogous to a universal constant in physics.²⁶ In fact, there is some variation in estimated convergence

²²Formally, β -convergence is an implication of (9) if $\log y_{i,\infty}^E$ is assumed constant across countries in addition to the assumption on $\log A_{i,0}$ made in (15).

²³See Barro and Sala-i-Martin (2004, chapters 11 and 12) for application of β -convergence tests to a variety of data sets. Homogeneity can reflect self-selection as pointed out by DeLong (1988). He argues that Baumol's (1986) conclusion that unconditional β -convergence occurred over 1870-1979 among a set of affluent (in 1979) countries is spurious for this reason.

²⁴Panel studies estimates of convergence rates have typically been substantially higher than cross-section estimates. Examples where this is true for regressions that only control for the Solow variables include Islam (1995) and Lee, Pesaran, and Smith (1998). The panel approach has possible interpretation problems which we discuss in Section VI.

²⁵For example, Barro and Sala-i-Martin (1991) present results for US states and regions as well as European regions; Barro and Sala-i-Martin (1992) for US states, a group of 98 countries and the OECD; Mankiw, Romer, and Weil (1992) for several large groups of countries; Sala-i-Martin (1996a,b) for US states, Japanese prefectures, European regions, and Canadian provinces; Cashin (1995) for Australian states and New Zealand; Cashin and Sahay (1996) for Indian regions; Persson (1997) for Swedish counties; and, Shioji (2001a) for Japanese prefectures and other geographic units.

²⁶An alternative view is expressed by Quah (1996b) who suggests that the 2% finding may be a statistical artifact that arises for reasons unrelated to convergence *per se*. At the most primitive level, like any endogenous variable, the rate of convergence is determined by preferences, technology, and endowments. Operationally, this means that the rate of convergence will depend on model parameters and exogenous variables. For example, as

rates, but the range is relatively small; estimates generally range between 1% and 3%, as noted by Barro and Sala-i-Martin (1992).²⁷

Despite the many confirmations of this result now in the literature, the claim of global conditional β -convergence remains controversial; here we review the primary problems with the β -convergence literature.

a. robustness with respect to choice of control variables

In moving from unconditional to conditional β -convergence, complexities arise in terms of the specification of steady-state income. The reason for this is the dependence of the steady-state on Z . Theory is not always a good guide in the choice of elements of Z ; differences in formulations of equation (18) have led to a “growth regression industry” as researchers have added plausibly relevant variables to the baseline Solow specification. As a result, one can identify variants of (18) where convergence appears to occur as $\hat{\beta} < 0$ as well as variants where divergence occurs, i.e. $\hat{\beta} > 0$.

We discuss issues of uncertainty in the specification of growth regressions below. Here we note here that one class of efforts to address model uncertainty has led to confirmatory evidence of conditional β -convergence. This approach assigns probabilities to alternative formulations of (18) and uses these probabilities to construct statements about β that average across the different models. Doppelhofer, Miller, and

stated above, in the augmented Solow model studied by Mankiw, Romer, and Weil (1992), the relationship between the rate of convergence and the parameters of the model is $\lambda_i = (1 - \alpha - \phi)(n_i + g + \delta)$. Barro and Sala-i-Martin (2004, p. 111-113) discuss the relationship for the case of the Ramsey-Cass-Koopmans model with an isoelastic utility function and a Cobb-Douglas production function. Given this dependence, the ubiquity of the estimated 2% rate of convergence, taken at face value, appears to suggest a remarkable uniformity of preferences, technologies, and endowments across the economic units studied.

²⁷Barro and Sala-i-Martin argue that this variation reflects unobserved heterogeneity in steady-state values with more variation being associated with slower convergence. However, in as much as it is correlated with variables included in the regression equations, unobserved heterogeneity renders the parameter estimators inconsistent, which renders the estimated convergence parameter hard to interpret.

Sala-i-Martin (2004) conclude the posterior probability that initial income is part of the linear growth model is 1.00 with a posterior expected value for β of -0.013; this leads to a point estimate of a convergence rate of 1.3% per annum, which is somewhat lower than the 2% touted in the literature; Fernandez, Ley, and Steel (2001a) also find that the posterior probability that initial income is part of the linear growth model is 1.00, despite using a different set of potential models and different priors on model parameters.²⁸ We therefore conclude that the evidence for conditional β -convergence appears to be robust with respect to choice of controls.

b. identification and nonlinearity: β – convergence and economic divergence

A second problem with the β -convergence literature is an absence of attention to the relationship between β -convergence and economic convergence as defined by eq. (20) or variations based upon it. Put differently, in the β -convergence literature there is a general failure to develop tests of the convergence hypothesis that discriminate between convergent economic models and a rich enough set of non-converging alternatives.

While $\beta < 0$ is an implication of the Solow growth model and so is an implication of the baseline convergent growth model in the literature, this does not mean that $\beta < 0$ is inconsistent with economically interesting non-converging alternatives. One such example is the model of threshold externalities and growth developed by Azariadis and Drazen (1990). In this model, there is a discontinuity in the aggregate production function for aggregate economies. This discontinuity means that the steady-state behavior of a given economy depends on whether its initial capital stock is above or below this threshold; specifically, this model may exhibit two distinct steady states. (Of course, there can be any number of such thresholds.) An important feature of the Azariadis-Drazen model is that data generated by economies that are described by it can exhibit statistical convergence even when multiple steady states are present.

To illustrate this, we follow an argument in Bernard and Durlauf (1996) based on a simplified growth regression. Suppose that for every country in the sample, the Solow

²⁸Fernandez, Ley, and Steel (2001a) do not report a posterior expected value for β .

variables X_i and additional controls Z_i are identical. Suppose as well that there is no technical change or population growth. Following the standard arguments for deriving a cross-country regression specification, the growth regression implied by the Azariadis-Drazen assumption on the aggregate production function is

$$\gamma_i = k + \beta \left(\log y_{i,0} - \log y_{l(i)}^* \right) + \varepsilon_i \quad (28)$$

where $l(i)$ indicates the steady state with which country i is associated and $y_{l(i)}^*$ denotes output per capita in that steady state; all countries associated with the same steady state thus have the same $\log y_{l(i)}^*$ value.

The threshold externality model clearly does not exhibit economic convergence as defined above so long as there are at least two steady states. Yet the data generated by a cross-section of countries exhibiting multiple steady states may exhibit statistical convergence. To see this, notice that for this stylized case, the cross-country growth regression may be written as

$$\gamma_i = k + \beta \log y_{i,0} + \varepsilon_i \quad (29)$$

Since the data under study are generated by (28), this standard regression is misspecified. What happens when (29) is estimated when (28) is the data generating process? Using population moments, the estimated convergence parameter β_{ols} will equal

$$\beta_{ols} = \beta \frac{\text{cov}\left(\left(\log y_{i,0} - \log y_{l(i)}^*\right), \log y_{i,0}\right)}{\text{var}(\log y_{i,0})} = \beta \left(1 - \frac{\text{cov}\left(\log y_{l(i)}^*, \log y_{i,0}\right)}{\text{var}(\log y_{i,0})} \right) \quad (30)$$

From the perspective of tests of the convergence hypothesis, the noteworthy feature of (30) is that one cannot determine the sign of β_{ols} a priori as it depends on

$1 - \frac{\text{cov}(\log y_{l(i)}^*, \log y_{i,0})}{\text{var}(\log y_{i,0})}$, which is a function of the covariance between the initial and

steady-state incomes of the countries in the sample. It is easy to see that it is possible for β_{ols} to be negative even when the sample includes countries associated with different steady states. Roughly speaking, one would expect $\beta_{ols} < 0$ if low-income countries tend to initially be below their steady states whereas high-income countries tend to start above their steady states. While we do not claim this is necessarily the case empirically, the example does illustrate how statistical convergence (defined as $\beta < 0$) may be consistent with economic nonconvergence. Interestingly, it is even possible for the estimated convergence parameter β_{ols} to be smaller (and hence imply more rapid convergence) than the structural parameter β in (28).

Below, we review evidence of multiple steady states in the growth process. At this stage, we would note two things. First, some studies have produced evidence of multiple regimes in the sense that statistical models consistent with multiple steady states appear to better fit the cross-country data than the linear Solow model, e.g. Durlauf and Johnson (1995). Second, other studies have produced evidence of parameter heterogeneity such that β appears to depend nonlinearly on initial conditions so that it is equal to 0 for some countries; Liu and Stengos (1999) find precisely this when they reject the specification of constant β for all countries in favor of a specification in which β depends on initial income. These types of findings imply the compatibility of observed growth patterns with the existence of permanent income differences between economies with identical population growth and savings rates and access to identical technologies.

c. endogeneity

A third criticism that is sometimes made of the empirical convergence literature is based on the failure to account for the endogeneity of the explanatory regressors in growth regressions. One obvious reason why endogeneity may matter concerns the consistency of the regression estimates. This concern has led some authors to propose instrumental variables approaches to estimating β . Barro and Lee (1994) analyze

growth data in the periods 1965 to 1975 and 1975 to 1985 and use 5-year lagged explanatory variables as instruments. Barro and Lee find that the use of instrumental variables has little effect on coefficient estimates. Caselli, Esquivel, and Lefort (1996) employ a generalized method of moments (GMM) estimator to analyze a panel variant of the standard cross-country growth regression; growth in the panel is measured in 5-year intervals for 1960-1985. Their analysis produces estimates of β on the order of 10%, which is much larger than the 2% typically found.

Endogeneity raises a second identification issue with respect to the relationship between β -convergence and economic convergence: this idea appears in Cohen (1996) and Goetz and Hu (1996). Focusing on the Solow regressors, the value of β can fail to illustrate how initial conditions affect expected future income differences if the population and saving rates are themselves functions of income. Hence, $\beta \geq 0$ may be compatible with at least partial economic convergence, if the physical and human capital savings rates depend, for example, on the level of income. In contrast, $\beta < 0$ may be compatible with economic divergence if the physical and human capital accumulation rates for rich and poor are diverging across time. As such, this critique is probably best understood as a debate over what variables are the relevant initial conditions for evaluating (22) and/or (23). Cohen (1996) argues that the conventional human capital accumulation equation, in which accumulation is proportional to per capita output, is misspecified, failing to account for feedbacks from the stock of human capital to the accumulation process. This feedback means that poor countries with low initial stocks of human capital fail to accumulate human capital as quickly as richer ones. Goetz and Hu (1996) directly focus on the feedback from income to human capital accumulation.

The implications of this form of endogeneity for empirical work on convergence are mixed. Cohen (1996) concludes that a proper accounting for the dependence of human capital accumulation on initial capital stocks reconciles conditional β -convergence with unconditional β -divergence for a broad cross-section. Goetz and Hu (1996), in contrast, find that estimates of the speed of convergence are increased if one accounts for the effect of income on human capital accumulation for counties in the US South. This seems to be an area that warrants much more work.

d. measurement error

As Abramovitz (1986), Baumol (1986), DeLong (1988), Romer (1990), and Temple (1998) point out, measurement errors will tend to bias regression tests towards results consistent with the hypothesis of β -convergence. This occurs because, by construction, $\gamma_{i,t}$ is measured with positive (negative) error when $\log y_{i,0}$ is measured with negative (positive) error so there tends to be a negative correlation between the measured values of the two variables even if there is none between the true values. To see this, we ignore the issue of control variables and consider the case where growth is described by $\gamma_i = k + \beta \log y_{i,0} + \varepsilon_i$ where ε_i is independent across observations. Suppose that \log output is measured with error so that the researcher only observes $\zeta_{i,t} = \log y_{i,t} + e_{i,t}$, $t = 0, T$ where $e_{i,t}$ is a serially uncorrelated random variable with variance σ_e^2 and distributed independently of $\log y_{i,s}$ and ε_i for all i and s . The regression of observed growth rates will, under these assumptions, obey the equation

$$T^{-1}(\zeta_{i,T} - \zeta_{i,0}) = k + \beta \zeta_{i,0} + T^{-1}e_{i,T} - \left(\frac{\beta T + 1}{T}\right)e_{i,0} + \varepsilon_i \quad (31)$$

This is a classic errors in variables problem; the term $\left(\frac{\beta T + 1}{T}\right)e_{i,0}$ is negatively correlated with $\zeta_{i,0}$ which induces a negative bias in the estimate $\hat{\beta}$. In other words, the regression of observed growth rates on observed initial incomes will tend to produce an estimated coefficient that is consistent with the β -convergence hypothesis even if the hypothesis is not reflected in the actual behavior of growth rates across countries. In practice, as Temple (1998) explains, the direction of the bias is made ambiguous by the possibilities that the $e_{i,t}$ are serially dependent and that other right-hand-side (conditioning) variables are also measured with error. The actual effect of measurement error on results then becomes an empirical matter to be investigated by individual researchers.

In studying the role of the level of human capital in determining the rate of growth, Romer (1990) estimates a growth equation that has among its explanatory variables the level of per capita income at the beginning of the sample period. Consistent with the conditional β -convergence hypothesis, he finds a negative and significant coefficient on this variable when the equation is estimated by ordinary least squares. Wary of the possibility and effects of measurement error in initial income, as well as in the human capital variable – the literacy rate – Romer also estimates the equation using the number of radios per 1000 inhabitants and (the log of) per capita newsprint consumption as instruments for initial income and literacy with the result that the coefficients on both variables become insignificant “suggesting” that the OLS results are “attributable to measurement error” (p. 278).

Temple (1998) uses the measurement error diagnostics developed by Klepper and Leamer (1984), Klepper (1988), and classical method-of-moments adjustments, to investigate the effects of measurement error on the estimated rate of convergence in MRW's augmented Solow model. He finds that allowing for the possibility of small amounts of unreliability in the measurement of initial income implies a lower bound on the estimated convergence rate just above zero – too low to elevate conditional convergence to the status of a stylized fact. Barro and Sala-i-Martin (2004, pp. 472-3) use lagged values of state personal income as instruments for initial income to check for the possible effects of measurement error in their β -convergence tests for the US states. They find little change in the estimated convergence rates and conclude that measurement error is not an important determinant of their results. Barro (1991) follows the same procedure for other data sets and reaches a similar conclusion about the unimportance of measurement error in his results.

Some authors have attempted to address the sources of measurement error. Dowrick and Quiggin (1997) is a notable example in this regard in their consideration of the role of price indices in affecting convergence tests. Specifically, they examine the effect of constant price estimates of GDP on β -convergence calculations and find that when the prices used to construct these measures are based on prices in advanced economies, tendencies towards convergence are understated.

e. effects of linear approximation

There is a body of research that explores the effects of the approximations that are employed to produce the linear regression models used to evaluate β -convergence. As outlined earlier, regression tests of the β -convergence hypothesis rely on a log-linear approximation to the law of motion in a one sector neoclassical growth model. In addition to the possibility that Taylor series approximations in the nonstochastic version of the model are inadequate, Binder and Pesaran (1999) show that the standard practice of adding a random term to the log-linearized solution of a nonstochastic growth model does not necessarily produce the same behavior as associated with the explicit solution of a stochastic model.

Efforts to explore the limits of the linear approximation used in empirical growth studies have generally concluded that the approximation is reasonably accurate. Romer (2001, p. 25 n. 18) claims that the approximation will be “quite reliable” in this context and Dowrick (2004) presents results showing that the approximation to the true transition dynamics is quite good in a Solow model with a single capital good and an elasticity of output with respect to capital of $2/3$. This is larger than the typical physical capital share but it is not an unreasonable number for the sum of the shares of physical and human capital. To test for nonlinearity, Barro (1991) adds the square of initial (1960) income to one of his regressions and finds a positive estimated coefficient implying that the rate of convergence declines as income rises and that it is positive only for incomes below \$10800 – a figure that exceeds all of the 1960 income levels in his sample. However, the t -ratio for the estimated coefficient on the square of initial income is just 1.4 which represents weak evidence against the adequacy of the approximation.

How should one interpret such findings? At one level, these studies conclude that the approximation used to derive the equation used in cross-section convergence studies appears to be reasonably accurate. It follows that the previously discussed nonlinearities in the growth process found by researchers investigating the possibility of multiple steady states do not reflect the inadequacy of the linear approximation used in most cross-section studies. Put differently, evidence of nonlinearity appears to reflect deeper factors than simple approximation error from the use of a first order Taylor series expansion.

iii. Distributional approaches to convergence

A second approach to convergence focuses on the behavior of the cross-section distribution of income in levels. Unlike the β -convergence approach, the focus of this literature has been less on the question of relative locations within the income distribution, i.e. whether one can expect currently poor countries to either equal or exceed currently affluent countries, but rather on the shape of the distribution as a whole. Questions of this type naturally arise in microeconomic analyses of income inequality, in which one may be concerned with whether the gap between rich and poor is diminishing, regardless of whether the relative positions of individuals are fixed over time.

a. σ -convergence

Much of the empirical literature on the cross-country income distribution has focused on the question of the evolution of the cross-section variance of $\log y_{i,t}$. For a set of income levels let $\sigma_{\log y,t}^2$ denote the variance across i of $\log y_{i,t}$. σ -convergence is said to hold between times t and $t+T$ if

$$\sigma_{\log y,t}^2 - \sigma_{\log y,t+T}^2 > 0 \quad (32)$$

This definition is designed, like β -convergence, to formalize the idea that contemporary income differences are transitory, but does so by asking whether the dispersion of these differences will decline across time.

Recent work has attempted to identify regression specifications from which one can infer σ -convergence. Friedman (1992) and Cannon and Duck (2000) argue that it is possible to produce evidence concerning σ -convergence from regressions of the form

$$\gamma_i = T^{-1}(\log y_{i,t+T} - \log y_{i,t}) = \alpha + \pi \log y_{i,t+T} + \varepsilon_i \quad (33)$$

To see why this is so, following Cannon and Duck (2000), observe that σ -convergence requires that $\sigma_{\log y_{i,t}, \log y_{i,t+T}} < \sigma_{\log y_{i,t}}^2$. The regression coefficient in (33) may be written as

$$\pi = T^{-1} \left(1 - \frac{\sigma_{\log y_{i,t}, \log y_{i,t+T}}}{\sigma_{\log y_{i,t+T}}^2} \right) \quad (34)$$

which means that $\pi < 0$ implies $\sigma_{\log y_{i,t}, \log y_{i,t+T}} < \sigma_{\log y_{i,t}}^2$. Postiveness definiteness of the variance/ covariance matrix for $\log y_{i,t}$ and $\log y_{i,t+T}$ requires that

$$\left(\sigma_{\log y_{i,t}, \log y_{i,t+T}} \right)^2 < \sigma_{\log y_{i,t}}^2 \sigma_{\log y_{i,t+T}}^2$$

Therefore, if $\pi < 0$, then it must be the case that (32) holds.

Hence a test that accepts null hypothesis that $\pi < 0$ by implication accepts the null hypothesis of σ -convergence. But even this type of test has some difficulties. As pointed out by Bliss (1999,2000), it is difficult to interpret tests of σ -convergence since these tests presume that the data generating process is not invariant; an evolving distribution for the data makes it difficult to think about test distributions under a null. Additional issues arise when unit roots are present.

One limitation to this approach is that it is not clear how one can formulate a sensible notion of conditional σ -convergence. A particular problem in this regard is that one would not want to control for initial income in forming residuals, which would render the concept uninteresting as it could be generated by nothing more than time-dependent heteroskedasticity in the residuals. On the other hand, omitting income would render the interpretation of the projection residuals problematic since initial income is almost certain to be correlated with the variables that have been included when the residuals are formed. An economically interesting formulation of conditional σ -convergence would be a useful contribution.

b. evolution of the world income distribution

Work on σ -convergence has helped stimulate the more general study of the evolution of the world income distribution. This work involves examining the cross-

section distribution of country incomes at two or more points in time in order to identify how this cross-section distribution has changed. Of particular interest in such studies is the presence or emergence of multiple modes in the distribution. Bianchi (1997) uses nonparametric methods to estimate the shape of the cross-country income distribution and to test for multiple modes in the estimated density. He finds evidence of two modes in densities estimated for 1970, 1980, and 1989. Moreover, he finds a tendency for the modes to become more pronounced and to move further apart over time. This evidence supports the ideas of a vanishing middle as the distribution becomes increasingly polarized into “rich” and “poor” and of a growing disparity between those two groups. While such polarization might be desirable, were it the case that middle income economies were becoming high income ones, Bianchi’s evidence suggests that much of this movement represents a transition from middle income to poor. Further, by “cutting” each of the estimated densities at the anti-mode between the two modes, Bianchi is able to measure mobility within the distribution by counting the crossings of the cut points. These crossings represent countries moving from one basin of attraction to the other. Just 3 of the possible 238 crossings are observed.²⁹ The implication is that there is very little mobility within the cross-country income distribution. The 20 or so countries in the “rich” basin of attraction in 1970 are still there in 1989 and similarly for the 100 or so countries starting in the “poor” basin.

Paap and van Dijk (1998) model the cross-country distribution of per capita income as the mixture of a Weibull and a truncated normal density. The Weibull portion captures the left-hand mode and right skewness in the data while the truncated normal portion captures the right-hand mode. This combination is selected after testing the goodness of fit of various combinations of the normal density (truncated at zero), gamma, log normal and Weibull distributions; the data set that is employed measures levels of real GDP per capita for 120 countries for the time period 1960 and 1989. They find a bimodal fitted density in each year with “poor” and “rich” components corresponding to the Weibull and truncated normal densities respectively. The computed means of these

²⁹Bianchi’s data contains 119 countries observed at 3 distinct years, so each country is capable of making two crossings. The only crossings observed are 1) Trinidad and Tobago, which moves down between 1980 and 1989, 2) Venezuela, which moves down between 1970 and 1980, and 3) Hong Kong, which moves up between 1970 and 1980.

components diverge over the sample period and the weight given to the poor component in the mixture jumps in the mid-1970's from about .72 to about .82 implying that the mean gap between rich and poor countries grew and the poor increased in number. The attention to levels rather than log levels makes it hard to evaluate the welfare significance of this increased dispersion.

Recently, analyses of the distributions of income and growth have focused on identifying differences in these distributions across time and across subsets of countries. Anderson (2003) studies changes in the world income distribution by using nonparametric density function estimates combined with stochastic dominance arguments to compare the distributions at different points in time.³⁰ These methods allow him to construct measures of polarization of the income distribution; polarization is essentially characterized by shifts in probability density mass that increase disparities between relatively rich and relatively poor economies. Anderson finds that between 1970 and 1995 polarization between rich and poor countries increased throughout the time period. Maasoumi, Racine, and Stengos (2003) analyze the evolution of the cross-country distributions of realized, predicted, and residual growth rates; fitted growth rates and residuals are formed from nonparametric growth regressions using the Solow variables. These authors find that the distributions of growth rates for OECD and non-OECD countries are persistently different between 1965 and 1995, with the OECD distribution's variance reducing over time whereas the non-OECD distribution appears to be becoming less concentrated. One finds the same results for fitted growth rates; in contrast it is difficult to identify dimensions along which the distributions of OECD and non-OECD growth rate residuals differ. The major methodological difference between these papers relative to Paap and van Dijk (1998) is that these analyses do not rely on a mixture specification.

Distributional approaches suggest the utility of convergence measures that are based on the complete properties of probability measures characterizing output for different economies. Letting $\mu_i(x)$ and $\mu_j(x)$ denote the probability density functions

³⁰Anderson (2004) discusses issues related to the interpretation and econometric implementation of these methods.

for the variable of interest in economies i and j respectively, Anderson and Ge (2004) propose computing the convergence statistic $CI_{i,j}$

$$CI_{i,j} = \int_{-\infty}^{\infty} \min(\mu_i(x), \mu_j(x)) dx \quad (35)$$

This statistic is bounded between 0 and 1; a value of zero means that the density functions never assign positive probability to any common intervals or values of x whereas a value of 1 means that the densities coincide on all positive probability intervals or values. Anderson and Ge (2004) refer to the case $CI_{i,j} = 1$ as complete convergence. This statistic differs from the convergence measure described by eq. (21) as it evaluates differences between current densities and not asymptotic ones, but they are clearly closely related.

In our view, this approach will likely prove useful in a range of contexts. In particular, if one is interested in comparing income distributions between two economies, the Anderson-Ge statistic is a natural metric. In growth contexts, it is less clear whether the higher moments that distinguish (22) from (35) are of major concern, at least in the context of current debates.

d. distribution dynamics

In a series of papers, Quah (1993a,b,1996a,b,c,1997) has persuasively criticized standard regression approaches to studying convergence issues for being unable to shed light on important issues of mobility, stratification, and polarization in the world income distribution. Rather than studying the average behavior of a representative country, Quah proposes a schema, which he calls “distribution dynamics”, for studying the evolution of the entire cross-country income distribution. One way of implementing this approach is to assume that the process describing the evolution of the distribution is time-invariant and first-order Markov. Discretizing the state space then permits representation of cross-country income distribution as a probability mass function, λ_t , with an associated transition matrix, M . Each row of M is a probability mass function describing the

distribution over states of the system after one transition given that the system is currently in the state corresponding to that row. The evolution of the income distribution can then be described by $\lambda_t = M'\lambda_{t-1}$ so that $\lambda_{t+s} = (M^s)'\lambda_t$ is the s -step-ahead probability mass function and $\lambda_\infty = M'\lambda_\infty$ defines the long-run (ergodic) mass function (if it exists). Quah (1993b, 1996b) takes this approach and finds that the estimated M implies a bimodal (“twin-peaked”) ergodic mass function indicating a tendency towards polarization in the evolution of the world income distribution.³¹

Updating Quah's analysis using more recent data, Kremer, Onatski, and Stock (2001) also find evidence of twin-peaks in the long-run distribution of per capita incomes. However, they find the rich (right-hand) peak to be much larger than the poor (left-hand) peak unlike Quah, who found similarly sized peaks at both ends of the distribution. Kremer, Onatski, and Stock's point estimates imply that most countries will ultimately move to the rich state although, during the transition period, which could last hundreds of years, polarization in the income distribution may worsen. They are also unable to reject the hypothesis that there is a single right-hand peak in the long-run distribution. Quah (2001) responds to these claims by arguing that the imprecision in the estimates of the ergodic distributions is such that it is not possible to reject a wide range of null hypotheses including, by construction, that of twin-peakedness. Importantly, as Quah notes his work and that of others, including Kremer, Onatski, and Stock, is consistent with the view that the global poor are many in number and likely to be so for a very long time.

In addition, as Quah (1996c, 1997, 2001) and Bulli (2001) discuss, the process of discretizing the state space of a continuous variable is necessarily arbitrary and can alter the probabilistic properties of the data. Especially relevant here is the fact that the shape of the ergodic distribution can be altered by changing the discretization scheme. Reichlin (1999) demonstrates that the dynamic behavior inferred from the analysis of Markov transition probabilities, and the apparent long-run implications of that behavior, are

³¹As Quah (1993b, footnote 4) explains, the estimated ergodic distributions “... should *not* be read as forecasts of what will happen in the future...” (his emphasis). Rather, he continues, they “... should be interpreted simply as characterizations of tendencies in the post-War history that actually realized.”

sensitive to the discretization scheme employed; this work also shows that the estimated ergodic distribution can be sensitive to small changes in the transition probabilities. Bulli (2001) addresses this critique and shows how to discretize the state space in a way that preserves the probabilistic properties of the data. Applying her method to cross-country income data she finds an estimated ergodic distribution quite different from that found by arbitrary discretization as well as being an accurate approximation to the distribution computed using a continuous state space method.

An alternative formulation of distribution dynamics that avoids discretization problems is proposed by Quah (1996c,1997) and models the cross-country income distribution at time t with the density function, $f_t(x)$. If the process describing the evolution of the distribution is again assumed to be time-invariant and first-order Markov, then density at time $t + \tau$, $\tau > 0$, will be $f_{t+\tau}(x) = \int_0^\infty g_\tau(x|z)f_t(z)dz$ where $g_\tau(x|z)$ is the τ -period-ahead density of x conditional on z . The function $g_\tau(x|z)$ is the continuous analog of the transition matrix M and, assuming it exists, the ergodic (long-run) density function, $f_\infty(x)$, implied by $g_\tau(x|z)$ is the solution to $f_\infty(x) = \int_0^\infty g_\tau(x|z)f_\infty(z)dz$. Using nonparametric methods, Quah (1996c,1997) estimates various $g_\tau(x|z)$ and finds strong evidence of twin-peakedness in the cross-country income distribution. The estimated ergodic densities presented by Bulli (2001) and Johnson (2004) support Quah's conclusions.

Azariadis and Stachurski (2003) derive the form of the $g_\tau(x|z)$ implied by a stochastic version of the model in Azariadis and Drazen (1990). Estimation of the model's parameters enables them to compute forward projections of the sequence of cross-country income distributions, and ultimately the ergodic distribution, implied by the model. Consistent with the work of Quah (1996c, 1997) they find bimodality to be a pervasive feature of the sequence of distributions for about 100 years. Eventually, however, all countries transition to the rich mode so the ergodic distribution is unimodal as found by Kremer, Onatski, and Stock (2001). As Quah (2001) notes, there is "as yet" no theory of inference for this case so reconciliation of this result with the view that the

ergodic distribution is bimodal cannot be done through formal statistical tests. However, while Quah (2001) observes that such a theory is an “obvious next step,” he suggests that we may be close to the limits of what can be reasonably inferred from the cross-country income data.

Johnson (2000) offers an interpretation of $g_\tau(x|z)$ which draws an analogy between the median of the conditional distribution and the law of motion of a non-stochastic one-variable dynamic system. The median is the function $m(x)$ such that $\int_0^{m(x)} g_\tau(z|x) dz = .5$ so that a country with income of $m(x)$ at time t has an equal chance of having a higher or lower income at time $t+\tau$. Consider a point x_0 such that $m(x_0) = x_0$ and suppose that, in some neighborhood of x_0 , $m(x) > x$ for $x < x_0$ and $m(x) < x$ for $x > x_0$ implying $\Pr(x_{t+\tau} > x_t) > .5$ for $x < x_0$ and $\Pr(x_{t+\tau} < x_t) > .5$ for $x > x_0$ so that, in this neighborhood, countries with incomes different from x_0 tend to move toward x_0 . In the long run we may expect to find many countries in the vicinity of x_0 creating the tendency for a mode in the ergodic density, $f_\infty(x)$, at x_0 . Similarly, in a non-stochastic one-variable dynamic system with the law of motion $x_{t+\tau} = m(x_t)$, the condition on the phase diagram for the local stability of a steady-state at x_0 is that the graph of $m(x)$ intersects the 45° line from above at x_0 . In both cases, x_0 is a point of accumulation in the sense that the long-run probability of finding countries in the vicinity of x_0 will tend to be high relative to that elsewhere. Conversely, just as steady states are unstable in the non-stochastic case when $m(x)$ crosses the 45° line from below, analogous points in the stochastic case tend to produce antimodes in the ergodic density.

While Quah's estimated $g_\tau(x|z)$ indicate a strong tendency towards polarization in the world income distribution, they do not reveal much about intra-distribution mobility. Bimodality is arguably of less concern in a normative sense if there is movement between the two modes than it is if there is none. Quah (1996c) studies the mobility within the distribution by computing, (through stochastic simulation) the mean

time for a “growth miracle” which he defines as passage from the 10th to 90th percentile of the distribution. He finds an expected time of 201 years for such a miracle to occur.

Quah’s methods have subsequently been applied to a range of contexts. Andres and Lamo (1995) apply these methods to the OECD, Lamo (2000) to the regions of Spain, Johnson (2000) to US states, Bandyopadhyay (2002) to the Indian states, and Andrade, Laurini, Madalozzo, and Valls Pereira (2004) to Brazilian municipalities. These methods have also been extended to broader notions of distributional dynamics. Fiaschi and Lavezzi (2004) develop an analysis of the joint distribution of income levels and growth rates; their findings are compatible with the existence of multiple equilibria in the sense that countries may become trapped in the lower part of the income distribution.

e. relationship between distributional convergence and the persistence of initial conditions

Distributional methods have proven important in establishing stylized facts concerning the world income distribution. At the same time, there has been relatively little formal effort to explore the implications of findings such as twin peaks for the empirical salience of alternative growth theories. Some potential implications of distributional dynamics for evaluating theories are suggested by Quah (1996c), who finds that conditioning on measures of physical and human capital accumulation similar to those used by Mankiw, Romer, and Weil (1992) and a dummy variable for the African continent has little effect on the dynamics of the cross-country income distribution. The polarization and immobility features are similar in both cases and conditioning increases the expected time for a growth miracle to 760 years.³² These results suggest that the heterogeneity revealed by the distributional approaches is, at least in part, due to the existence of convergence clubs.

³²Other efforts to find determinants of polarization and immobility have produced mixed results. For the OECD countries, Andres and Lamo (1995) condition on the steady state implied by the Solow model and find little decrease in the tendency to polarization unless country specific effects are permitted. Lamo (2000) finds only a small increase in mobility for Spanish regions after conditioning on interregional migration flows. Bandyopadhyay (2002) shows that infrastructure spending and education measures appear to contribute to polarization between rich and poor states of India.

That being said, in general, it is relatively difficult to interpret properties of the cross-country income distribution in the context of economic convergence in the sense of (22). To see why this is so, it is useful to focus on the absence of a clear relationship between β -convergence, which measures the relative growth of rich versus poor countries and σ -convergence, which focuses explicitly on the distribution of countries. These two convergence notions do not have any necessary implications for one another, i.e. one may hold when the other does not. For our purposes, what is important is that σ -convergence is not an implication of β -convergence and so does not speak directly to the question of the transience of contemporary income differences. The erroneous assertion that β -convergence implies σ -convergence is known as Galton's fallacy and was introduced into the modern economic growth context by Friedman (1992) and Quah (1993a).

To understand the fallacy, suppose that log per capita output in each of N countries obeys the AR(1) process

$$\log y_{i,t} = \alpha + \zeta \log y_{i,t-1} + \varepsilon_{i,t} \quad (36)$$

where $0 < \zeta < 1$ and the random variables $\varepsilon_{i,t}$ are i.i.d across countries and time. For this model, each country will, by definition (22), exhibit convergence as any contemporaneous difference in output between two countries will disappear over time. Further, it is easy to see, using $\gamma_i = T^{-1}(\log y_{i,t+T} - \log y_{i,t})$, that the regression of growth on a constant and initial income will exhibit β -convergence. This is immediate when one considers growth between t and $t+1$ which means that growth obeys

$$\gamma_{i,t} = \alpha + (\zeta - 1) \log y_{i,t-1} + \varepsilon_{i,t} \quad (37)$$

where $\zeta - 1 < 0$ by assumption. In this model, by construction, the unconditional population variance of log output is constant because the reduction in cross-section variance associated with the tendency of high-income countries to grow more slowly than low-income countries is offset by the presence of the random shocks $\varepsilon_{i,t}$. This indicates

why σ -convergence is not a natural implication of long run independence from initial conditions; rather σ -convergence captures the evolution of the cross-section income distribution towards an invariant measure. This suggests that an important next step in the distributional approach to convergence is the development of tools which will allow distribution methods to more directly adjudicate substantive growth questions as they relate to the persistence of initial conditions.

iv. time series approaches to convergence

A final approach to convergence is based on time series methods. This approach is largely statistical in nature, which allows various hypotheses about convergence to be precisely defined, and thereby reveals appropriate strategies for formal testing. A disadvantage of the approach is that it is not explicitly tied to particular growth theories. Bernard and Durlauf (1995,1996), Evans (1998) and Hobijn and Franses (2000) provide a systematic framework for time series convergence tests.

Following Bernard and Durlauf (1995), a set of countries I is said to exhibit convergence if

$$\lim_{T \rightarrow \infty} \text{Proj}(\log y_{i,t+T} - \log y_{j,t+T} | F_t) = 0 \quad \forall i, j \in I \quad (38)$$

where $\text{Proj}(a|b)$ denotes the projection of a on b and F_t denotes some information set; operationally, this information set will typically contain various functions of time and current and lagged values of $\log y_{i,t}$ and $\log y_{j,t}$. Relative to our previous discussion, this definition represents a form of unconditional convergence that is closely related to (22). One can modify the definition to apply to the residual of per capita income after it has been projected on control variables such as savings rates in order to produce a definition of conditional convergence, but this has apparently not been done in the empirical literature.

In evaluating (38), researchers have generally focused on whether deterministic or stochastic trends are present in $\log y_{i,t} - \log y_{j,t}$; the presence of such trends immediately

implies a violation of (38). As such, time series tests of convergence have typically been implemented using unit root tests. One reason for this focus is that the presence of unit roots in $\log y_{i,t} - \log y_{j,t}$ allows for an extreme and therefore particularly interesting form of divergence between economies since a unit root implies that the difference $\log y_{i,t} - \log y_{j,t}$ will, with probability 1, become arbitrarily large at some point in the future.

The use of unit root and related time series tests has important implications for the sorts of countries that may be tested. Time series tests presuppose that $y_{i,t}$ may be thought of as generated by an invariant process in either levels or first differences, i.e., either levels or first differences may be modeled as the sum of deterministic terms plus a Wold representation for innovations. Such an assumption has significant economic content. As argued by Bernard and Durlauf (1996) countries that start far from their invariant distributions and are converging towards them, as occurs for countries that are in transition to the steady-state in the Solow-Swan model, will be associated with $\log y_{i,t} - \log y_{j,t}$ series that do not fulfill this requirement. Hence, tests of (38) can produce erroneous results if applied to such economies. To see this intuitively, suppose that for country i , $\log y_{i,t} = \log y_{i,t+1}$ for all t , so that country i has converged to a constant steady-state. Suppose that country j has the same steady-state as country i and is monotonically converging to this state so that $\log y_{i,t} > \log y_{j,t}$ for all observations. Then $\log y_{i,t} - \log y_{j,t} > 0$ for all t in the sample; which means that the series has a nonzero mean and tests that fail to account for the fact that the density of $\log y_{i,t} - \log y_{j,t}$ is changing across time can easily give erroneous inferences. For example one may use a test and conclude $\log y_{i,t} - \log y_{j,t}$ possesses a nonzero mean and erroneously interpret this as evidence against convergence, when the fact that the process does not have a time-invariant mean is ignored. This argument suggests that time series convergence tests are really only appropriate for advanced economies that may plausibly be thought of as characterized by invariant distributions.

Generally, the first generation of these tests rejected convergence for countries as well as other economic units. For example, Bernard and Durlauf (1995), studying 15

advanced industrialized economies between 1900 and 1989 based on data developed in Maddison (1982,1989), find little evidence that convergence is occurring; Hobijn and Franses (2000) similarly find little evidence of convergence across 112 countries taken from the Penn World Table for the period 1960-1989. The findings of nonconvergence in output levels are echoed in recent work by Pesaran (2004a) who employs convergence definitions that explicitly focus on the probability of large deviations, i.e. eq. (25). He finds little evidence of output level convergence using either the Maddison or Penn World Table data.

Relatively little explicit attention has been paid to the question of systematically identifying convergence clubs using time series methods. One exception is Hobijn and Franses (2000) who employ a clustering algorithm to identify groups of converging countries.³³ Their algorithm finds many small clusters in their sample of 112 countries – depending on the particular rule used to determine cluster membership, they find 42 or 63 clusters with most containing just two or three countries. Hobijn and Franses view these clusters as convergence clubs but it is not clear that they represent groups of countries in distinct basins of attraction of the growth process. Absent controls for structural characteristics, these groupings could simply reflect the pattern of differences in those characteristics rather than differences in long-run outcomes due to differences in initial conditions. Moreover, the Bernard and Durlauf (1996) argument about the substantive economic assumptions that underlie time series methods for studying convergence seems applicable here. Given the breadth of the sample used by Hobijn and Franses, it is unlikely that it contains only data generated by countries whose behavior is near their respective steady-states; such an assumption is much more plausible for restricted samples such as the OECD countries. The clusters they find could thus reflect, in many cases at least, transition dynamics rather than convergence clubs. An important extension of this work would be the exploration of how one can distinguish convergence clubs from

³³Corrado, Martin, and Weeks (2004) extend this approach to allow for time variation in clusters. They conclude that there is substantial evidence of club convergence as opposed to overall convergence for European regions. A nice feature of their analysis is the effort to interpret the clubs that are identified statistically with alternative economic theories, and conclude that geographic proximity and demographic similarity correlate with their observed clusters.

what may be called “transition” clubs, i.e. groups of countries exhibiting similar transition dynamics.

A number of studies of time series convergence have criticized these claims of nonconvergence; these criticisms are based upon inferential issues that have arisen in the general unit roots literature. One of these issues concerns the validity of unit root tests in the presence of structural breaks in $\log y_{i,t} - \log y_{j,t}$; as argued initially by Perron (1989), the failure to allow for structural breaks when testing for unit roots can lead to spurious evidence in support of the null hypothesis that a unit root is present. An initial analysis of this type in cross-country contexts is Greasley and Oxley (1997) who, imposing breaks exogenously, find convergence for Denmark and Sweden whereas the sort of test employed by Bernard and Durlauf (1995) does not. The role of breaks in time series convergence tests is systematically studied in Li and Papell (1999). An important feature of their analysis is that Li and Papell avoid exogenous imposition of trend breaks and in fact find that the dates of these breaks exhibit some heterogeneity, although many of them cluster around World War II. Li and Papell find that the evidence for OECD convergence is more mixed than did Bernard and Durlauf (1995) in the sense that allowing for trend breaks reduces the number of country pairs that fail to exhibit convergence. Related findings are due to Carlino and Mills (1993) who study US regions and reject convergence except under specifications that allow for a trend break in 1946. These conclusions are shown by Loewy and Papell (1996) to hold even if one allows potential trend breaks to be endogenously determined by the data.

While the analysis of trend breaks and convergence tests is valuable because of its implications about the time series structure of output differences between countries, studies of this type suffer from some interpretation problems. The presence of the regime break is presumably suggestive of an absence of convergence in the sense of (22) or (38), since it implies that there is some component of $\log y_{i,t} - \log y_{j,t}$ that will not disappear over a sufficiently long time horizon. The time series definition of convergence is violated by any long-term predictability in output differences. Hence, claims by authors that allowing for data breaks produces evidence of convergence begs the question of what is meant by convergence. That being said, the sort of violation of (22) or (38) implied by a trend break is different from the type implied by a unit root. In particular, a break

associated with the level of output means that the output difference between two countries is always bounded, unlike the unit root case.

A distinct line of criticism of time series convergence tests is due to Michelacci and Zaffaroni (2000) who argue that convergence tests based on the presence of unit roots may perform badly when the true processes exhibit long memory. Let $\gamma(L)u_{i,j,t}$ denote the moving average representation for $\log y_{i,t} - \log y_{j,t}$. Suppose that the k 'th coefficient in the representation has the property that

$$\gamma_k \propto k^{d-1}, \quad 0 < d < 1 \quad (39)$$

In this case, shocks die out at a hyperbolic rather than geometric rate, which is one definition of long memory in a time series process. Michelacci and Zaffaroni (2000) show that if output deviations exhibit long memory, one can reconcile the claim of β -convergence with time series evidence of divergence, i.e. the failure of various tests to reject the presence of a unit root in per capita output deviations. This is a potentially important reconciliation of these two distinct testing strategies.

That being said, the plausibility of a long memory characterization has yet to be established in the economics literature. One problem is that there is an absence of a body of economic theory that predicts the presence of long memory.³⁴ The existing theoretical justifications of long memory processes derive from aggregation arguments originating with Granger (1980); the conditions under which aggregation produces long memory do not have any particular empirical justification. In addition, there are questions concerning the ability of conventional statistical methods to allow one to distinguish between long memory models and various alternatives. Diebold and Inoue (2001) indicate how long memory may be spuriously inferred for series subject to regime shifts, so the strength of evidence of long memory cited by Michelacci and Zaffaroni (2000)

³⁴There are at least two reasons why unit roots stem naturally from existing economic theories. First, technology shocks are generally modeled as permanent. Second, Euler equations often produce unit root or near unit root like conditions. The random walk theory of stock prices is one example of this, in which risk neutral agents produce unpredictability of stock price changes as an equilibrium.

may be questioned. Nevertheless, the Michelacci-Zaffaroni argument is important, not least because it focuses attention on the role in growth empirics of size and power issues that arise in all unit root contexts.

Time series approaches to convergence are melded with analysis related to σ -convergence in Evans (1996) who considers the cross-section variance of growth rates at time t ,

$$\sigma_t^2 = \frac{1}{N} \sum_i (\log y_{i,t} - \overline{\log y_t})^2 \quad (40)$$

where $\overline{\log y_t} = \frac{1}{N} \sum_i \log y_{i,t}$ and N is the cardinality of I . Evans observes that σ_t^2 may be represented as a unit root process with a quadratic time trend when there is no cointegration among the series $\log y_{i,t}$. This leads Evans to suggest a time-series test of convergence based on unit root tests applied to σ_t^2 . Employing this test, Evans concludes that there is convergence to a common trend among 13 industrial countries. One interpretation problem with this analysis is that it allows different countries to possess different deterministic trends in per capita output albeit with the same trend growth rate. Such differences are obviously germane with respect to convergence as an economic concept being consistent, for example, with the club and conditional convergence hypotheses but not with the unconditional convergence hypothesis. Evans (1997) provides a time series approach to estimating rates of convergence. He shows that OLS applied to equation (18) yields a consistent estimator of β , and hence the rate of convergence, only if (i) each $\log y_{i,t} - \overline{\log y_t}$ obeys an AR(1) process having the same AR(1) parameter lying strictly between 0 and 1; and, (ii) the control variables, X_i and Z_i , account for all cross-country heterogeneity. He argues that neither condition is likely to hold and offers an alternative method of measuring the rate of convergence based on the supposition that $\log y_{i,t} - \overline{\log y_t}$ follows an AR(q) process with lag polynomial $\Lambda(L)$. Again, this specification allows countries to follow different parallel balanced

growth paths and Evans defines the rate of convergence for economy i as the rate at which $\log y_{i,t}$ “is expected to revert toward its balanced growth path far in the future.” He shows that, given that it is a real, distinct, positive fraction, the dominant root of the polynomial $z^q \Lambda(z^{-1})$ equals one minus this rate. Evans computes estimates of the convergence rates and their 90% confidence intervals for a sample of 48 countries over the period 1950-90 and for the contiguous US states over the period 1929-91. For the states, about a third of the point estimates are negative and about two-thirds of the confidence intervals contain zero, while for the countries, about half of the point estimates are negative and all but two of the confidence intervals contain zero. However, in spite of these positive estimated average convergence rates of 15.5% and 5.9% respectively, Evans' analysis fails to yield persuasive evidence in favor of the conditional convergence hypothesis since, in most cases, the hypothesis of a convergence rate of zero cannot be rejected at the 10% level of significance.

Later sections of the chapter will discuss how growth researchers can draw on time series data in other ways. One popular route has been to use panel data, with repeated observations on each country or region. Another method is to use techniques broadly similar to those of event studies in empirical finance, and trace out the consequences of specific events, such as major political or economic reforms. We will consider these approaches in Section VI.iii below.

v. sources of convergence or divergence

Abramowitz (1986), Baumol (1986), DeLong (1988) and many others, both before and since, view convergence as the process of follower countries “catching up” to leader countries by adopting their technologies. Some more recent contributors, such as Barro (1991) and Mankiw, Romer, and Weil (1992), adopt the view that convergence is driven by diminishing returns to factors of production.³⁵ In the neoclassical model, if

³⁵ When an economy is below its steady-state value of capital per efficiency unit of labor, the marginal product of capital is relatively high (and is higher than in the steady state). As a result, a given investment rate translates into relatively high output growth. Capital grows as well but, because of diminishing returns, the capital-output ratio rises and the marginal product of capital declines, causing the growth of output and capital to slow.

each country has access to the same aggregate production function the steady-state is independent of an economy's initial capital and labor stocks and hence initial income. In this model, long-run differences in output reflect differences in the determinants of accumulation, not differences in the technology used to combine inputs to produce output. Mankiw (1995, p. 301), for example, argues that for “understanding international experience, the best assumption may be that all countries have access to the same pool of knowledge, but differ by the degree to which they take advantage of this knowledge by investing in physical and human capital.” Even if one relaxes the assumption that countries have access to the same production function, convergence in growth rates can still occur so long as each country’s production function is concave in capital per efficiency unit of labor and each country experiences the same rate of labor-augmenting technical change.

Klenow and Rodríguez-Clare (1997a) challenge this “neoclassical revival” with results suggesting that differences in factor accumulation are, at best, no more important than differences in productivity in explaining the cross-country distribution of output per capita. They find that only about half of the cross-country variation in the 1985 level of output per worker is due to variation in human and physical capital inputs while a mere 10% or so of the variation in growth rates from 1960 to 1985 reflects differences in the growth of these inputs. The differences between the results of Mankiw, Romer, and Weil (1992) and the findings of Klenow and Rodríguez-Clare (1997a) in their reexamination of Mankiw, Romer and Weil have two principal origins. First, citing concerns about the endogeneity of the input quantities, Klenow and Rodríguez-Clare (1997a) eschew estimation of the capital shares and choose to impute parameters based on the results of other studies. Second, they modify Mankiw, Romer and Weil's measure of human capital accumulation by supplementing secondary school enrollment rates using data on primary enrollment. This yields a measure of human capital accumulation with less cross-country variation than that used by Mankiw, Romer and Weil. This one modification decreases

Eventually, the economy converges to a steady state in which capital and output grow at the same rate and the marginal product of capital is sustained at a constant level by labor-augmenting technical progress. Dowrick and Rogers (2002) find that both diminishing returns and technology transfer are important contributors to the convergence process. See also Bernard and Jones (1996) and Barro and Sala-i-Martin (1997).

the relative contribution of cross-country variation in human and physical capital inputs to variation in the 1985 level of output per worker to 40% from the 78% found by Mankiw, Romer, and Weil. Prescott (1998) and Hall and Jones (1999) confirm the view that differences in inputs are unable to explain observed differences in output and Easterly and Levine (2001, p. 177) state that “[t]he 'residual' (total factor productivity, TFP) rather than factor accumulation accounts for most of the income and growth differences across countries.”

Unlike many authors, who estimate TFP as a residual after assuming a common Cobb-Douglas production function, Henderson and Russell (2004) use a non-parametric production frontier approach (data envelopment analysis) to decompose the 1965 to 1990 growth of labor productivity into (i) shifts in the (common, worldwide) production frontier (technological change); (ii) movements toward (or away from) the frontier (technological catch-up); and, (iii) capital accumulation. They find a dominant role for capital accumulation in the growth of the cross-country mean of labor productivity with human and physical capital each accounting for about half of that role.³⁶ They also observe that the distribution of labor productivity became more dispersed from 1965 to 1990 and their results suggest that physical and human capital accumulation were largely responsible for the increased dispersion.

The results of Henderson and Russell (2004) and those of the previous authors are, however, more consistent than it may seem. Klenow and Rodríguez-Clare (1997a), Hall and Jones (1999) and Barro and Sala-i-Martin (2004) argue that the standard growth accounting decomposition overstates the contribution of capital accumulation to output growth by attributing to capital the effect on output of increases in capital induced by increases in TFP. This effect also applies to Henderson and Russell's approach and adjusting for it provides some reconciliation of their findings with those of Klenow and Rodríguez-Clare (1997a), Prescott (1998) and Hall and Jones (1999). The standard growth accounting formula attributes a fraction (equal to labor's share of output) of the

³⁶Note that any misspecification of the production function due to the Cobb-Douglas assumption in other studies will tend to increase the apparent variation in TFP relative to that found by Henderson and Russell (2004) under the weaker assumption of constant returns to scale. In a rare effort to evaluate the Cobb-Douglas specification, Duffy and Papageorgiou (2000) reject it in favor of a more general CES functional form.

growth in output per worker to growth in TFP and a fraction (equal to capital's share of output) to capital accumulation despite the fact that, in the steady-state, growth in output per worker is entirely due to technological progress (Barro and Sala-i-Martin (2004, p. 457-60) and Klenow and Rodríguez-Clare (1997a p. 75, fn. 4)). The total effect of technological progress on output growth can thus be estimated by dividing labor's share into the estimated growth rate of TFP. Interpreting "capital" broadly, labor's share is about 1/3 suggesting that this effect is about three times the rate of growth of TFP. Henderson and Russell (2004, Table 5, row (a)) find that, on average, about 90% of the increase in output per worker over the 1965 to 1990 period is attributable to the accumulation of human and physical capital with increases in TFP accounting for the remaining 10%. Applying the adjustment discussed above suggests that technological progress accounts for about 30% of the growth in output per worker over this period while capital accumulation, due to transition dynamics, accounts for the remainder.

As well as determining the relative contributions of inputs and TFP to the cross-country variation in output and output growth, some have studied what features of the cross-country output distribution are explained by the cross-country distributions of inputs and TFP. Henderson and Russell (2004) document the emergence of a second mode in the cross-country distribution of output per worker between 1965 and 1990 and find changes in efficiency (the distance from the world technological frontier) to be largely responsible. A primary role for TFP in determining the shape of the long-run distribution of output per capita is found by Feyrer (2003) who uses Markov transition matrices estimated with data from 90 countries over the period 1970 to 1989 to estimate the ergodic distributions of output per capita, the capital-output ratio, human capital per worker, and TFP. He finds that the long-run distributions of both output per capita and TFP are bimodal while those of both the capital-output ratio and human capital per worker are unimodal. This result, Feyrer observes, has potentially important implications for theoretical modelling of development traps. It suggests that models of multiple equilibria that give rise to equilibrium differences in TFP are more promising than

models that emphasize indeterminacy in capital intensity or educational attainment.³⁷ It is also consistent with Quah's (1996c) finding that conditioning on measures of physical and human capital accumulation (and a dummy variable for the African continent) has little effect on the dynamics of the cross-country income distribution.

As discussed in Section III.iii.d, the shapes of ergodic distributions computed from transition matrices estimated with discretized data are not, in general, robust to changes in the way in which the state space is discretized. To avoid these problems, Johnson (2004) extends Feyrer's analysis using Quah's (1996c,1997) continuous state-space methods and finds evidence of bimodality in the long-run distributions of both the capital-output ratio and TFP in addition to that in the long-run distribution of output per capita. This finding is broadly consistent with data produced by a version of the Solow growth model that includes a threshold externality à la Azariadis and Drazen (1990) but may be partly due to the computation of TFP after supposing a Cobb-Douglas production function across countries. Accordingly, some care must be exercised when drawing conclusions from these results.

More generally, in much of the development accounting literature cited above, TFP is measured as a residual under the assumption of a concave worldwide production function. Durlauf and Johnson (1995) present evidence contrary to that assumption and in support of the implied multiple steady states in the growth process. It seems likely that the imposition of a concave production function in this case will tend to exaggerate the measured differences in TFP and so confound inferences about the importance of TFP variation.³⁸ While Henderson and Russell (2004)'s approach is nonparametric and free from any assumption of a particular technology *per se*, it estimates the world technology frontier by fitting a convex cone to data on outputs and inputs. The imposed convexity of the production set prevents the method from discovering any nonconvexities that may exist and, in addition to masking the presence of multiple steady states, convexifying

³⁷Romer (1993) discusses the intellectual origins of the centrality of capital accumulation in models of economic development and argues that “idea gaps are central to the process of economic development” (p. 548).

³⁸Graham and Temple (2003) show that the existence of multiple steady states can increase the variance and accentuate bimodality in the observed cross-country distribution of TFP.

these nonconvexities would tend to overstate the cross-country variation in TFP. The extent to which our current understanding of the relative contributions of variation in inputs and variation in TFP to the observed variation in income levels is influenced by the effects on measured TFP of a misspecified worldwide technology remains an open research question.

Despite these concerns and the differences in the precise estimates found by different researchers, it is clear that cross-country variation in inputs falls short of explaining the observed cross-country variation in output. The result that the TFP residual, a “measure of our ignorance” computed as the ratio of output to some index of inputs, is an important (perhaps the dominant) source of cross-country differences in long-run economic performance is useful but hardly satisfying and the need for a theory of TFP expressed by Prescott (1998) is well founded. Research such as Acemoglu and Zilibotti (2001) and Caselli and Coleman (2003) are promising contributions to that agenda.

V. Statistical models of the growth process

While the convergence hypothesis plays a uniquely prominent role in empirical growth studies, it by no means represents the bulk of empirical growth studies. The primary focus of empirical growth papers may be thought of as a general exploration of potential growth determinants. This work may be divided into three main categories: 1) studies designed to establish that a given variable does or does not help explain cross-country growth differences, 2) efforts to uncover heterogeneity in growth and 3) studies that attempt to uncover nonlinearities in the growth process. While analyses of these types are typically motivated by formal theories, operationally they represent efforts to develop statistical growth models that are consistent with certain types of specification tests.

Section V.i discusses the analysis of how specific determinants affect growth. We describe the range of different variables that have appeared in growth regressions and consider alternative methodologies for analyzing growth models in the presence of

uncertainty about which regressors should be included to define the “true” growth model. Section V.ii addresses issues of parameter heterogeneity. The complexity of the growth process and the plethora of new growth theories suggest that the mapping of a given variable to growth is likely a function of both observed and unobserved factors; for example, the effect of human capital investment on growth may depend on the strength of property rights. We explore methods to account for parameter heterogeneity and consider the evidence that has been adduced in support of its presence. Section V.iii focuses on the analysis of nonlinearities and multiple regimes in the growth process. Endogenous growth theories are often highly nonlinear and can produce multiple steady states in the growth process, both of which have important implications for econometric practice. This subsection explores alternative specifications that have been employed to allow for nonlinearity and multiple regimes and describes some of the main findings that have appeared to date.

i. specifying explanatory variables in growth regressions

In the search for a satisfactory statistical model of growth, the main area of effort has concerned the identification of appropriate variables to include in linear growth regressions, this generally amounts to the specification of Z in equation (18). Appendix 2 provides a survey of different regressors that have been proposed in the growth literature with associated studies that either represent the first use of the variable or a well known use of the variable.³⁹ The table contains 145 different regressors, the vast majority of which have been found to be statistically significant using conventional standards.⁴⁰ One reason why so many alternative growth variables have been identified is due to questions of measurement. For example, a claim that domestic freedom affects growth leaves unanswered how freedom is to be measured. We have therefore organized the body of growth regressors into 43 distinct growth “theories” (by which we mean conceptually

³⁹Our choices of which studies to include should not be taken to reflect any stance on any cases where there is disagreement about priority as to who first proposed a variable.

⁴⁰Of course, the high percentage of statistically significant growth variables reflects publication bias as well as data mining.

distinct growth determinants); each of these theories is found to be statistically significant in at least one study.

As Appendix 2 indicates, the number of growth regressors that have been identified approaches the number of countries available in even the broadest samples. And this regressor list does not consider cases where interactions between variables or nonlinear transformations of variables have been included as regressors; both of which are standard ways of introducing nonlinearities into a baseline growth regression. This plethora of potential regressors starkly illustrates one of the fundamental problems with empirical growth research, namely, the absence of any consensus on which growth determinants ought to be included in a growth model. In this section, we discuss efforts to address the question of variable choice in growth models.

To make this discussion concrete, define S_i as the set of regressors which a researcher always retains in a regression and let R_i denote additional controls in the regression, so that

$$\gamma_i = \psi S_i + \pi R_i + \varepsilon_i \quad (41)$$

Notice that the inclusion of a variable in S does not mean the researcher is certain that it influences growth, only that that it will be included in all models under consideration. To make this concrete, one can think of an exercise in which one wants to consider the relationship between initial income and growth. A researcher may choose to include initial income and the other Solow growth regressors in every specification of the model, but may in contrast be interested in the effects of different non-Solow growth regressors on inferences about the initial income/growth connection.

If one takes the regressors that comprise R as fixed, then statements about elements of ψ are straightforward. A frequentist approach to inference will compute an estimate of the parameter $\hat{\psi}$ with an associated distribution that depends on the data generating process; Bayesian approaches will compute a posterior probability density of ψ given the researcher's prior, the data, and the assumption that the linear model is correctly specified, i.e. the choice of variables in R corresponds to the "true" model. Designating

the available data as D and a particular model as m , this posterior may be written as $\mu(\psi|D,m)$.

The basic problem in developing statistical statements either about $\hat{\psi}$ or $\mu(\psi|D,m)$ is that there do not exist good theoretical reasons to specify a particular model m . This is *not* to say that the body of growth theories may not be used to identify candidates for R . Rather, the problem is that growth theories are, using a phrase due to Brock and Durlauf (2001a), openended. Theory openendedness means that the growth theories are typically compatible with one another. For example, a theory that institutions matter for economic growth is not logically inconsistent with a theory that emphasizes the role of geography in growth. Hence, if one has a set of K potential growth theories, all of which are logically compatible with one another (and all subsets of theories), there exist $2^K - 1$ potential theoretical specifications of the form (41), each one of which corresponds to a particular combination of theories.

One approach to resolving the problem of model uncertainty is based on identifying variables whose empirical importance is robust across different model specifications. This is the idea behind Levine and Renelt's (1992) use of extreme bounds analysis (Leamer (1983) and Leamer and Leonard (1983)) to assess growth determinants. To see how extreme bounds analysis may be applied to the assessment of robustness of growth determinants, suppose that one has specified a space of possible models M . For model m , the growth process is

$$\gamma_i = \psi_m S_i + \pi_m R_{i,m} + \varepsilon_{i,m} \quad (42)$$

where the subscripts m reflect the model specific nature of the parameters and associated residuals. One can compute $\hat{\psi}_m$ for every model in M . Motivated by Leamer (1983), Levine and Renelt employ the rule that there is strong evidence that a given regressor in S , call it s_i , robustly affects growth if the sign of the associated regression coefficient $\hat{\psi}_{i,m}$ is constant and the coefficient estimate is statistically significant across all model

specifications in M . In this analysis the S vector is composed of a variable of interest and other variables whose presence is held fixed across specifications.

In the Levine and Renelt (1992) analysis, S includes a constant, the initial income, the investment share of GDP, secondary school enrollment rates, and population growth; these variables proxy for those suggested by the Solow model. Models are distinguished by alternative combinations of 1 to 3 variables taken from a set of 7 variables; these correspond to alternative choices of $R_{i,m}$. Based on the constant sign and statistical significance criteria, Levine and Renelt (1992) conclude that the only robust growth determinants among the elements of S_i are initial income and the share of investment in GDP. These two findings are confirmed in subsequent work by Kalaitzidakis, Mamuneas, and Stengos (2000) who allow for potential nonlinearities in (41). Specifically, they consider partially linear versions of (41), so that

$$\gamma_i = \psi_m S_i + f_m(\pi R_{i,m}) + \varepsilon_{i,m} \quad (43)$$

Note that the function $f(\cdot)$ is allowed to vary across specifications of R . As in Levine and Renelt (1992), Kalaitzidakis, Mamuneas, and Stengos conclude that initial income and physical capital investment rates are robust determinants of growth. Unlike Levine and Renelt, they also find that inflation volatility and exchange rate distortions are robust; this is interesting as it is an example where the failure to account for nonlinearity in one set of variables masks the importance of another.

From a decision-theoretic perspective, the extreme bounds approach is a problematic methodology. The basic difficulty, discussed in detail in Brock and Durlauf (2001a) and Brock, Durlauf, and West (2003) is that if one is interested in ψ_l because one is considering whether to change $s_{i,l}$, by one unit, i.e. one is advising country i on a policy change, the extreme bounds standard corresponds to a very risk averse way of responding to model uncertainty. Specifically, suppose that for a policymaker, $El(s_{i,l}, m)$ represents the expected loss associated with the current policy level in country i . We assume that one is only interested in the case where an increase in the policy raises

growth, which means we will assume that it is necessary for $\hat{\psi}_{l,m} > 0$ in order to conclude that one should make the change. One can approximate the t -statistic rule, i.e. requiring that the coefficient estimate for s_l be statistically significant in order to justify a policy as implying that

$$El(s_{i,l} + 1, m) - El(s_{i,l}, m) = (\hat{\psi}_{l,m} - 2sd(\hat{\psi}_{l,m})) > 0 \quad (44)$$

where $sd(\hat{\psi}_{l,m})$ is the estimate of the standard deviation associated with $\hat{\psi}_{l,m}$ and the statistical significance level required for the coefficient is assumed to correspond to a t -statistic of 2. This loss function may look odd, but it is in fact the sort of loss function implicitly assumed whenever one relies on t -statistics to make policy decisions. Extreme bounds analysis requires that (44) holds for every model in M . This requires that $El(s_{i,l})$, the expected loss for a policymaker when one conditions only on the policy variable, has the property that

$$El(s_{i,l} + 1) - El(s_{i,l}) > 0 \Rightarrow El(s_{i,l} + 1, m) - El(s_{i,l}, m) > 0 \quad \forall m \quad (45)$$

This means that the policymaker must have minimax preferences with respect to model uncertainty, i.e. he will make the policy change only if it yields a positive expected payoff under the least favorable model in the model space. While there are reasons to believe that in practice, individuals assess model uncertainty differently than within-model uncertainty⁴¹, the extreme risk aversion embedded in (45) seems hard to justify.

Even when one moves away from decision-theoretic considerations, extreme bounds analysis is somewhat difficult to interpret as a statistical procedure. Hoover and Perez (2004), for example, show that the use of extreme bounds analysis can lead to the conclusion that many growth determinants are fragile even when they are part of the data

⁴¹See discussion in Brock, Durlauf, and West (2003) of the Ellsberg Paradox.

generating process. They also find that the procedure has poor power properties in the sense that some regressors that do not matter may spuriously appear to be robust.⁴²

The concern that extreme bounds analysis represents an excessively conservative approach to evaluating empirical results led Sala-i-Martin (1997a,b) to propose a different way to evaluate the robustness of findings. Within a model, suppose there is an evaluative criterion for $\hat{\psi}_m$ that is used to determine whether the variable s_i matters for the growth process. One example of such a standard is whether or not $\hat{\psi}_{i,m}$ is statistically significant at some level. Sala-i-Martin first proposes averaging the statistical significance levels via

$$\hat{S}_i = \sum_m \hat{\omega}_m \hat{S}_{i,m} \quad (46)$$

where $\hat{S}_{i,m}$ is the statistical significance level associated with $\hat{\psi}_m$ and $\hat{\omega}_m$ is the weight assigned to model m , $\sum_m \hat{\omega}_m = 1$. Sala-i-Martin (1997a,b) employs weights determined by the likelihoods of each model as well as employing equal weighting. Second, Sala-i-Martin (1997a,b) proposes examining the percentage of times a variable appears statistically significant with a given sign; a variable whose sign and statistical significance holds across 95% of the different models estimated is regarded as robust. This approach finds that initial income, the investment to GDP ratio and secondary school education are all robust determinants of growth. Sala-i-Martin (1997a,b) extends this analysis to the evaluation of additional variables and finds a number also are robust by his criteria.

While these approaches have the important advantage over extreme bounds analysis of accounting for the informational content of the entire distribution of $\hat{\psi}_m$, the procedures do not have any decision-theoretic or conventional statistical justification. We are unaware of any statistical interpretation to averaged significance levels. Further, little is understood about the statistical properties of these procedures. Hoover and Perez

⁴²For further discussion of extreme bounds analysis, see Temple (2000b) and the references therein.

(2004), for example, find that the second Sala-i-Martin procedure has poor size properties, in the sense that “true” growth determinants are still likely to fail to be identified.

Dissatisfaction with extreme bounds analysis and the variants we have described have led some authors to embed the determinants of robust growth regressors in a general model selection context. Hendry and Krolzig (2004) and Hoover and Perez (2004) both employ general-to-specific modeling methodologies generally associated with the research program of David Hendry (cf. Hendry 1995) to select one version of (41) out of the model space. In both papers, the linear model that is selected out of the space of possible models is one where growth is determined by years an economy is open, the rate of equipment investment, a measure of political instability based on the number of coups and revolutions, a measure of the percentage of the population that is Confucian and a measure of the percentage of the population that is Protestant.

Methodologically, these papers in essence employ algorithms which choose a particular regression model from a space of models through comparisons based on a set of statistical tests. The extent to which one finds this approach appealing is a function of the extent to which one is sympathetic to the general methodological foundations of the Hendry research program; we avoid such an extended evaluation here, but simply note that like other general prescriptions the program remains controversial, especially the extent to which it relies on automatic model selection procedures that do not possess a clear decision-theoretic justification. As such, it is somewhat unclear how to evaluate the output of the procedure in terms of the objectives of a researcher. That being said, the automated procedures Hendry works with have the important virtue that they can facilitate identifying small sets of models that are well supported by available data. Identification of such models is important, for example, in forecasting, where Hendry’s procedures appear to have a strong track record.

In our judgment, the most promising current approach to accounting for model uncertainty employs model averaging techniques to construct parameter estimates that formally address the dependence of model-specific estimates on a given model. Examples where model averaging has been applied to cross-country growth data include Brock and Durlauf (2001a), Brock, Durlauf, and West (2003), Doppelhofer, Miller, and

Sala-i-Martin (2004), Fernandez, Ley, and Steel (2001a) and Masanjala and Papageorgiou (2004). The basic idea in this work is to treat the “true” growth model⁴³ as an unobservable variable. In order to account for this variable, each element m in the model space M is associated with a posterior model probability $\mu(m|D)$. By Bayes’ rule,

$$\mu(m|D) \propto \mu(D|m)\mu(m) \quad (47)$$

where $\mu(D|m)$ is the likelihood of the data given the model and $\mu(m)$ is the prior model probability. These model probabilities are used to eliminate the dependence of parameter analysis on a specific model. For frequentist estimates, averaging is done across the model-specific estimates $\hat{\psi}_m$ to produce an estimate $\hat{\psi}$ via

$$\hat{\psi} = \sum_m \hat{\psi}_m \mu(m|D) \quad (48)$$

whereas for the Bayesian context, the dependence of the posterior probability measure of the parameter of interest, $\mu(\psi|D, m)$ on the model choice is eliminated via standard conditional probability arguments, i.e.

$$\mu(\psi|D) = \sum_{m \in M} \mu(\psi|D, m)\mu(m|D) \quad (49)$$

Brock, Durlauf, and West (2003) argue that the strategy of constructing posterior probabilities that are not model-dependent is the appropriate one when the objective of the statistical exercise is to evaluate alternative policy questions such as whether to

⁴³In this discussion, we will assume that one of the models in the model space M is the correct specification of the growth process. When none of the model specifications is the correct one, this naturally affects the interpretation of the model averaging procedure.

change elements of S_i by one unit. Notice that this approach assumes that the goal of the exercise is to study a parameter, i.e. ψ , not to identify the best growth model.

Model averaging approaches are still quite new in the growth literature, so many questions exist as to how to implement the procedure. One issue concerns the specification of priors on parameters within a model. Brock and Durlauf (2001a), Brock, Durlauf and West (2003), and Doppelhofer, Miller, and Sala-i-Martin (2004) assume a diffuse prior on the model specific coefficients. The advantage of this prior is that, when the errors are normal with known variance, the posterior expected value of ψ , conditional on the data D and model m , is the ordinary least squares estimator $\hat{\psi}_m$. The disadvantage of this approach is that since the diffuse prior on the regression parameters is improper, one has to be careful that the posterior model probabilities associated with the prior are interpretable. For this reason, Doppelhofer, Miller, and Sala-i-Martin (2004) eschew reference to their methodology as strictly Bayesian. That being said, so long as the posterior model probabilities include appropriate penalties for model complexity, (and Brock and Durlauf (2001a), Brock, Durlauf, and West (2003), and Doppelhofer, Miller, and Sala-i-Martin (2004) all compute posterior model probabilities using BIC adjusted likelihoods) we do not see any conceptual problem in interpreting this approach as strictly Bayesian. Fernandez, Ley, and Steel (2001a) and Masanjala and Papageorgiou (2004) employ proper priors and therefore avoid such concerns.⁴⁴ We are unaware of any evidence that the choice of prior for the within-model regression coefficients is of great importance in terms of empirical inferences for the growth contexts that have been studied; Masanjala and Papageorgiou (2004) in fact compare results using using proper priors with the improper priors we have described and find that the choice of prior is unimportant.

A second unresolved issue concerns the specification of the prior model probabilities $\mu(m)$. In the model averaging literature, the general assumption has been to assign equal prior probabilities to all models in M . This prior may be interpreted as assuming that the prior probability that a given variable appears in the “true” model is .5

⁴⁴Fernandez, Ley, and Steel (2001b) provide a general analysis of proper model specific priors for model averaging exercises.

and that the probability that one variable appears in the model is independent of whether others appear. Doppelhofer, Miller, and Sala-i-Martin (2004) consider modifications of this prior in which the probability that a given variable appears in the true model is $p < .5$; these alternative probabilities are chosen in order to assign greater weight to more parsimonious growth models, i.e. models in which fewer regressors appear.⁴⁵

Brock and Durlauf (2001a) and Brock, Durlauf, and West (2003) argue against the assumption that the probability that one regressor should appear in a growth model is independent of the inclusion of others. The basic problem with priors that assume independence is analogous to the red bus/blue bus problem in discrete choice theory; namely, some regressors are quite similar to others, e.g. alternative measures of trade openness, whereas other regressors are quite disparate, e.g. geography and institutions. Brock, Durlauf, and West (2003) propose a tree structure to organize model uncertainty for linear growth models. First, they argue that growth models suffer from theory uncertainty. Hence, one can identify alternative classes of models based on what growth theories are included. Second, for each specification of a body of theories to be embedded, they argue there is specification uncertainty. A given set of theories requires determining whether the theories interact, whether they are subject to threshold effects or other types of nonlinearity, etc. Third, for each theory and model specification, there is measurement uncertainty. The statement that weather affects growth does not specify the relevant empirical proxies, e.g. the number of sunny days, average temperature, etc. Finally, each choice of theory, specification and measurement is argued to suffer from heterogeneity uncertainty, which means that it is unclear which subsets of countries obey a common linear model. Brock, Durlauf, and West (2003) argue that one should assign priors that account for the interdependences implied by this structure in assigning model probabilities. Appendix 2 follows this approach in organizing growth regressors according to theory.

Doppelhofer, Miller, and Sala-i-Martin (2004) and Fernandez, Ley, and Steel (2001a) employ model averaging methods to identify which growth regressors should be

⁴⁵In our judgment, this presumption is unappealing as our own prior beliefs suggest that the true growth model is likely to contain many distinct factors. One implication of the openness of growth theories is that the simultaneous importance of many factors is certainly plausible.

included in linear growth models. These analyses do not distinguish between variables to be included in all regressions and variables whose inclusion determines alternative models; all variables are pooled and all possible combinations are considered. Doppelhofer, Miller, and Sala-i-Martin (2004) working with 31 potential growth determinants, conclude, weighting prior models so that the expected number of included regressors is 7 (this corresponds to a prior probability of variable inclusion of about .25), that four variables have posterior model inclusion probabilities above .9: initial income, fraction of GDP in mining, number of years the economy has been open,⁴⁶ and fraction of the population following Confucianism. Working with a universe of 41 potential growth determinants, Fernandez, Ley, and Steel find that, under the assumption that the prior probability that a given variable appears in the correct growth model is .5, four variables have posterior model inclusion probabilities above .9: initial income, fraction of the population following Confucianism, life expectancy, and rate of equipment investment.

Brock and Durlauf (2001a) and Masanjala and Papageorgiou (2004) employ model averaging to study the reason for the poor growth performance of sub-Saharan Africa. Brock and Durlauf (2001a) reexamine Easterly and Levine's (1997a) finding that ethnic heterogeneity helps explain sub-Saharan Africa's growth problems. This reanalysis finds that the Easterly and Levine (1997) claim is robust in the sense that ethnic heterogeneity helps explain why growth in sub-Saharan Africa had stagnated relative to the rest of the world. On the other hand, Brock and Durlauf (2001a) also find that ethnic heterogeneity does not appear to explain growth patterns in the rest of the world. This leads to the unresolved question of why ethnic heterogeneity has uniquely strong growth effects in sub-Saharan Africa. Masanjala and Papageorgiou (2004) conduct a general analysis of the determinants of sub-Saharan African growth versus the world as a whole and conclude that the relevant growth variables for Africa are quite different. In particular, variation in sub-Saharan growth is much more closely associated with the share of the economy made up by primary commodities production. They also find, contrary to Doppelhofer, Miller, and Sala-i-Martin (2004) that the share of mining in the economy is a robust determinant of growth in Africa but not the world as a whole.

⁴⁶Sachs and Warner (1995) use this variable as an index of overall openness of an economy.

Finally, model averaging has been applied by Brock, Durlauf, and West (2003) to analyze the question of how to employ growth regressions to evaluate policy recommendations. Specifically, the paper assesses the question of whether a policymaker should favor a reduction of tariffs for sub-Saharan African countries; the analysis assumes that the policymaker possesses mean/variance preferences with respect to the effects of changes in current policies with a constant tradeoff of mean against standard deviation of 1 to 2. The analysis finds strong support for a tariff reduction in that it concludes that a policymaker with these preferences should support a tariff reduction for any of the countries in sub-Saharan Africa unless the policymaker has a very strong prior that sub-Saharan African countries obey a distinct linear growth process from the rest of the world. In the case where the policymaker has a strong prior that sub-Saharan Africa is “different” from the rest of the world, there is sufficient uncertainty about the relationship between tariffs and growth for these countries that a change in the rates cannot be justified; the strong prior in essence means that the growth experiences of non-African countries have little effect on the precision of estimates of growth behavior that are constructed using data on sub-Saharan African countries in isolation.

ii. parameter heterogeneity

From its earliest stages, the use of linear growth models has generated considerable unease with respect to the statistical foundations of the exercise. Arguably, the data for very different countries cannot be seen as realizations associated with a common data generating process (DGP). For econometricians that have been trained to search for a good approximation to a DGP, the modeling assumptions and procedures of the growth literature can look arbitrary. One expression of this concern is captured in a famous remark in Harberger (1987): “What do Thailand, the Dominican Republic, Zimbabwe, Greece, and Bolivia have in common that merits their being put in the same regression analysis?”

Views differ on the extent to which this objection is fundamental. There is general agreement that, when studying growth, it will be difficult to recover a DGP even if one exists. In particular, the prospects for recovering causal effects are clearly weak. Those

who are only satisfied with the specification and estimation of a structural model, in which parameters are either ‘deep’ or correspond to precisely defined causal effects within a coherent theoretical framework, will be permanently disappointed.⁴⁷ The growth literature must have a less ambitious goal, namely to investigate whether or not particular hypotheses have any support in the data. In practice, growth researchers are looking for patterns and systematic tendencies that can increase our understanding of the growth process, in combination with historical analysis, case studies, and relevant theoretical models. Another key aim of empirical growth research, which is harder than it looks at first sight, is to communicate the degree of support for any patterns identified by the researcher.

The issue of parameter heterogeneity is essentially that raised by Harberger. Why should we expect disparate countries to lie on a common surface? Clearly this criticism could be applied to most empirical work in social science, whether the data points reflect the actions and characteristics of individuals and firms, or the aggregations of their choices that are used in macroeconometrics. What is distinctive about the growth context is not so much the lack of a common surface, as the way in which the sample size limits the scope for addressing the problem. In principle, one response would be to choose a more flexible model that has a stronger chance of being a good approximation to the data. Yet this can be hard, and an inherently fragile procedure, when the sample is rarely greater than 100 observations.

If parameter heterogeneity is present, the consequences are potentially serious, except in a special case. If a slope parameter varies randomly across units, and is distributed independently of the variables in the regression and the disturbances, the coefficient estimate should be an unbiased estimate of the mean of the parameter. The assumption of independence is not one, however, that may be expected in light of the body of growth theories. For example, when estimating the relationship between growth and investment, the marginal effect of investment will almost certainly be correlated with aspects of the economic environment that should also be included in the regression.

⁴⁷Note that this reflects the shortcomings of economic theory as well as those of data and econometric analysis.

The solution to this general problem is to change the specification in a way that allows greater flexibility in estimation. There are many ways of doing this. One approach is to consider more general functional forms than the canonical Solow regression which for comparison purposes we restate as:

$$\gamma_i = k + \beta \log y_{i,0} + \pi_n \log(n_i + g + \delta) + \pi_K \log s_{K,i} + \pi_H \log s_{H,i} + \varepsilon_i \quad (50)$$

Liu and Stengos (1999) estimate a semiparametric partially linear version of this model, namely

$$\gamma_i = k + f_\beta(\log y_{i,0}) + \pi_n \log(n_i + g + \delta) + \pi_K \log s_{K,i} + f_{\pi_H}(\log s_{H,i}) + \varepsilon_i \quad (51)$$

where $f_\beta(\cdot)$ and $f_{\pi_H}(\cdot)$ are arbitrary (except for variance smoothness requirements) functions. One important finding is that the value of $f_\beta(\log y_{i,0})$ is only negative when initial per capita income exceeds about \$1800. They also find a threshold effect in secondary school enrollment rates (their empirical proxy for $\log s_{H,i}$) so the variable is only associated with a positive impact on growth if it exceeds about 15%. Banerjee and Duflo (2003) use this same regression strategy to study nonlinearity in the relationship between changes in inequality and growth; their specification estimates a version of (51) where initial income and human capital savings enter linearly (along with some additional non-Solow variables) but with the addition on the right hand side of the function $f_G(G_{i,t} - G_{i,t-5})$ where $G_{i,t}$ is the Gini coefficient. Using a panel of 45 countries and 5 year growth averages, their analysis produces an estimate of $f_G(\cdot)$ which has an inverted U shape. One limitation of such studies is that they only allow for nonlinearity for a subset of growth determinants, an assumption that has little theoretical justification and is, from a statistical perspective, ad hoc; of course the approach is more general and less ad hoc than simply assuming linearity as is done in most of the literature.

Durlauf, Kourtellos, and Minkin (2001) extend this approach and estimate a version of the augmented Solow model that allows the parameters for each country to vary as functions of initial income, i.e.

$$\gamma_i = k(y_{i,0}) + \beta(y_{i,0}) \log y_{i,0} + \pi_n(y_{i,0}) \log(n_i + \delta + g) + \pi_K(y_{i,0}) \log s_{K,i} + \pi_H(y_{i,0}) \log s_{H,i} + \varepsilon_i \quad (52)$$

This formulation means that each initial income level defines a distinct Solow regression; as such it shifts the focus away from nonlinearity towards parameter heterogeneity, although the model is of course nonlinear in $y_{i,0}$. This approach reveals considerable parameter heterogeneity especially among the poorer countries. Durlauf, Kourtellos, and Minkin (2001) confirm Liu and Stengos (1999) in finding that $\beta(y_{i,0})$ is positive for low $y_{i,0}$ values and negative for higher ones. They also find that $\pi_K(y_{i,0})$ fluctuates greatly over the range of $y_{i,0}$ values in their sample. This work is extended in Kourtellos (2003a) who finds parameter dependence on initial literacy and initial life expectancy. The varying coefficient approach is also employed in Mamuneas, Savvides, and Stengos (2004) who analyze annual measures of total factor productivity for 51 countries. They consider a regression model of TFP in which the coefficient on human capital in the regression is allowed to depend on human capital both in isolation and in conjunction with a measure of trade openness (other coefficients are held constant). Constancy of the human capital coefficient is rejected across a range of specifications.

At a minimum, it generally makes sense for empirical researchers to test for neglected parameter heterogeneity, either using interaction terms or by carrying out diagnostic tests. Chesher (1984) showed that White's information matrix test can be used in this context. For the normal linear model with fixed regressors, Hall (1987) showed that, asymptotically, the information matrix test corresponds to a joint test for heteroskedasticity and non-normality. Later in the chapter, we discuss how evidence of heteroskedasticity should sometimes be seen as an indicator of misspecification.

Other authors have attempted to employ panel data to identify parameter heterogeneity without the imposition of a functional relationship between parameters and various observable variables. An important early effort is Canova and Marcet (1995). Defining $s_{i,t}$ as the logarithm of the ratio of a country's per capita income to the time t international aggregate value, Canova and Marcet estimate models of the form

$$s_{i,t} = a_i + \rho_i s_{i,t-1} + \varepsilon_{i,t}. \quad (53)$$

The long-run forecast of $s_{i,t}$ is given by $\frac{a_i}{1-\rho_i}$ with $1-\rho_i$ being the rate of convergence towards that value. Canova and Marcet estimate their model using data on the regions of Europe and on 17 western European countries. Restricting the parameters a_i and ρ_i to be constant across i gives a standard β -convergence test and yields an estimated annual rate of convergence of approximately 2%. On the other hand, allowing for heterogeneity in these parameters produces a “substantial”, statistically significant, dispersion of the implied long-run $s_{i,t}$ forecasts. Moreover, those forecasts are positively correlated with $s_{i,0}$, the initial values of $s_{i,t}$, implying a dependence of long-run outcomes on initial conditions contrary to the convergence hypothesis. For the country-level data, differences in initial conditions explain almost half the cross-sectional variation in long-run forecasts; in contrast, the role of standard control variables such as rates of physical and human capital accumulation and government spending shares is minor. The latter finding must be tempered by the fact that the sample variation in these controls is less than that in Barro (1991) or Mankiw, Romer, and Weil (1992), for example.

A similar approach is taken by Maddala and Wu (2000) who consider models of the form

$$\log y_{i,t} = \alpha_i + \rho_i \log y_{i,t-1} + u_{i,t} \quad (54)$$

which is of course very similar to the model analyzed by Marcet and Canova (1995). Employing shrinkage estimators for α_i and ρ_i , they conclude that convergence rates, measured as $\beta_i = -\log \rho_i$ exhibit substantial heterogeneity.

iii. nonlinearity and multiple regimes

In this section we discuss several papers that have attempted to disentangle the roles of heterogeneous structural characteristics and initial conditions in determining growth performance. These studies employ a wide variety of statistical methods in attempting to identify how initial conditions affect growth. Despite this, there is substantial congruence in the conclusions of these papers as these studies each provide evidence of the existence of convergence clubs even after accounting for variation in structural characteristics.

An early contribution to this literature is Durlauf and Johnson (1995) who use classification and regression tree (CART) methods to search for nonlinearities in the growth process as implied by the existence of convergence clubs.⁴⁸ The CART procedure identifies subgroups of countries that obey a common linear growth model based on the Solow variables. These subgroups are identified by initial income and literacy, a typical subgroup l is defined by countries whose initial income lies within the interval $\underline{y}_{l,y} \leq y_{i,0} < \bar{y}_{l,y}$ and whose literacy rate L_i lies in the interval $\underline{L}_{l,L} \leq L_i < \bar{L}_{l,L}$. The number of subgroups and the boundaries for the variable intervals that define them are chosen by an algorithm that trades off model complexity (i.e. the number of subgroups) and goodness of fit. Because the procedure sequentially splits the data into finer and finer subgroups, it gives the data a tree structure.

⁴⁸A detailed discussion of regression tree methods appears in Breiman, Friedman, Olshen and Stone (1984). The technical appendix of Durlauf and Johnson (1995) presents a treatment tailored to the specific question of identifying multiple regimes in growth models. Regression tree methods suffer from the absence of a well-developed asymptotic theory for testing the number of regimes that are present in a data set, but the procedure is consistent in the sense that under relatively weak conditions, if there are a finite number of regimes, as the sample size grows to infinity, the correct model will be revealed.

Durlauf and Johnson (1995) also test the null hypothesis of a common growth regime against the alternative hypothesis of a growth process with multiple regimes in which economies with similar initial conditions tend to converge to one another. Using income per capita and the literacy rate (as a proxy for human capital) to measure the initial level of development and, using the same cross-country data set as Mankiw, Romer, and Weil, Durlauf, and Johnson reject the single regime model required for global convergence. That is, even after controlling for the structural heterogeneity implied by Mankiw, Romer, and Weil's augmented version of the Solow model, there is a role for initial conditions in explaining variation in cross-country growth behavior.

Durlauf and Johnson's (1995) findings of multiple convergence clubs appear to be reinforced by subsequent research. Papageorgiou and Masanjala (2004) note that one possible source for Durlauf and Johnson's findings may occur due to the misspecification of the aggregate production function. As observed in Section II, the linear representation of the Solow model represents an approximation around the steady-state when the aggregate production function is Cobb-Douglas. Papageorgiou and Masanjala estimate a version of the Solow model based on a constant elasticity of substitution (CES) production function rather than the Cobb-Douglas, following findings in Duffy and Papageorgiou (2000). They then examine the question of whether or not Durlauf and Johnson's multiple regimes remain under the CES specification. Using Hansen's (2000) approach to sample splitting and threshold estimation, they find statistically significant evidence of thresholds in the data. The sample splits they estimate divide the data in four distinct growth regimes and are broadly consistent with those found by Durlauf and Johnson.⁴⁹

⁴⁹Motivated by the debate over trade openness and growth, Papageorgiou (2002) applies Hansen's method to the Durlauf and Johnson data with the trade share added to the set of variables on which sample splits may occur. He finds that this variable divides the middle-income countries into high and low growth groups obeying different growth processes; however openness does not appear to matter for high and low income countries. This suggests the importance of further work on which variables are most appropriate in characterizing threshold effects. Using the regression tree approach with a large collection of candidate split variables, Johnson and Takeyama (2001) find evidence of thresholds in US state economic growth behavior defined by variables likely to be proxies for capital/labor ratio, agglomeration effects, and communication effects.

These findings are extended in recent work due to Tan (2004) who employs a procedure known as GUIDE (generalized, unbiased interaction detection and estimation) to identify subgroups of countries which obey a common growth model.⁵⁰ Relative to CART, the GUIDE algorithm has two advantages: 1) the algorithm explicitly looks for interactions between explanatory variables when identifying splits and 2) some within model testing supplements the penalties for model complexity and thereby reduces the tendency for CART procedures to produce an excessive number of splits in finite samples. Tan (2004) finds strong evidence that measures of institutional quality and ethnic fractionalization define convergence clubs across a wide range of countries. He also finds weaker evidence that geography distinguishes the growth process for sub-Saharan Africa from the rest of the world.

Further research has corroborated the evidence of multiple regimes using alternative statistical methods. One approach that has proven useful is based on projection pursuit methods⁵¹. Desdoigts (1999) uses these methods in an attempt to separate the roles of microeconomic heterogeneity and initial conditions in the growth experiences of a group of countries and identifies groups of countries with relatively homogeneous growth experiences based on data about the characteristics and initial conditions of each country. The idea of projection pursuit is to find the orthogonal projections of the data into low dimensional spaces that best display some interesting feature of the data. A well-known special case of projection pursuit is principal components analysis. In principal components analysis, one takes only as many components as are necessary to account for “most” of the variation in the data. Similarly, in projection pursuit one should only consider as many dimensions as needed to account for “most” of the clustering in the data.

Desdoigts finds several interesting clusters. The first is the OECD countries. The two projections identifying this cluster put most of their weight on the primary and secondary school enrollment rates, the 1960 income gap and the rate of growth in the labor force. The prominence of variables that Desdoigts argues are proxies for initial

⁵⁰GUIDE originates in Loh (2002).

⁵¹Projection pursuit is developed in Friedman and Tukey (1974) and Friedman (1987). Appendix A of Desdoigts (1999) provides a useful primer.

conditions among those defining the projections leads him to conclude that initial conditions are more important in defining this cluster than are other country characteristics. Reapplication of the clustering method to the remaining (non-OECD) countries yields three sub-clusters that can be described as Africa, Southeast Asia, and Latin America. Here the projections put most weight on government consumption, the secondary school enrollment rate and investment in electrical machinery and transportation equipment. Most of these variables are argued to proxy for structural characteristics of the economies, suggesting that they, rather than initial conditions, are responsible for the differences in growth experiences across the three geographic sub-clusters. Nevertheless, this approach relies on the judgment of the researcher in determining which variables proxy for initial conditions and which proxy for structural characteristics.

Further evidence of the utility of projection pursuit methods may be found in Kourtellos (2003b). Unlike Desdoigts, Kourtellos (2003b) uses projection pursuit to construct models of the growth process. Formally, he estimates models of the form

$$\gamma_i = \sum_{l=1}^L f_l (y_{i,0}\beta_l + X_i\psi_l + Z_i\pi_l) + \varepsilon_i \quad (55)$$

Each element in the summation represents a distinct projection. Kourtellos uncovers evidence of two steady-states, one for low initial income and low initial human capital countries.

A third approach to multiple regimes is employed by Bloom, Canning, and Sevilla (2003) based on the observation that if long-run outcomes are determined by fundamental forces alone, the relationship between exogenous variables and income levels ought to be unique. If initial conditions play a role there will be multiple relationships – one for each basin of attraction defined by initial conditions. If there are two (stochastic) steady states, and large shocks are sufficiently infrequent,⁵² the system will, under suitable regularity

⁵²The assumed rarity of large shocks implies that movements between basins of attraction of each of the steady states are sufficiently infrequent that they can be ignored in

conditions, exhibit an invariant probability measure that can be described by a “reduced form” model in which the long-run behavior of $\log y_{i,t}$ depends only on the exogenous variables, m_i , such as

$$\log y_{i,t} = \log y_1^*(m_i) + u_{1,i,t} \text{ with probability } p(m_i) \quad (56)$$

and

$$\log y_{i,t} = \log y_2^*(m_i) + u_{2,i,t} \text{ with probability } 1 - p(m_i) \quad (57)$$

where $u_{1,i,t}$ and $u_{2,i,t}$ are independent, zero-mean deviations from the steady-state log means $\log y_1^*(m_i)$ and $\log y_2^*(m_i)$ respectively, and $p(m_i)$ is the probability that country i is in the basin of attraction of the first of the two steady states. From the perspective of the econometrician, $\log y_{i,t}$ thus obeys a mixture process. The two steady states associated with (56) and (57) are possibly interpretable as a low-income regime or poverty trap and as a high-income or perpetual growth regime respectively. Bloom, Canning and Sevilla estimate a linear version of this model using 1985 income data from 152 countries with the absolute value of the latitude of the (approximate) center of each country as the fundamental exogenous variable. They are able to reject the null hypothesis of a single regime model in favor of the alternative of a model with two regimes – a high-level (manufacturing and services) steady state in which income is independent of latitude and a low-level (agricultural) steady-state in which income depends on latitude (presumably through its influence on climate). In addition, the probability of being in the high-level steady state is found to rise with latitude.

A final approach to multiple regimes is due to Canova (2004) who introduces a procedure for panel data that estimates the number of groups and the assignment of countries or regions to these groups, drawing on Bayesian ideas. This approach has the

estimation. This assumption is consistent with, for example, Bianchi’s (1997) finding of very little mobility in the cross-country income distribution.

important feature that it allows for parameter heterogeneity across-countries within a given subgroup. The researcher can order the countries or regions by various criteria (for example, output per capita in the pre-sample period) and the estimation procedure then chooses break points and group membership in such a way that the predictive ability of the overall model is maximized. This approach is applied to autoregressive models of per capita output as in eq. (54) above.

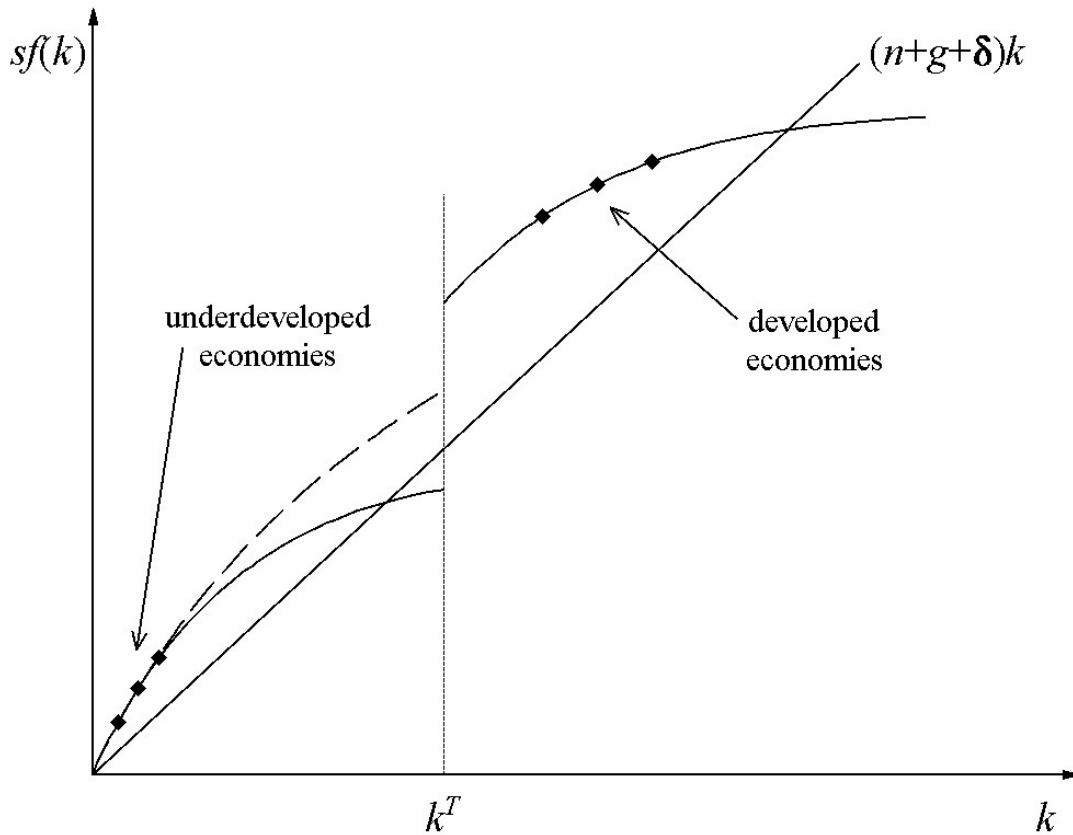
Using data on per capita income data in the regions of Europe, Canova (2004) finds that ordering the data by initial income maximizes the marginal likelihood of the model and breaks the data into 4 clusters. The estimated mean steady-states for each group are significantly different from each other implying that the groups are convergence clubs. The differences in the means are also economically important with the lowest and highest being 45% and 115% of the overall average respectively. Canova finds little across-group mobility especially among those regions that are initially poor. Using data on per capita income in the OECD countries, two clusters are found and, again, initial per capita income is the preferred ordering variable. The estimated model parameters imply an “economically large” long-run difference in the average incomes of countries in the two groups with little mobility between them.

In assessing these analyses, it is important to recognize an identification problem in attempting to link evidence of multiple growth regimes to particular theoretical growth models. As argued in Durlauf and Johnson (1995), this identification problem relates to whether evidence of multiple regimes represents evidence of multiple steady-states as opposed to nonlinearity in the growth process.

To see why this is so, suppose that one has identified two sets of countries that obey separate growth regimes with regime membership determined by a country’s initial capital stock, i.e. there exists a capital threshold k^T that divides the two groups of countries. An example of this can be seen in Figure 8. Clearly, the two sets of countries do not obey a common linear model but it is not clear whether or not multiple steady-states exist. The output behavior of low capital countries is compatible with either the solid or dashed curve in the lower part of the figure, but only the solid curve produces multiple steady-states. The identification problem stems from the fact that one does not have observations that allow one to distinguish differences in the long-run behavior of

countries that start with capital stocks in the vicinity of k^T . This argument does not depend on growth regimes determined by the capital stock but it does depend on whether or not the variable or variables that define the regimes are growing over time, as would occur for initial income or initial literacy. For growing variables, the possibility exists that countries currently associated with low levels of the variable will in the future exhibit behaviors that are similar to those countries which are currently associated with high levels of the variables.

Figure 8: Nonlinearity Versus Multiple Steady-States



How might evidence of multiple steady-states be achieved? One possibility is via the use of structural models in empirical analysis. While this has not been done econometrically, Graham and Temple (2003) follow this strategy and calibrate a two-sector general equilibrium model with increasing returns to scale in nonagricultural production. Their empirically motivated choice of calibration parameters produces a model which implies that some countries are in a low-output equilibrium. Another possibility is to exploit time series variation in a single country to identify the presence of jumps from one equilibrium or steady state to another.

VI. Econometric issues I: alternative data structures

Our discussion of growth econometrics now shifts from general issues of hypothesis testing and model specification to explore specific econometric issues that arise in the estimation of growth models. This section reviews econometric issues that arise for the different types of data structures that appear in growth analyses. By data structures, we refer to features such as whether the data are observed in cross-section, time series, or panel as well as to whether particular data series are conceptualized as endogenous or exogenous. At the risk of stating the obvious, choices of method involve significant trade-offs, which depend partly on statistical considerations and partly on the economic context. This means that attempts at universal prescriptions are misguided, and we will try to show the desirability of matching techniques to the economic question at hand.

One example, to be discussed further below, would be the choice between panel data methods and the estimation of separate time series regressions for each country. The use of panel data is likely to increase efficiency and allow richer models to be estimated, but at the expense of potentially serious biases if the parameter homogeneity assumptions are incorrect. This trade-off between robustness and efficiency is another running theme of our survey. The scientific solution would be to base the choice of estimation method on a context-specific loss function, but this is clearly a difficult task, and in practice more subjective decisions are involved.

This section has four main elements. Section VI.i examines econometric issues that arise in the use of time series data to study growth, emphasizing some of the drawbacks of this approach. Section VI.ii discusses the many issues that arise when panel data are employed, an increasingly popular approach to growth questions. We consider the estimation of dynamic models in the presence of fixed effects, and alternatives to standard procedures. Section VI.iii describes another increasingly popular approach, namely the use of “event studies” to analyze growth behavior, based on studying responses to major shocks such as policy reforms. Section VI.iv examines endogeneity and the use of instrumental variables. We argue that the use of instrumental variables in growth contexts is more problematic than is often appreciated and suggest

the importance of combining instrumental variable choice with a systematic approach to model selection.

i. time series approaches

At first glance, the most natural way to understand growth would be to examine time series data for each country in isolation. As we saw previously, growth varies substantially over time, and countries experience distinct events that contribute to this variation, such as changes in government and in economic policy.

In practice a time series approach runs into substantial difficulties. One key constraint is the available data. For many developing countries, some of the most important data are only available on an annual basis, with limited coverage before the 1960s. Moreover, the listing of annual data in widely used sources and online databases can be misleading, because some key variables are measured less frequently. For example, population figures are often based primarily on census data, while measures of average educational attainment are often constructed by interpolating between census observations using school enrollments. When examining published data, it is not always clear where this kind of interpolation has been used. The true extent of information in the time series variation may be less than appears at first glance, and conventional standard errors on parameter estimates will be misleading when interpolated data are used.

Even where reliable data are available, some key growth determinants display relatively little time variation, a point that has been emphasized by Easterly et al (1993), Easterly (2001) and Pritchett (2000a). There do exist other variables that appear to show significant variation, but this variation may not correspond to the concept the researcher has in mind. An example would be political stability. Since Barro (1991), researchers have sometimes used the incidence of political revolutions and coups as a measure of political instability. The interpretation of such an index clearly varies depending on the length of the time period used to construct it. If the hypothesis of interest relates to underlying political uncertainty (say, the ex-ante probability of a transfer of power) then the observations on political instability would need to be averaged over a long time

period. The variation in political instability at shorter horizons only casts light on a different hypothesis, namely the direct impact of revolutions and coups.

There are other significant problems with the time series approach. The hypotheses of most interest to growth theorists are mainly about the evolution of potential output, not deviations from potential output such as business cycles and output collapses. Since measured output is a noisy indicator of potential output, it is easy for the econometric modeling of a growth process to be contaminated by business cycle dynamics. A simple way to illustrate this would be to consider what happens if measured log output is equal to the log of potential output plus a random error. If log output is trend stationary, this is a classical measurement error problem. When lags of output or the growth rate are used as explanatory variables, the parameter estimates will be inconsistent.

Such problems are likely to be even more serious in developing countries, where large slumps or crises are not uncommon, and output may deviate for long periods from any previous structural trend (Pritchett (2000a)). We have already seen the extent to which output behaves very differently in developing countries compared to OECD members, and a major collapse in output is not a rare event. There may be no underlying trend in the sense commonly understood, and conventional time series methods should be applied with caution. Some techniques that are widely used in the literature on business cycles in developed countries, such as the Hodrick-Prescott filter, will often be inappropriate in the context of developing countries.

The problem of short-run output instability extends further. It is easy to construct examples where the difference between observed output and potential output is correlated with variables that move up and down at high frequencies, with inflation being one obvious example. This means that time series studies of inflation and growth based on observed output will find it hard to isolate reliably an effect of inflation on potential output; for further discussion see Temple (2000a). When considerations like these are combined with the paucity of the available data, it appears a hard task to learn about long-term growth using time series regressions, especially when developing countries are the main focus of interest.

Nevertheless, despite these problems, there are some hypotheses for which time series variation can be informative. We have already seen the gains from time series approaches to convergence issues. Jones (1995) and Kocherlakota and Yi (1997) show how time series models can be used to discriminate between different growth theories. To take the simplest example, the AK model of growth predicts that the growth rate will be a function of the share of investment in GDP. Jones points out that investment rates have trended upwards in many OECD countries, with no corresponding increase in growth rates. Although this might be explained by offsetting changes in other growth determinants, it does provide evidence against simple versions of the AK model.

Jones (1995) and Kocherlakota and Yi (1997) develop a statistical test of endogenous growth models based on regressing growth on lagged growth and a lagged policy variable (or the lagged investment rate, as in Jones). Exogenous growth models predict that the coefficients on the lagged policy variable should sum to zero, indicating no long-run effect of permanent changes in this variable on the growth rate. In contrast, some endogenous growth models imply that the sum of coefficients should be non-zero. A simple time series regression then provides a direct test of the predictions of these models. More formally, as in Jones (1995), for a given country i one can investigate a dynamic relationship for the growth rate $\gamma_{i,t}$ where

$$\gamma_{i,t} = A(L)\gamma_{i,t-1} + B(L)z_{i,t} + \varepsilon_{i,t} \quad (58)$$

where z is the policy variable or growth determinant of interest, and $A(L)$ and $B(L)$ are lag polynomials assumed to be compatible with stationarity. The hypothesis of interest is whether $B(1) \neq 0$. If the sum of the coefficients in the lag polynomial $B(L)$ is significantly different from zero, this implies that a permanent change in the variable z will affect the growth rate indefinitely. As Jones (1995) explicitly discusses, this test is best seen as indicating whether a policy change affects growth over a long horizon, rather than firmly identifying or rejecting the presence of a long-run growth effect in the theoretical sense of that term. The theoretical conditions under which policy variables

affect the long-run growth rate are remarkably strict, and many endogenous growth models are best seen as new theories of potentially sizeable level effects.⁵³

This approach is closely related to Granger-causality testing, where the hypothesis of interest would be the explanatory power of lags of $Z_{i,t}$ for $\gamma_{i,t}$ conditional on lagged values of $\gamma_{i,t}$. Blomstrom, Lipsey, and Zejan (1996) carry out Granger-causality tests for investment and growth using panel data with five-year subperiods. They find strong evidence that lagged growth rates have explanatory power for investment rates, but much weaker evidence for causality in the more conventional direction from investment to growth. Hence, the partial correlation between growth and investment found in many cross-section studies may not reflect a causal effect of investment. In a similar vein, Campos and Nugent (2002) find that, once Granger-causality tests are applied, the evidence that political instability affects growth may be weaker than usually believed.

The motivation for these two studies, and others like them, is that evidence of temporal precedence helps to build a case that one variable is influenced by another. When this idea is extended to panels, an underlying assumption is that timing patterns and effects will be similar across units (countries or regions). Potential heterogeneity has sometimes been acknowledged, as in the observation of Campos and Nugent (2002) that their results are heavily influenced by the African countries in the sample. The potential importance of these factors is also established in Binder and Brock (2004) who, by using panel methods to allow for heterogeneity in country-specific dynamics, find feedbacks from investment to growth beyond those that appear in Blomstrom, Lipsey, and Zejan (1996).

A second issue is more technical. Since testing for Granger-causality using panel data requires a dynamic model, the use of a standard fixed effects (within groups) estimator is likely to be inappropriate when individual effects are present. We discuss this further in section VI below. One potential solution is the use of instrumental variable procedures, as in Campos and Nugent (2002). In the context of investment and growth, a comprehensive examination of the associated econometric issues has been carried out by

⁵³See Temple (2003) for more discussion of this point and the long-run implications of different growth models.

Bond, Leblebicioglu, and Schiantarelli (2004). Their work shows that these issues are more than technicalities: unlike Blomstrom, Lipsey, and Zejan (1996), they find strong evidence that investment has a causal effect on growth.

A familiar objection to the more ambitious interpretations of Granger-causality is that much economic behavior is forward-looking (see for example Klenow and Rodriguez-Clare (1997b)). The movements of stock markets are one instance where temporal sequences can be misleading about causality. Similarly, when entrepreneurs or governments invest heavily in infrastructure projects, or when unusually high inflows of foreign direct investment are observed, the fact that such investments precede strong growth does not establish a causal effect.

ii. panel data

As we emphasized above, the prospects for reliable generalizations in empirical growth research are often constrained by the limited number of countries available. This constraint makes parameter estimates imprecise, and also limits the extent to which researchers can apply more sophisticated methods, such as semiparametric estimators.

A natural response to this constraint is to use the within-country variation to multiply the number of observations. Using different episodes within the same country is ultimately the only practical substitute for somehow increasing the number of countries. To the extent that important variables change over time, this appears the most promising way to sidestep many of the problems that face growth researchers. Moreover, as the years pass and more data become available, the prospects for informative work of this kind can only improve.

We first discuss the implementation and advantages of panel data estimators in more detail, and then some of the technical issues that arise in the context of growth. Perhaps not surprisingly, these methods introduce a set of problems of their own, and should not be regarded as a panacea. Too often, panel data results are interpreted without sufficient care and risk leading researchers astray. In particular, we highlight the care needed in interpreting estimates based on fixed effects.

We will use T to denote the number of time series observations in a panel of N countries or regions. At first sight, T should be relatively high in this context, because of the availability of annual data. But the concerns about time series analysis raised above continue to apply. Important variables are either measured at infrequent intervals, or show little year-to-year variation that can be used to identify their effects. Moreover, variation in growth rates at annual frequencies may give very misleading answers about the longer-term growth process. For this reason, most panel data studies in the growth field have averaged data over five or ten year periods. Given the lack of data before 1960, this implies that growth panels not only have relatively few cross-sectional units (the number of countries employed is often between 50 and 100) but also very low values of T , often 5 or 6 at most.⁵⁴

Most empirical growth models estimated using panel data are based on the hypothesis of conditional convergence, namely that countries converge to parallel equilibrium growth paths, the levels of which are a function of a few variables. A corollary is that an equation for growth (essentially the first difference of log output) should contain some dynamics in lagged output. In this case, the growth equation can be rewritten as a dynamic panel data model in which current output is regressed on controls and lagged output, as in Islam (1995). In statistical terms this is the same model, the only difference of interpretation being that the coefficient on initial output (originally β) is now $1 + \beta$:

$$\log y_{i,t} = (1 + \beta) \log y_{i,t-1} + \psi X_{i,t} + \pi Z_{i,t} + \alpha_i + \mu_t + \varepsilon_{i,t} \quad (59)$$

This regression is a general panel analog to the cross-section regression (18). In this formulation, α_i is a country-specific effect and μ_t is a time-specific effect. The inclusion of time-specific effects is important in the growth context, not least because the means of the log output series will typically increase over time, given productivity growth at the world level.

⁵⁴This is true of the many published studies that have used version 5.6 of the Penn World Tables. Now that more recent data are available, there is more scope for estimating panels with a longer time dimension.

Inclusion of a country-specific effect allows permanent differences in the level of income between countries that are not captured by $X_{i,t}$ or $Z_{i,t}$. In principle, one can also allow the parameters $1+\beta$, ψ , and π to differ across i ; Lee, Pesaran, and Smith (1997,1998) do this for the coefficients for $\log y_{i,t-1}$ and a linear time trend (the latter allowing for steady-state differences in the rate of technological change, corresponding to non-parallel growth paths in the steady state).

The vast majority of panel data growth studies use a fixed effects (within-group) estimator rather than a random effects estimator. Standard random effects estimators require that the individual effects α_i are distributed independently of the explanatory variables, and this requirement is clearly violated for a dynamic panel such as (59) by construction, given the dependence of $\log y_{i,t}$ on α_i .

Given the popularity of fixed effects estimators, it is important to understand how these estimators work. In a fixed effects regression there is a full set of country-specific intercepts, one for each country, and inference proceeds conditional on the particular countries observed (a natural choice in this context). Identification of the slope parameters, usually constrained to be the same across countries, relies on variation over time within each country. The “between” variation, namely the variation across countries in the long-run averages of the variables, is not used.

The key strength of this method, familiar from the microeconomic literature, is the ability to address one form of unobserved heterogeneity: any omitted variables that are constant over time will not bias the estimates, even if the omitted variables are correlated with the explanatory variables. Intuitively, the country-specific intercepts can be seen as picking up the combined effects of all such variables. This is the usual motivation for using fixed effects in the growth context, especially in estimating conditional convergence regressions, as is further discussed in Islam (1995), Caselli Esquivel, and Lefort (1996) and Temple (1999). A particular motivation for the use of fixed effects arises from the Mankiw, Romer, and Weil (1992) implementation of the Solow model. As discussed in Section III, their version of the model implies that one determinant of the level of the steady-state growth path is the initial level of efficiency ($A_{i,0}$) and cross-section heterogeneity in it should usually be regarded as unobservable,

cf. eq. (15). Islam (1995) explicitly develops a specification in which this term is treated as a fixed effect, while world growth and common shocks are incorporated using time-specific effects.

The use of panel data methods to address unobserved heterogeneity can bring substantial gains in robustness, but is not without costs. The fixed-effects identification strategy cannot be applied in all contexts. Sometimes a variable of interest is measured at only one point in time. Even where variables are measured at more frequent intervals, some are highly persistent, in which case the within-country variation is unlikely to be informative. At one extreme, some explanatory variables of interest are essentially fixed factors, like geographic characteristics or ethnolinguistic diversity. Here the only available variation is “between-country”, and empirical work will have to be based on cross-sections or pooled cross-section time-series. Alternatively a two-stage hybrid of these methods can be used, in which a panel data estimator is used to obtain estimates of the fixed effects, which are then explicitly modeled in a second stage as in Hoeffler (2002). As we discuss further below, an important direction for future panel data work may be the analysis of the information content of country-specific effects.

A common failing of panel data studies based on within-country variation is that researchers do not pay enough attention to the dynamics of adjustment. There are many panel data papers on human capital and growth that test only whether a change in school enrollment or years of schooling has an immediate effect on aggregate productivity, which seems an implausible hypothesis. Another example, given by Pritchett (2000a), is the use of panels to study inequality and growth. All too often, changes in the distribution of income are implicitly expected to have an immediate impact on growth. Yet many of the relevant theoretical papers highlight long-run effects, and there is a strong presumption that much of the short-run variation in measures of inequality is due to measurement error. In these circumstances, it is hard to see how the available within-country variation can shed much useful light.

There is also a more general problem. Since the fixed effects estimator ignores the between-country variation, the reduction in bias typically comes at the expense of higher standard errors. Another reason for imprecision is that either of the devices used to eliminate the country-specific intercepts – the within-groups transformation or first-

differencing – will tend to exacerbate the effect of measurement error.⁵⁵ As a result, it is common for researchers using panel data models with fixed effects, especially in the context of small T , to obtain imprecise sets of parameter estimates.

Given the potentially unattractive trade-off between robustness and efficiency, Barro (1997), Temple (1999), Pritchett (2000a) and Wacziarg (2002) all argue that the use of fixed effects in empirical growth models has to be approached with care. The price of eliminating the misleading component of the between variation – namely, the variation due to unobserved heterogeneity – is that all the between variation is lost.

There are alternative ways to reveal this point, but consider the random effects GLS estimator of the slope parameters, which will be more efficient than the within-country estimator for small T when the random-effects assumptions are appropriate. This GLS estimator can be written as a matrix-weighted average of the within-country estimator and the between-country estimator, which is based on averaging the data over time and then estimating a simple cross-section regression by OLS.⁵⁶ The weights on the two sets of parameter estimates are the inverses of their respective variances. The corollary of high standard errors using within-country estimation, indicating that the within-country variation is relatively uninformative, is that random effects estimates based on a panel of five-yearly averages are very similar to OLS estimates based on thirty-year averages (Wacziarg (2002)). Informally, the random effects estimator sees the between-country variation as offering the greatest scope for identifying the parameters.⁵⁷

This should not be surprising: growth episodes within countries inevitably look a great deal more alike than growth episodes across countries, and therefore offer less identifying variation. Restricting the analysis to the within variation eliminates one source of bias, but immediately makes it harder to identify growth effects with any degree of precision. This general problem is discussed in Pritchett (2000a). Many of the

⁵⁵See Arellano (2003, p. 47-51) for a more formal treatment of this issue.

⁵⁶This result holds for the GLS estimator of the random effects model. In practice, since the true variance components are unknown, feasible GLS must be used.

⁵⁷Of course, this does not imply that the random effects estimator is the best choice; as we have seen, the underlying assumptions for consistency of the estimator are necessarily invalid for a dynamic panel. Instead, our discussion is intended to draw attention to the trade-off between bias and efficiency in deciding whether or not to use fixed-effects estimation.

explanatory variables currently used in growth research are either highly stable over time, or tending to trend in one direction. Educational attainment is an obvious example. Without useful identifying variation in the time series data, the within-country approach is in trouble. Moreover, growth is quite volatile at short horizons. It will typically be hard to explain this variation using predictors that show little variation over time, or that are measured with substantial errors. The result has been a number of panel data studies suggesting that a given variable “does not matter” when a more accurate interpretation is that its effect cannot be identified using the data at hand.

Some of these problems suggest a natural alternative to the within-country estimator, which is to devote more attention to modeling the heterogeneity, rather than treating it as unobserved (Temple (1999)). To put this differently, current panel data methods treat the individual effects as nuisance parameters. As argued by Durlauf and Quah (1999) this is clearly inappropriate in the growth context. The individual effects are of fundamental interest to growth economists because they appear to be a key source of persistent income differences. This suggests that more attention should be given to modeling the heterogeneity rather than finding ways to eliminate its effects.⁵⁸

Depending on the sources of heterogeneity, even simple recommendations, such as including a complete set of regional dummies, can help to alleviate the biases associated with omitted variables. More than a decade of growth research has identified a host of fixed factors that could be used to substitute for country-specific intercepts. A growth model that includes these variables can still exploit the panel structure of the data, and overall this approach has clear advantages in both statistical and economic terms. It means that the between variation is retained, rather than entirely thrown away, while the explicit modeling of the country-specific effects is directly informative about the sources of persistent income and growth differences.

In practice, the literature has focused on another aspect of using panel data estimators to investigate growth. Nickell (1981) showed that within-groups estimates of a

⁵⁸Note that fixed-effects estimators could retain a useful role, because it would be natural to compare their parameter estimates with those obtained using a specific model for the heterogeneity. Where the estimates of common parameters, such as the coefficient on the lagged dependent variable, are different across the two methods, this could indicate the chosen model for the heterogeneity is misspecified.

dynamic panel data model can be badly biased for small T , even as N goes to infinity. The direction of this bias is such that, in a growth model, output appears less persistent than it should (the estimate of β is too low) and the rate of conditional convergence will be overestimated.

In other areas of economics, it has become increasingly common to avoid the within-groups estimator when estimating dynamic models. The most widely-used alternative strategy is to difference the model to eliminate the fixed effects, and then use two stage least squares or GMM to address the correlation between the differenced lagged dependent variable and the induced MA(1) error term. To see the need for instrumental variable procedures, first-difference (59) to obtain

$$\Delta \log y_{i,t} = (1 + \beta)\Delta \log y_{i,t-1} + \Delta X_{i,t}\psi + \Delta Z_{i,t}\pi + \Delta \mu_i + \varepsilon_{i,t} - \varepsilon_{i,t-1} \quad (60)$$

and note that (absent an unlikely error structure) the $\log y_{i,t-1}$ component of $\Delta \log y_{i,t-1}$ will be correlated with the $\varepsilon_{i,t-1}$ component of the new composite error term, as is clearly seen by considering equation (59) lagged one period. Hence, at least one of the explanatory variables in the first-differenced equation will be correlated with the disturbances, and instrumental variable procedures are required.

Arellano and Bond (1991), building on work by Holtz-Eakin, Newey, and Rosen (1988), developed the GMM approach to dynamic panels in detail, including methods suitable for unbalanced panels and specification tests. Caselli, Esquivel, and Lefort (1996) applied their estimator in the growth context and, as discussed above, this approach yielded a much faster rate of conditional convergence than found in cross-section studies.

The GMM approach is typically based on using lagged levels of the series as instruments for lagged first differences. If the error terms in the levels equation (ε_{it}) are serially uncorrelated then $\Delta \log y_{i,t-1}$ can be instrumented using $\log y_{i,t-2}$ and earlier lagged levels (where available). This corresponds to a set of moment conditions that can be used to estimate the first-differenced equation by GMM. Bond (2002) provides an accessible introduction to this approach.

As an empirical strategy for growth research, this has some appeal, because it could alleviate biases due to measurement error and endogenous explanatory variables. In practice, many researchers are skeptical that lags are suitable instruments. It is easy to see that a variable such as educational attainment may influence output with a considerable delay, so that the exclusion of lags from the growth equation can look arbitrary. More generally, the GMM approach relies on a lack of serial correlation in the error terms of the growth equation (before differencing). Although this assumption can be tested using the methods developed in Arellano and Bond (1991), and can also be relaxed by an appropriate choice of instruments, it is nevertheless restrictive in some contexts.

Another concern is that the explanatory variables may be highly persistent, as is clearly true of output. Lagged levels can then be weak instruments for first differences, and the GMM panel data estimator is likely to be severely biased in short panels. Bond, Hoeffler, and Temple (2001) illustrate this point by comparing the Caselli, Esquivel, and Lefort (1996) estimates of the coefficient on lagged output with OLS and within-group estimates. Since the OLS and within-group estimates of β are biased in opposing directions then, leaving aside sampling variability and small-sample considerations, a consistent parameter estimate should lie between these two extremes (see Nerlove (1999,2000)). Formally, when the explanatory variables other than lagged output are strictly exogenous, we have

$$p \lim \hat{\beta}_{WG} < p \lim \hat{\beta} < p \lim \hat{\beta}_{OLS} \quad (61)$$

where $\hat{\beta}$ is a consistent parameter estimate, $\hat{\beta}_{WG}$ is the within-groups estimate and $\hat{\beta}_{OLS}$ is the estimate from a straightforward pooled OLS regression. For the data set and model used by Caselli, Esquivel and Lefort, this large-sample prediction is not valid, which raises a question mark over the reliability of the first-differenced GMM estimates.

One device that can be informative in short panels is to make more restrictive assumptions about the initial conditions. If the observations at the start of the sample are distributed in a way that is representative of steady-state behavior, in a sense that can be made more precise, efficiency gains are possible. Assumptions about the initial

conditions can be used to derive a “system” GMM estimator, of the form developed and studied by Arellano and Bover (1995) and Blundell and Bond (1998), and also discussed in Ahn and Schmidt (1995) and Hahn (1999). In this estimator, not only are lagged levels used as instruments for first differences, but lagged first differences are used as instruments for levels, which corresponds to an extra set of moment conditions.

There is some Monte Carlo evidence (Blundell and Bond (1998)) that this estimator is more robust than the Arellano-Bond method in the presence of highly persistent series. As also shown by Blundell and Bond (1998), the necessary assumptions can be seen in terms of an extra restriction, namely that the deviations of the initial values of $\log y_{i,t}$ from their long-run values are not systematically related to the individual effects.⁵⁹ For simplicity, we focus on the case where there are no explanatory variables other than lagged output. The required assumption on the initial conditions is that, for all $i = 1, \dots, N$ we have

$$E\left[(\log y_{i,1} - \bar{y}_i)\alpha_i\right] = 0 \quad (62)$$

where the \bar{y}_i are the long-run values of the $\log y_{i,t}$ series and are therefore functions of the individual effects α_i and the autoregressive parameter β . This assumption on the initial conditions ensures that

$$E\left[\Delta \log y_{i,2}\alpha_i\right] = 0 \quad (63)$$

and this together with the mild assumption that the changes in the errors are uncorrelated with the individual effects, i.e.

$$E\left[\Delta \varepsilon_{i,t}\alpha_i\right] = 0 \quad (64)$$

⁵⁹Note that the long-run values of log output are evolving over time when time-specific effects are included in the model.

implies $T - 2$ extra moment conditions of the form

$$E\left[\Delta \log y_{i,t-1}(\alpha_i + \varepsilon_{i,t})\right] = 0 \text{ for } i = 1, \dots, N \text{ and } t = 3, 4, \dots, T \quad (65)$$

Intuitively, as is clear from the new moment conditions, the extra assumptions ensure that the lagged first difference of the dependent variable is a valid instrument for untransformed equations in levels since it is uncorrelated with the composite error term in the levels equation. These extra moment conditions can then be combined with the more conventional conditions used in the Arellano-Bond method. This builds in some insurance against weak identification, because if the series are persistent and lagged levels are weak instruments for first differences, it may still be the case that lagged first differences have some explanatory power for levels.⁶⁰

In principle, the validity of the restrictions on the initial conditions can be tested using the incremental Sargan statistic (or C statistic) associated with the additional moment conditions. Yet the validity of the restriction should arguably be evaluated in wider terms, based on some knowledge of the historical forces giving rise to the observed initial conditions. This point – that key statistical assumptions should not always be evaluated only in statistical terms – is one that we will return to later.

Alternatives to GMM have been proposed. Kiviet (1995,1999) derives an analytical approximation to the Nickell bias that can be used to construct a bias-adjusted within-country estimator for dynamic panels. The simulation evidence reported in Judson and Owen (1999) and Bun and Kiviet (2001) suggests that this estimator performs well relative to standard alternatives when N and T are small. One minor limitation is that it cannot yet be applied to an unbalanced panel. A more serious limitation, relative to GMM, is that it does not address the possible correlation between the explanatory variables and the disturbances due to simultaneity and measurement error. Nevertheless, for researchers determined to use fixed effects estimation, there is a clear case for implementing this bias adjustment, at least as a complement to other methods.

⁶⁰An alternative approach would be to use small-sample bias adjustments for GMM panel data estimators, such as those described in Hahn, Hausman, and Kuersteiner (2001).

A further issue that arises when estimating dynamic panel data models is that of parameter heterogeneity. If a slope parameter such as β varies across countries, and the explanatory variable is serially correlated, this will induce serial correlation in the error term. If we focus on a simple case where a researcher wrongly assumes $\beta_i = \beta$ for all $i = 1, \dots, N$ then the error process for a given country will contain a component that resembles $(\beta_i - \beta) \log y_{i,t-1}$. Hence there is serial correlation in the errors, given the persistence of output. The estimates of a dynamic panel data model will be inconsistent even if GMM methods are applied.

This problem was analyzed in more general terms by Robertson and Symons (1992) and Pesaran and Smith (1995) and has been explored in great depth for the growth context by Lee, Pesaran, and Smith (1997, 1998). Since an absence of serial correlation in the disturbances is usually a critical assumption for the GMM approach, parameter heterogeneity can be a serious concern. Some of the possible solutions, such as regressions applied to single time series, or the pooled mean group estimator developed by Pesaran, Shin, and Smith (1999), have limitations in studying growth for reasons already discussed. An alternative solution is to split the sample into groups that are more likely to share similar parameter values. Groupings by regional location or level of development are a natural starting point.

Perhaps the state of the art in analyzing growth using panel data and allowing for parameter heterogeneity is represented by Phillips and Sul (2003). They allow for heterogeneity in parameters not only across countries, but also over time. Temporal heterogeneity is rarely investigated in panel studies, but may be important, especially if observed growth patterns combine transitional dynamics towards a country's steady state with fluctuations around that steady-state. Phillips and Sul find some evidence of convergence towards steady states for OECD economies as well as US regions.

We close our discussion of panel data approaches by noting some unresolved issues in their application. It is important to be aware how panel data methods change the substantive interpretation of regression results, and care is needed when moving between the general forms of the estimators and the economic hypotheses under study. Relevant examples occur in analyses of β -convergence. If one finds β -convergence in a panel

study having allowed for fixed effects, the interpretation of this finding is very different than if one finds evidence of convergence in the absence of fixed effects. Specifically, the presence of fixed effects represents an immediate violation of our convergence definitions (20) or (22) as different economies must exhibit steady-state differences in per capita income regardless of whether they have identical saving rates and population growth rates.⁶¹ Fixed effects may even control for the presence of unmodelled determinants of steady state growth, an identification problem analogous to the one that was previously discussed in the context of interpreting the control variables Z in equations (17) and (18) above. Similarly, allowing for differences in time trends for per capita output, as done in Lee, Pesaran, and Smith (1997,1998) means that the finding of extremely rapid β -convergence is consistent with long-run divergence of per capita output across the economies they study; the long-run balanced growth paths are no longer parallel. In an interesting exchange, Lee, Pesaran, and Smith (1998) criticize Islam (1995) for failing to allow for different time trends across countries. In response, Islam (1998) argues that Lee, Pesaran, and Smith are assessing an economically uninteresting form of convergence when they allow for trend differences. This debate is an excellent example of the issues of interpretation that are raised in moving between specific economic hypotheses and more general statistical models.

One drawback of many current panel studies is that the construction of the time series observations can appear arbitrary. There is no inherent reason why 5 or 10 years represent natural spans over which to average observations. Similarly, there is arbitrariness with respect to which time periods are aggregated. A useful endeavor would be the development of tools to ensure that panel findings are robust with respect to the assumptions employed in creating the panel from the raw data.

More fundamentally, the empirical growth literature has not fully addressed the question of the appropriate time horizons over which growth models should be assessed. For example, it remains unclear when business cycle considerations (or instances of output collapses) may be safely ignored when modeling the growth process. While cross-section studies that examine growth over 30-40 year periods might be exempt from this

⁶¹ The impact of controlling fixed effects for interpreting β -convergence is recognized in the conclusion to Islam (1995).

consideration, it is less clear that panel studies employing 5-year averages are genuinely informative about medium-run growth dynamics.

iii. event study approaches

Although we have focused on the limitations of panel data methods, it is clear that the prospects for informative work of this kind should improve over time. The addition of further time periods is valuable in itself, and the history of developing countries in the 1980s and 1990s offers various events that introduce richer time series variation into the data. These events include waves of democratization, macroeconomic stabilization, financial liberalization, and trade liberalization, and panel data methods can be used to investigate their unfolding consequences for growth.

An alternative approach has become popular, and proceeds in a similar way to event studies in the empirical finance literature. In event studies, researchers look for systematic changes in asset returns after a discrete event, such as a profits warning. In fields outside finance, before-and-after studies like this have proved an informative way to gauge the effects of devaluations (see Pritchett (2000a) for references), of inflation stabilization (Easterly (1996)) and the consequences of the debt crisis for investment, as in Warner (1992).

Pritchett (2000a) argues that there is a great deal of scope for studying the growth impact of major events and policy changes in a similar way. The obvious approach is to study the time paths of variables such as output growth, investment and TFP growth, examined before and after such events. In empirical growth research, Henry (2000,2003) has applied this form of analysis to the effects of stock market liberalization on investment and growth, Giavazzi and Tabellini (2004) have considered economic and political liberalizations, while Wacziarg and Welch (2003) have studied the effects of trade liberalization. Depending on the context, one can also study the response of other variables in a way that is informative about the channels of influence. For example, in the case of trade liberalization, it is natural to study the response of the trade share, as in the work of Wacziarg and Welch.

The rigor of this method should not be overplayed. As with any other approach to empirical growth, one has to be cautious about inferring a causal effect. This is clear from exploring the analogy with treatment effects, a focus of recent research in microeconometrics and labor economics.⁶² In the study of growth, the treatments – such as democratization – are clearly not exogenously assigned, but are events that have arisen endogenously. Moreover, the treatment effects will be heterogeneous and could depend, for example, on whether a policy change is seen as temporary or permanent (Pritchett (2000a)). In these circumstances, the ability to quantify even an average treatment effect is strongly circumscribed. It may be possible to identify the direction of effects, and here the limited number of observations does have one advantage. With a small number of cases to examine, it is easy for the researcher to present a graphical analysis that allows readers to gauge the extent of heterogeneity in responses, and the overall pattern. At the very least, this offers a useful complement to regression-based methods.

iv. endogeneity and instrumental variables

A final set of data-based issues concerns the identification of instrumental variables in cross-section and time series contexts. An obvious and frequent criticism of growth regressions is that they do little to establish directions of causation. At one level, there is the standard problem that two variables may be correlated but jointly determined by a third. It is very easy to construct growth examples. Variables such as growth and political stability could be seen as jointly determined equilibrium outcomes associated with, say, a particular set of institutions. In this light, a correlation between growth and political stability, even if robust in statistical terms, does not appear especially informative about the structural determinants of growth.

There are many instances in growth research when explanatory variables are clearly endogenously determined (in the economic, not the statistical sense). The most familiar example would be a regression that relates growth to the ratio of investment to

⁶²This connection with the treatment effect literature is sometimes explicitly made, as in Giavazzi and Tabellini (2004) and Persson and Tabellini (2003). The connection helps to understand the limitations of the evidence, but the scope for resolving the associated identification problems may be limited in cross-country data sets.

GDP. This may tell us that the investment share and growth are associated, but stops short of identifying a causal effect. Even if we are confident that a change in investment would affect growth, in a sense this just pushes the relevant question further back, to an understanding of what determines investment.

When variables are endogenously determined in the economic sense, there is also a strong chance that they will be endogenous in the technical sense, namely correlated with the disturbances in the structural equation for growth. To give an example, consider what happens if political instability lowers growth, but slower economic growth feeds back into political instability. The estimated regression coefficient will tend to conflate these two effects and will be an inconsistent estimate of the causal effect of instability.⁶³

Views on the importance of these considerations differ greatly. One position is that the whole growth research project effectively capsizes before it has even begun, but Mankiw (1995) and Wacziarg (2002) have suggested an alternative view. According to them, one should accept that reliable causal statements are almost impossible to make, but use the partial correlations of the growth literature to rule out some possible hypotheses about the world. Wacziarg uses the example of the negative partial correlation between corruption and growth found by Mauro (1995). Even if shown to be robust, this correlation does not establish that somehow reducing corruption will be followed by higher growth rates. But it does make it harder to believe some of the earlier suggestions, rarely based on evidence, that corruption could be actively beneficial.

One approach is to model as many as possible of the variables that are endogenously determined. A leading example is Tavares and Wacziarg (2001), who estimate structural equations for various channels through which democracy could influence development. In their analysis, democracy affects growth via factors such as its effect on human capital accumulation, physical capital accumulation, inequality and government expenditures. They conclude the net effect of democracy on growth is

⁶³Although this ‘reverse causality’ interpretation of endogeneity is popular and important, it should be remembered that a correlation between an explanatory variable and the error term can arise for other reasons, including omitted variables and measurement error. As we discuss, it is important to bear this more general interpretation of the error term in mind when judging the plausibility of exclusion restrictions in instrumental variable procedures.

slightly negative, despite the positive contributions that are made from the role of democracy in promoting greater human capital and reduced inequality.

This approach has some important advantages in both economic and statistical terms. It can be informative about underlying mechanisms in a way that much empirical growth research is not. From a purely statistical perspective, if the structural equations are estimated jointly by methods such as three stage least squares or full information maximum likelihood, this is likely to bring efficiency gains. That said, systems estimation is not necessarily the best route: it has the important disadvantage that specification errors in one of the structural equations could contaminate the estimates obtained for the others.

The most common response to the endogeneity of growth determinants has been the application of instrumental variable procedures to a single structural equation, with growth as the dependent variable. As mentioned in Section IV, two growth studies that employ instrumental variables estimators based on lagged explanatory variables are Barro and Lee (1994) and Caselli, Esquivel, and Lefort (1996). Appendices 3 and 4 describe a wide range of other instrumental variables that have been proposed for the Solow variables and other growth determinants respectively, where the focus has been on the endogeneity of particular variables. The variety of instruments that have been proposed illustrates that it is relatively straightforward to find an instrument that is correlated with the endogenous explanatory variable(s).

This apparent success may be illusory. In our view, the belief that it is easy to identify valid instrumental variables in the growth context is deeply mistaken. We regard many applications of instrumental variable procedures in the empirical growth literature to be undermined by the failure to address properly the question of whether these instruments are valid, i.e. whether they may be plausibly argued to be uncorrelated with the error term in a growth regression. When the instrument is invalid, instrumental variables estimates will of course be inconsistent. Not enough is currently known about the consequences of “small” departures from validity, but it is certainly possible to envisage circumstances under which ordinary least squares would be preferable to instrumental variables on, say, a mean square error criterion.

A common misunderstanding, perhaps based on confusing the economic and statistical versions of “exogeneity”, is that predetermined variables, such as geographical characteristics, are inevitably strong candidates for instruments. There is, however, nothing in the predetermined nature of these variables to ensure either that they are not direct growth determinants or that they are uncorrelated with omitted growth determinants. Even if we take the extreme (from the perspective of being predetermined) example of geographic characteristics, there are many channels through which these could affect growth, and therefore many ways in which they could be correlated with the disturbances in a growth model. Brock and Durlauf (2001a) use this type of reasoning to make a very general critique of the use of instrumental variables in growth economics, basing it on the notion of theory open-endedness that we have described earlier. Since growth theories are mutually compatible, the validity of an instrument requires a positive argument that it cannot be a direct growth determinant or correlated with an omitted growth determinant. For many of the instrumental variables that have been proposed, this is clearly not the case.

Discussions of the validity of instruments inevitably suffer from some degree of imprecision because of the need to make qualitative and subjective judgments. When one researcher claims that it is implausible that a given instrument is valid, unless this claim is made on the basis of a joint model of the instruments and the variable of original interest, another researcher can always simply reject the assertion as unpersuasive. To be clear, this element of subjectivity does not mean that arguments about validity are pointless.⁶⁴ Rather, one must recognize that not all statistical questions can be adjudicated on the basis of mathematical analysis.

To see how different instruments might be assigned different levels of plausibility, we consider two examples. Brock and Durlauf (2001a) single out Frankel and Romer’s (1999) geographic instruments as an example where instrument validity

⁶⁴Put differently, one does not require a precise definition of what makes an instrument valid in order to argue whether a given instrument is valid or not. To take an example due to Taylor (1998), the absence of a precise definition of money does not weaken my belief that the currency in my wallet is a form of money, whereas the computer on which this paper is written is not. To claim such arguments cannot be made is known as the Socratic fallacy.

appears suspect as such variables are likely correlated with features of a country's economic, political, legal, and social institutions.⁶⁵ In our view, the large body of theoretical and empirical evidence on the role of institutions on growth, as well as even a cursory reading of history, renders the orthogonality assumptions required to use the instruments questionable.⁶⁶ For example, it is a standard historical claim that the fact that Great Britain is an island had important implications for its political development. While Frankel (2003) suggests that this worry is contrived, the argument against instrument validity flows quite naturally from modern growth theory and the many possible ways in which geographic characteristics such as remoteness could influence development.

As an example where instrument validity may be more plausible, consider Cook (2002a). He employs measures of damage caused by World War II as instruments for various growth regressors such as savings rates. The validity of Cook's instruments again relies on the orthogonality of World War II damage with omitted postwar growth determinants. It may be that levels of wartime damage had consequences for post-War growth performance in other respects (such as institutional change) but this argument is perhaps less straightforward than in the case of geographic characteristics.

To be clear, this discussion is nowhere near sufficient to conclude that Frankel and Romer's instruments are invalid whereas Cook's are valid. Rather our point is that conclusions concerning the relative plausibility of one set of instruments versus another need to rest on explicit arguments. It is not enough to appeal to a variable being predetermined, because this does not ensure that it is uncorrelated with the disturbances in the structural equation being estimated. A key implication of our discussion is that historical information has a vital role to play in facilitating formal growth analyses and evaluating exclusion restrictions.

⁶⁵While questions about the validity of instrumental variables arise in virtually all contexts, the force of these concerns differs across contexts. For example, in rational expectations models, lagged variables are natural instruments with respect to variables that, from the perspective of the theoretical model, are martingale differences, as occurs for excess holding returns. Objections to particular instruments in these contexts typically rely on alternative specifications of preferences or some other modification of the economic logic of the original model. This is quite different from the openness of growth theories.

⁶⁶The body of work on institutions and growth excellently summarized in Acemoglu, Johnson, and Robinson (2004) is strongly supportive of this claim.

This discussion of instrumental variables indicates another important, albeit neglected, issue in empirical growth analysis: the relationship between model specification and instrumental variable selection. One cannot discuss the validity of particular instruments independently from the choice of the specific growth determinants under study. An important outstanding research question is whether model uncertainty and instrumental variable selection can be integrated simultaneously into some of the methods we have described, including model averaging and automated model selection. The recent work of Hendry and Krolzig (2005) on automated methods includes an ambitious approach to systematic model selection for simultaneous equation models in which identifying restrictions are determined by the data.

VII. Econometric issues II: data and error properties

In this section we consider a range of questions that arise in growth econometrics from the properties of data and errors. Starting with data issues, Section VII.i examines how one may handle outliers in growth data. Section VII.ii examines the problem of measurement error. This is an important issue since there are good reasons to believe that the quality of the data is sometimes poor for less developed economies. In Section VII.iii we consider the case where data are not even measured, i.e. are missing. Turning to issues of the properties of model errors, Section VII.iv examines the analysis of heteroskedasticity in growth contexts. Finally, Section VII.v addresses the problem of cross-section correlation in model errors.

i. outliers

Empirical growth researchers often work with small data sets and estimate relatively simple models. In these circumstances, OLS regressions are almost meaningless unless they have been accompanied by systematic investigation of the data, including the sensitivity of the results to outlying observations.

There are various reasons why some observations may be unrepresentative. It is possible for variables to be measured with error for that particular region or country. Alternatively, the model specified by the researcher may omit a relevant consideration, and so a group of country observations will act as outliers. By construction, least squares estimates can be highly sensitive to the presence of small groups of observations. The practical implication is that OLS can give a misleading account of the patterns in the majority of the data. The dangers of using OLS were forcibly expressed by Swartz and Welsch (1986, p. 171): “In a world of fat-tailed or asymmetric error distributions, data errors, and imperfectly specified models, it is just those data in which we have the least faith that often exert the most influence on the OLS estimates”.

Some researchers respond to this concern using leverage measures or single-case diagnostics such as Cook’s distance statistic. There are well-known problems with these approaches, because where more than one outlier is present, its effect can easily be hidden by another (known in the statistics literature as “masking”). By far the best response is to use a more robust estimator, such as least trimmed squares, at least as a preliminary way of investigating the data.⁶⁷ These issues are discussed in more detail in Temple (1998,2000b).

ii. measurement error

We now turn to a more general discussion of measurement error. It is clear that measurement errors are likely to be pervasive, especially in data that relate to developing countries. Concepts that appear straightforward in economic models can present huge measurement problems in practice, as in the example of the capital stock discussed by Pritchett (2000b). Yet relatively few empirical studies of growth consider the impact of measurement error in any detail.

⁶⁷This estimator should not be confused with trimmed least squares, and other methods based on deleting observations with high residuals in the OLS estimates. A residual-based approach is inadequate for obvious reasons.

The best-known statistical result applies to a bivariate model where the independent variable is measured with error.⁶⁸ The estimate of the slope coefficient will be biased towards zero, even in large samples, because measurement error induces covariance between the observable form of the regressor and the error term. This attenuation bias is well known, but sometimes misleads researchers into suggesting that measurement error will only mask effects, a claim that is not true in general. When there are multiple explanatory variables, but only one is measured with error, then typically all the parameter estimates will be biased. Some parameter estimates may be biased away from zero and, although the direction of the bias can be estimated consistently, this is rarely done. When several variables are measured with error, the assumption that measurement error only hides effects is even less defensible.

Where measurement error is present, the coefficients are typically not identified unless other information is used. The most popular solution is to use instrumental variables, if an instrument can be found which is likely to be independent of the measurement error. A more complex solution is to exploit higher-order sample moments to construct more sophisticated estimators, as in Dagenais and Dagenais (1997). These procedures may be unreliable in small samples since the use of higher-order moments will make them especially sensitive to outliers.

Sometimes partial identification is possible, in the sense that bounds on the extent of measurement error can be used to derive consistent estimates of bounds on the slope parameters. Although it can be difficult for researchers to agree on sensible bounds on the measurement error variances, there are easier ways of formulating the necessary restrictions, as discussed by Klepper and Leamer (1984). Their reverse regression approach was implemented by Persson and Tabellini (1994) and Temple (1998), but has rarely been used by other researchers. Another strategy is to investigate sensitivity to varying degrees of measurement error, based on method-of-moments corrections. Again, this is easy to implement in linear models, and should be applied more routinely than it is at present. Temple (1998) provides a discussion of both approaches in the context of

⁶⁸This and the following discussion assume classical measurement error. Under more general assumptions, it is usually even harder to identify the consequences of measurement error for parameter estimates.

estimating technology parameters and the rate of conditional convergence within the Mankiw, Romer, and Weil (1992) model.

iii. missing data

Some countries never appear in growth data sets, partly by design: it is common to leave out countries with very small populations, oil producers, and transition economies. These are countries that seem especially unlikely to lie on a regression surface common to the majority of the OECD countries or the developing world. Countries with small populations should not be allowed to carry a great deal of weight in attempting to draw generalizations about growth for larger countries.

Other countries are left out for different reasons. When a nation experiences political chaos, or lacks economic resources, the collection of national accounts statistics will be a low priority. This means that countries like Afghanistan, Ethiopia and Somalia rarely appear in comparative growth studies. In other cases, countries appear in some studies but not in others, depending on the availability of particular variables of interest.

Missing data are of course a potentially serious problem. If one started from a representative data set and then deleted countries at random, this would typically increase the standard errors but not lead to biased estimates. More serious difficulties arise if countries are missing in a systematic way, because then parameter estimates are likely to be biased. This problem is given relatively little attention in mainstream econometrics textbooks, despite a large body of research in the statistics literature.

A variety of solutions are possible, with the simplest being one form or another of imputation, with an appropriate adjustment to the standard errors. Hall and Jones (1999) and Hoover and Perez (2004) are among the few empirical growth studies to carry out imputation in a careful and systematic way. This approach may be especially useful when countries are missing from a data set because a few variables are not observed for their particular cases. It is then easy to justify using other available information to predict the missing data, and thereby exploit the additional information in the variables that are observed. Alternative approaches to missing data are also available, based on likelihood or Bayesian methods, which can be extended to handle missing observations.

iv. heteroskedasticity

It is common in cross-section regressions for the underlying disturbances to have a non-constant variance. As is well known, the coefficient estimates remain unbiased, but OLS is inefficient and the estimates of the standard errors are biased. Most empirical growth research simply uses the heteroskedasticity-consistent standard errors developed by Eicker (1967) and White (1980). These estimates of the standard errors are consistent but not unbiased, which suggests that alternative solutions to the problem may be desirable. For data sets of the size found in cross-country empirical work, the alternative estimators developed by MacKinnon and White (1985) are likely to have better finite sample properties, as discussed in Davidson and MacKinnon (1993) and supported by simulations in Long and Ervin (2000).

There are at least two other concerns with the routine application of White's heteroskedasticity correction as the only response to heteroskedasticity. The first is that by exploiting any structure in the variance of the disturbances, using weighted least squares, it may be possible to obtain efficiency gains. The second and more fundamental objection is that heteroskedasticity can often arise from serious model misspecification, such as omitted variables or neglected parameter heterogeneity. Evidence of heteroskedasticity should then prompt revisions of the model for the conditional mean, rather than mechanical adjustments to the standard errors. See Zietz (2001) for further discussion and references.

v. cross-section error correlation

An unresolved issue in growth econometrics is the treatment of cross-section correlation in model errors. Such correlation may have important consequences for inference; as noted by DeLong and Summers (1991) in the growth context, failure to account for cross-sectional dependence can lead to incorrect calculation of standard errors and hence, incorrect inferences. One would certainly expect cross-sectional dependence to be present when studying growth. For example, countries that are geographically close together, or trading partners, may experience common shocks.

Whether this effect is sizeable remains an open question, but one that might be addressed using ideas developed in Baltagi et al. (2003) and Driscoll and Kraay (1998), among others. In the context of growth regressions, work on cross-section dependence may be divided into two lines. One direction concerns the identification of the presence of cross-section dependence. Pesaran (2004a) develops tests for cross-section dependence that do not rely on any prior ordering; this framework in essence sums the cross-section sample error correlations in a panel and evaluates whether they are consistent with the null hypothesis that the population correlations are zero. Specifically, he proposes (recalling that N denotes the cross-section dimension and T the time dimension) a cross-section dependence statistic CD

$$CD = \sqrt{\frac{2T}{N(N-1)}} \left(\sum_{i=1}^{N-1} \sum_{j=i+1}^N \hat{\rho}_{i,j} \right) \quad (66)$$

where $\hat{\rho}_{i,j}$ is the sample correlation between $\varepsilon_{i,t}$ and $\varepsilon_{j,t}$; Pesaran gives conditions under which this statistic converges to a Normal (0,1) random variable (as N and T become infinite) under the null hypothesis of no cross-section correlation. This test statistic is based on earlier work by Breusch and Pagan (1980) and appears to possess good finite sample properties in comparison to this earlier work. Using a country-level panel, Pesaran (2004a) finds strong rejections of the null of no cross-section correlation both for the world as a whole as well as within several geographic groupings.

The second and primary direction for the analysis of cross-section correlation has been concerned not so much with testing for its presence, but rather accounting for its presence in growth exercises. One approach relies on formulating a statistical model of the dependence. Phillips and Sul (2003) model the residuals in a growth panel as

$$\varepsilon_{i,t} = \delta_i \theta_t + u_{i,t} \quad (67)$$

where θ_t and $u_{i,t}$ are independent random variables; $u_{i,t}$ is assumed to be i.i.d. across countries and across time. Phillips and Sul (2002) describe the properties of panel estimators under this assumption.

Another possibility in analyzing cross-sectional dependence is to treat the problem as one of spatial correlation in errors. The problem of spatial correlation has been much studied in the regional science literature, and statisticians in this field have developed spatial analogues of many time series concepts, see Anselin (2001) for an overview. Spatial methods have, in our view, an important role to play in growth econometrics. However, when these methods are adapted from the spatial statistics literature, they raise the problem of identifying the appropriate notion of space. One can imagine many reasons for cross-section correlation. If one is interested in technological spillovers, it may well be the case that in the space of technological proximity, the United Kingdom is closer to the United States than is Mexico. Put differently, unlike the time series and spatial cases, there is no natural cross-section ordering to elements in the standard growth data sets. Following language due to Akerlof (1997) countries are perhaps best thought of as occupying some general socio-economic-political space defined by a range of factors; if one could identify their locations, then spatial methods could be implemented.

An interesting approach to addressing the relevant spatial location of countries is pursued by Conley and Ligon (2002). In their analysis, they attempt to construct estimates of the spatial covariation of the residuals ε_i in a cross-section. In order to do this, they construct different measures of socioeconomic distance between countries. They separately consider geographic distance (measured between capital cities), as well as measures of the costs of transportation between these cities. Once a distance metric is constructed, these are used to construct a residual covariance matrix. Estimation methods for this procedure are developed in Conley (1999). Conley and Ligon (2002) find that allowing for cross-section dependence in this way is relatively unimportant in terms of appropriate calculation of standard errors for growth model parameters. Their methods could be extended to allow for comparisons of different variables as the source for cross-section correlation as is done in Conley and Topa (2002) in the context of residential neighborhoods. A valuable generalization of this work would be the modeling of cross-

section correlations as a function of multiple variables. Such an analysis would make further progress on the measurement of distances in socioeconomic space, which, as we have suggested, presumably are determined by multiple channels.

A generally unexplored possibility for studying cross-section dependence in growth (and other contexts) is to model these correlations structurally as the outcome of spillover effects.⁶⁹ The theoretical literature on social interactions studies cross-sectional dependences in precisely this way (see Brock and Durlauf (2001b) for a survey of this literature). While such models have the potential for providing firm microfoundations for cross-section dependence, the presence of such spillovers has consequences for identification that are not easily resolved (Brock and Durlauf (2001b), Manski (1993)) and which have yet to be explored in growth contexts; Binder and Pesaran (2001) and Brock and Durlauf (2001b) analyze identification and estimation problems for intertemporal environments that are particularly germane to growth contexts.

VIII. Conclusions: the future of growth econometrics

In this section, we offer some closing thoughts on the most promising directions for empirical growth research. We are not the first authors to set out manifestoes for the field, and we explicitly draw on previous contributions, many of which deserve wider currency. It is also interesting to compare the current state of the field against the verdicts offered in the early survey by Levine and Renelt (1991). One dominant theme will be that the empirical study of growth requires an eclectic approach, and that the field has been harmed by a tendency for research areas to evolve independently, without enough interaction.⁷⁰ This is not simply a question of using a variety of techniques: it also means that there needs to be a closer connection between theory and evidence, a willingness to

⁶⁹An exception is Easterly and Levine (1997b).

⁷⁰To give a specific example, the macroeconomic literature on international technology differences only rarely acknowledges relevant work by trade economists, including estimates of the Heckscher-Ohlin-Vanek model that suggest an important role for technology differences. See Klenow and Rodriguez-Clare (1997b) for more discussion.

draw on ideas from areas such as trade theory, and more attention to particular features of the countries under study.

We start with Pritchett (2000a), who lists three questions for growth researchers to address:

- What are the conditions that initiate an acceleration of growth or the conditions that set off sustained decline?
- What happens to growth when policies – trade, macroeconomic, investment – or politics change dramatically in episodes of reform?
- Why have some countries absorbed and overcome shocks with little impact on growth, while others seem to have been overwhelmed by adverse shocks?

This agenda seems to us very appropriate, not least because it focuses attention on substantive economic issues rather than the finer points of estimating aggregate production functions. The importance of the first of Pritchett's questions is evident from the many instances where countries have moved from stagnation to growth and vice versa. A paper by Hausmann, Pritchett, and Rodrik (2004) explicitly models transitions to fast growth ("accelerations") and makes clear the scope for informative work of this kind. The second question we have discussed above, and research in this vein is becoming prominent, as in Henry (2000, 2003), Giavazzi and Tabellini (2004), and Wacziarg and Welch (2003). Here, one of the major challenges will be to relax the (sometimes only implicit) assumption that policies are randomly assigned. Finally, an important paper by Rodrik (1999) has addressed the third question, namely what determines varying responses to major shocks.

In all three cases, it is clear that econometric work should be informed by detailed studies of individual countries, such as those collected in Rodrik (2003). Too much empirical growth research proceeds without enough attention to the historical and institutional context. For example, a newcomer to this literature might be surprised at the paucity of work that integrates growth regression findings with, say, the known consequences of the 1980s debt crisis.

Another reason for advocating case studies is that much of the empirical growth literature essentially points only to reduced-form partial correlations. These can be useful, but it is clear that we often need to move beyond this. A partial correlation is more persuasive if it can be supported by theoretical arguments. The two combined are more persuasive if there is evidence of the intermediating effects or mechanisms that are emphasized in the relevant theory. There is plenty of scope for informative work that tries to isolate mechanisms by which variables such as financial depth, inequality, and political institutions shape the growth process. Wacziarg (2002), in particular, highlights the need for a structural growth econometrics, one that aims to recover channels of causation, and hence supports (or undermines) the economic significance of the partial correlations identified in the literature.

A more extreme view is that growth econometrics should be supplanted by the calibration of theoretical models. Klenow and Rodriguez-Clare (1997b) emphasize the potential of such an approach and note that Mankiw, Romer, and Weil's (1992) influential analysis can be seen partly as a comparison of estimated parameter values with those associated with specific theoretical models. Relatively little of the empirical work that has followed has achieved a similarly close connection between theory and evidence, and this has been a recurring criticism of the literature (for example, Levine and Renelt (1991) and Durlauf (2001)).

It may be premature to say that econometric approaches should be entirely replaced by calibration exercises, but the two methods could surely inform each other more often than at present. Calibrated models can help to interpret parameter estimates, not least in comparing the magnitude of the estimates with the implications of plausible models. Klenow and Rodriguez-Clare (1997b) discuss examples of this in more detail. At the same time, the partial correlations identified in growth econometrics can help to act as a discipline on model-building and can indicate where model-based quantitative investigations are most likely to be fruitful. This role for growth econometrics is likely to be especially useful in areas where the microeconomic evidence used to calibrate structural models is relatively weak, or the standard behavioral assumptions may be flawed.

The need for a tighter connection between theory and evidence is especially apparent in certain areas. The workhorse model for many empirical growth papers continues to be Solow-Swan, a closed economy model which leaves out aspects of interdependence that are surely important. Howitt (2000) has shown that growth regression evidence can be usefully reinterpreted in the light of a multi-country theoretical model with a role for technology diffusion. More generally, there is a need for researchers to develop empirical growth frameworks that acknowledge openness to flows of goods, capital and knowledge. These issues are partly addressed by the theoretical analysis of Barro, Mankiw, and Sala-i-Martin (1995) and empirical work that builds on such ideas deserves greater prominence. Here especially, research that draws on the quantitative implications of specific models, as in the work of Eaton and Kortum (1999, 2001) on technology diffusion and the role of imported capital goods, appears to be an important advance.

The neglect of open economy aspects of the countries under study is mirrored elsewhere. Much of the empirical literature uses a theoretical framework that was originally developed to explain the growth experiences of the USA and other developed nations. Yet this framework is routinely applied to study developing countries, and there appears plenty of scope for models that incorporate more of the distinctive features of poorer countries. These could include the potentially important roles of agricultural employment, dualism, and structural change, and in some cases, extensive state involvement in production. This is an area in which empirical growth researchers have really only scratched the surface.

Some of these issues are connected to an important current research agenda, namely the need to distinguish between different types of growth and their distributional consequences. For example, the general equilibrium effects of productivity improvements in agriculture may be very different to those in services and industry. Identifying the nature of “pro-poor” growth will require more detailed attention to particular features of developing countries. Given that the main source of income for the poor is usually labour income, growth researchers will need to integrate their models with theory and evidence from labour economics, in order to study how growth and labour markets interact.

Agénor (2004) considers some of the relevant issues, and again this appears to be a vital direction for future research.

Ideally, research along these various lines will utilize not only statistics, but also the power of case studies in generating hypotheses, and in deepening our understanding of the economic, social and political forces at work in determining growth outcomes. Case studies may be especially valuable in two areas. The first of these is the study of technology transfer. As emphasized in the survey by Klenow and Rodriguez-Clare (1997b), we do not know enough about why some countries are more successful than others in climbing the “ladder” of product quality and technological complexity. What are the relative contributions of human capital, foreign direct investment and trade? In recent years some of these issues have been intensively studied at the microeconomic level, especially the role of foreign direct investment and trade, but there remains work to be done in mapping firm and sector-level evidence into a set of aggregate implications.

A second area in which case studies are likely to prove valuable is the study of political economy, in its modern sense. It is a truism that economists, particularly those considering development, have become more aware of the need to account for the two-way interaction between economics and politics. A case can be made that the theoretical literature has outpaced the empirical literature in this regard. Studies of individual countries, drawing on both economic theory and political science, would help to close this gap.

Thus far, we have highlighted a number of limitations of existing work, and directions in which further research seems especially valuable. Some of the issues we have considered were highlighted much earlier by Levine and Renelt (1991), and that might lead to pessimism over the long-term prospects of this literature.⁷¹ This also shows that our prescriptions for future research could seem rather pious, since the improvements we recommend are easier said than done. We end our review by considering some areas in which genuine progress has been made, and where further progress appears likely.

One reason for optimism is the potential that recently developed model averaging

⁷¹Only now are researchers beginning to engage with some of the issues they raised, such as the varying conditions under which it is appropriate to use international rather than national prices in making productivity comparisons and constructing capital stocks.

methods have for shedding new light on growth questions. These methods help to address the model selection and robustness issues that have long been identified as a major weakness of cross-country growth research. By framing the problem explicitly in terms of model uncertainty, in the way envisaged by Leamer (1978), it is possible to consider many candidate explanatory variables simultaneously, and identify which effects appear to be systematic features of the data, as reflected in posterior probabilities of inclusion. The Bayesian approach to model averaging also provides an index of model adequacy, the posterior model probability, that is easy to interpret, and that allows researchers to gauge the extent of overall model uncertainty. Above all, researchers can communicate the degree of support for a particular hypothesis with more faith that the results do not depend on an arbitrary choice of regression specification. Although the application of Bayesian model averaging inevitably has limitations of its own, it appears more rigorous than many of the alternatives, and we expect a number of familiar growth questions to be revisited using these methods.

Another reason for optimism is that the quality of available data is likely to improve over time. The development of new and better data has clearly been one of the main achievements of the empirical growth literature since the early 1990s, and one that was not foreseen by critics of the field. Researchers have developed increasingly sophisticated proxies for drivers of growth that appeared resistant to statistical analysis. One approach, pioneered in the growth literature by Knack and Keefer (1995) and Mauro (1995), has been country-specific ratings compiled by international agencies. Such data increasingly form the basis for measures of corruption, government efficiency, and protection of property rights. More recent work such as that of Kaufmann et al. (1999a,1999b,2003) has established unusually comprehensive measures of various aspects of institutions.

The construction of proxies is likely to make increasing use of latent variable methods. These aim to reduce a set of observed variables to a smaller number of indicators that are seen as driving the majority of the variation in the original data, and that could represent some underlying variable of interest. For example, the extent of democracy is not directly observed, but is often obtained by applying factor analysis or extracting principal components from various dimensions of political freedom. There are

obvious dangers with this approach, but the results can be effective proxies for concepts that are otherwise hard to measure.⁷² They also help to overcome the dimensionality problem associated with cross-country empirical work. To be successfully employed, the rigorous use of a latent variable as a regressor will generally need to acknowledge the presence of measurement error.⁷³

Using latent variables makes especially good sense under one view of the proper aims of growth research. It is possible to argue that empirical growth studies will never give good answers to precise hypotheses, but can be informative at a broader level. For example, a growth regression is unlikely to tell us whether the growth effect of inflation is more important than the effect of inflation uncertainty, because these two variables are usually highly correlated. It may even be difficult to distinguish the effects of inflation from the effects of sizeable budget deficits.⁷⁴ Instead a growth regression might be used to address a less precise hypothesis, such as the growth dividend of macroeconomic stability, broadly conceived. In this context, it is natural to use latent variable approaches to measure the broader concept.

Another valuable development is likely to be the creation of rich panel data sets at the level of regions within countries. Regional data offer greater scope for controlling for some variables that are hard to measure at the country level, such as cultural factors. By comparing experiences across regions, there may also be scope for identifying events that correspond more closely to natural experiments than those found in cross-country data. Work such as that by Besley and Burgess (2000,2002,2004) using panel data on Indian

⁷²A relevant question, not often asked, is how high the correlation between the proxy and the true predictor has to be for the estimated regression coefficient on the proxy to be of the “true” sign. Krasker and Pratt (1986,1987) have developed methods that can be used to establish this under surprisingly general assumptions.

⁷³In principle this can be addressed by structural equation modeling, using software like EQS or LISREL to estimate a system of equations that includes explicit models for latent variables, an approach used elsewhere in the social sciences. Most economists are not familiar with this approach, and this makes the assumptions and results hard to communicate. It is therefore not clear that a full latent variable model should be preferred to a simpler solution, such as one of those we discuss in the measurement error section above.

⁷⁴As Sala-i-Martin (1991) has argued, various specific indicators of macroeconomic instability should perhaps be seen as symptoms of some deeper, underlying characteristic of a country.

states shows the potential of such an approach. In working with such data more closely, one of the main challenges will be to develop empirical frameworks that incorporate movements of capital and labour between regions: clearly, regions within countries should only rarely be treated as closed economies. Shioji (2001b) is an example of how analysis using regional data can take this into account.

Even with better data, at finer levels of disaggregation, the problem of omitted variables can only be alleviated, not resolved. It is possible to argue that the problem applies equally to historical research and case studies, but at least in these instances, the researcher may have some grasp of important forces that are difficult to quantify. Since growth researchers naturally gravitate towards determinants of growth that can be analyzed statistically, there is an ever-present danger that the empirical literature, even taken as a whole, yields a rather partial and unbalanced picture of the forces that truly matter. Even a growth model with high explanatory power, in a statistical sense, has to be seen as a rather provisional set of ideas about the forces that drive growth and development.

This brings us to our final points. We once again emphasize that empirical progress on the major growth questions requires attention to the evidence found in qualitative sources such as historical narratives and studies by country experts. One example we have given in the text concerns the validity of instrumental variables: understanding the historical experiences of various countries seems critical for determining whether exclusion restrictions are plausible. In this regard work such as that of Acemoglu, Johnson, and Robinson (2001,2002) is exemplary. More generally, nothing in the empirical growth literature suggests that issues of long-term development can be disassociated from the historical and cultural factors that fascinated commentators such as Max Weber. Where researchers have revisited these issues, as in Barro and McCleary (2003), the originality resides less in the conception of growth determinants and more in the scope for new statistical evidence. Of course, the use of historical analysis also leads back to the value of case studies, a point that has recurred throughout this discussion.

In conclusion, growth econometrics is an area of research that is still in its infancy. To its credit, the field has evolved in response to the substantive economic questions that arise in growth contexts. The nature of the field has also led

econometricians to introduce a number of statistical methods into economics, including classification and regression tree algorithms, robust estimation, threshold models and Bayesian model averaging, that appear to have wide utility. As with any new literature, especially one tackling questions as complex as these, it is possible to identify significant limitations of the existing evidence and the tools that are currently applied. But the progress that has been made in growth econometrics in the brief time since its emergence gives reason for continued optimism.

Appendix 1: Data

Key to the 102 countries

AGO, Angola, ARG, Argentina, AUS, Australia, AUT, Austria, BDI, Burundi, BEL, Belgium, BEN, Benin, BFA, Burkina Faso, BGD, Bangladesh, BOL, Bolivia, BRA, Brazil, BWA, Botswana, CAF, Central African Republic, CAN, Canada, CHE, Switzerland, CHL, Chile, CHN, China, CIV, Cote d'Ivoire, CMR, Cameroon, COG, Rep. of Congo, COL, Colombia, CRI, Costa Rica, CYP, Cyprus, DNK, Denmark, DOM, Dominican Republic, ECU, Ecuador, EGY, Egypt, ESP, Spain, ETH, Ethiopia, FIN, Finland, FJI, Fiji, FRA, France, GAB, Gabon, GBR, United Kingdom, GHA, Ghana, GIN, Guinea, GMB, The Gambia, GNB, Guinea-Bissau, GRC, Greece, GTM, Guatemala, GUY, Guyana, HKG, Hong Kong, HND, Honduras, IDN, Indonesia, IND, India, IRL, Ireland, IRN, Iran, ISR, Israel, ITA, Italy, JAM, Jamaica, JOR, Jordan, JPN, Japan, KEN, Kenya, KOR, Rep. of Korea, LKA, Sri Lanka, LSO, Lesotho, MAR, Morocco, MDG, Madagascar, MEX, Mexico, MLI, Mali, MOZ, Mozambique, MRT, Mauritania, MUS, Mauritius, MWI, Malawi, MYS, Malaysia, NAM, Namibia, NER, Niger, NGA, Nigeria, NIC, Nicaragua, NLD, Netherlands, NOR, Norway, NPL, Nepal, NZL, New Zealand, PAK, Pakistan, PAN, Panama, PER, Peru, PHL, Philippines, PNG, Papua New Guinea, PRT, Portugal, PRY, Paraguay, ROM, Romania, RWA, Rwanda, SEN, Senegal, SGP, Singapore, SLV, El Salvador, SWE, Sweden, SYR, Syria, TCD, Chad, TGO, Togo, THA, Thailand, TTO, Trinidad & Tobago, TUR, Turkey, TWN, Taiwan, TZA, Tanzania, UGA, Uganda, URY, Uruguay, USA, USA, VEN, Venezuela, ZAF, South Africa, ZAR, Dem. Rep. Congo, ZMB, Zambia, ZWE, Zimbabwe

Extrapolation

Where data on GDP per worker for the year 2000 are missing from PWT 6.1, but are available for 1996 or after, we extrapolate using the growth rate between 1990 and the latest available year. This procedure helps to alleviate the biases that can occur when countries are missing from the sample for systematic reasons, such as political or economic collapse.

The countries involved are Angola (extrapolated from 1990-1996), Botswana (1999), Central African Republic (1998), Democratic Republic of Congo (1997), Cyprus (1996), Fiji (1999), Guyana (1999), Mauritania (1999), Namibia (1999), Papua New Guinea (1999), Singapore (1996), and Taiwan (1998).

Appendix 2: Variables in Cross-Country Growth Regressions

+/- = sign of coefficient in the corresponding growth regression

? = sign not reported

* = claimed to be significant

_ = claimed to be insignificant

R.H.S. Variables		Studies
Capitalism		<ul style="list-style-type: none"> • Hall and Jones (1999) (+,*)
Capital account liberalization		<ul style="list-style-type: none"> • Eichengreen and Leblang (2003) (+,*)
Corruption		<ul style="list-style-type: none"> • Mauro (1995) (-,*) • Welsch (2003) (-,*)
Democracy	Minimum levels	<ul style="list-style-type: none"> • Barro (1996) (1997) (+,*)
	...Higher levels	<ul style="list-style-type: none"> • Barro (1996) (1997) (-,*)
	Overall	<ul style="list-style-type: none"> • Alesina et al. (1996) (?,_) • Minier (1998) (+,*)
	‘Voice’	<ul style="list-style-type: none"> • Dollar and Kraay (2003) (-,*)
Demographic Characteristics	Share of Population 15 or below	<ul style="list-style-type: none"> • Barro and Lee (1994) (-,*)
	Share of Population 65 or over	<ul style="list-style-type: none"> • Barro and Lee (1994) (?,_)
	Growth of 15-65 population share	<ul style="list-style-type: none"> • Bloom and Sachs (1998) (+,*)
Education	College Level	<ul style="list-style-type: none"> • Barro and Lee (1994) (-,_)
	Female (level)	<ul style="list-style-type: none"> • Barro and Lee (1994) (-,*) • Barro (1996) (1997) (-,*) • Caselli, et al. (1996) (+,*) • Forbes (2000) (-,*)
	Female (growth)	<ul style="list-style-type: none"> • Barro and Lee (1994) (-,*)
	Male (level)	<ul style="list-style-type: none"> • Barro and Lee (1994) (+,*) • Barro (1996) (+,*) • Caselli, et al. (1996) (-,*) • Forbes (2000) (+,*)
	Male (growth)	<ul style="list-style-type: none"> • Barro and Lee (1994) (+,*)

	Overall (level)	<ul style="list-style-type: none"> • Azariadis and Drazen (1990) (+,*) • Barro (1991) (+,*) • Knowles and Owen (1995) (+,_) • Easterly and Levine (1997a) (+,*) • Krueger and Lindahl (2000) (+,*) • Bils and Klenow (2000) (+,*)
	Primary Level	<ul style="list-style-type: none"> • Sachs and Warner (1995) (+,_) • Barro (1997) (-,_)
	Secondary Level	<ul style="list-style-type: none"> • Sachs and Warner (1995) (+,_)
	Initial Income * Male Schooling	<ul style="list-style-type: none"> • Barro (1997) (-,*)
	Proportion of Engineering Students	<ul style="list-style-type: none"> • Murphy, et al. (1991) (+,*)
	Proportion of Law Students	<ul style="list-style-type: none"> • Murphy, et al. (1991) (-,*)
Ethnicity and Language	Ethno-Linguistic Fractionalization	<ul style="list-style-type: none"> • Easterly and Levine (1997a) (-,*) • Sala-i-Martin (1997a,b) (?,_) • Alesina, et al. (2003) (-,*)
	Language Diversity	<ul style="list-style-type: none"> • Masters and McMillan (2001) (-,*/_)
Fertility		<ul style="list-style-type: none"> • Barro (1991) (1996) (1997) (-,*) • Barro and Lee (1994) (-,*)
Finance	Stock Markets	<ul style="list-style-type: none"> • Levine and Zervos (1998) (+,*) • Beckaert, et al. (2001) (+,*) • Beck and Levine (2004) (+,*)
	Banks	<ul style="list-style-type: none"> • Beck and Levine (2004) (+,*)
	Dollarization	<ul style="list-style-type: none"> • Edwards and Magendzo (2003) (+,_)
	Depth	<ul style="list-style-type: none"> • Berthelemy and Varoudakis (1995) (+,*) • Odedokun (1996) (+,*) • Ram (1999) (+,_) • Rousseau and Sylla (2001) (+,*) • Deidda and Fattouh (2002) (+,_)

		<ul style="list-style-type: none"> Demetriades and Law (2004) (+,*)
	Competition*development	<ul style="list-style-type: none"> Claessens and Laeven (2003) (+,*)
	Repression	<ul style="list-style-type: none"> Roubini and Sala-i-Martin (1992) (-,*) Easterly (1993) (-,*)
	Sophistication	<ul style="list-style-type: none"> King and Levine (1993) (+,*) Levine and Zervos (1993) (+,robust) Easterly and Levine (1997a) (+,*) Sala-i-Martin (1997a,b) (?,_)
	Credit	<ul style="list-style-type: none"> Levine and Renelt (1992) (+,not robust) De Gregorio and Guidotti (1995) (+,*)
	Growth rate	
	Volatility	<ul style="list-style-type: none"> Levine and Renelt (1992) (+,not robust)
Foreign Direct Investment		<ul style="list-style-type: none"> Blonigen and Wang (2004) (+,_)
Fraction of mining in GDP		<ul style="list-style-type: none"> Hall and Jones (1999) (+,*)
Geography	Absolute Latitude	<ul style="list-style-type: none"> Sala-i-Martin (1997a,b) (+,*) Bloom and Sachs (1998) (+,*) Masters and McMillan (2001) (-,_) Easterly and Levine (2001) (+,*) Rodrik et al. (2004) (+,*)
	Disease Ecology	<ul style="list-style-type: none"> McCarthy, et al. (2000) (+,*) McArthur and Sachs (2001) (+,*) Easterly and Levine (2002) (-,*) Sachs (2003) (-,*)
	Frost days	<ul style="list-style-type: none"> Masters and McMillan (2001)(+,*) Masters and Sachs (2001) (+,*)

	Land locked	<ul style="list-style-type: none"> Easterly and Levine (2001) (-,*)
	Coastline (length)	<ul style="list-style-type: none"> Bloom and Sachs (1998) (+,*) Masters and Sachs (2001) (+,*) Bloom, et al. (2003) (+,*)
	Arable land	<ul style="list-style-type: none"> Masters and Sachs (2001) (+,*)
	Rainfall	<ul style="list-style-type: none"> Masters and Sachs (2001) (+,*) Bloom, et al. (2003) (+,*)
	Variance of Rainfall	<ul style="list-style-type: none"> Bloom, et al. (2003) (-,*)
	Maximum Temperature	<ul style="list-style-type: none"> Bloom, et al. (2003) (-,*)
Government	Consumption (growth)	<ul style="list-style-type: none"> Kormendi and Meguire (1985) (+,_)
	Consumption (level)	<ul style="list-style-type: none"> Barro (1991) (-,*) Sachs and Warner (1995) (-,*) Barro (1996) (-,*) Caselli, et al. (1996) (+,*) Barro (1997) (-,*) Acemoglu, et al. (2002) (-,_)
	Deficits	<ul style="list-style-type: none"> Levine and Renelt (1992) (-,not robust) Fischer (1993) (-,*) Nelson and Singh (1994) (+,_) Easterly and Levine (1997a) (-,*) Bloom and Sachs (1998) (+,*)
	Investment	<ul style="list-style-type: none"> Barro (1991) (+,_) Sala-i-Martin (1997a,b) (?,_) Kelly (1997) (+,*)
	Various Expenditures	<ul style="list-style-type: none"> Levine and Renelt (1992) (-,not robust)
	Military Expenditures	<ul style="list-style-type: none"> Aizenman and Glick (2003) (-,*) Guaresma and Reitschuler (2003) (-,*)
	Military Expenditures under threat	<ul style="list-style-type: none"> Aizenman and Glick (2003) (+,*)
	Various Taxes	<ul style="list-style-type: none"> Levine and Renelt (1992) (?,not

		robust)
Growth Rate	of the G-7 Countries	<ul style="list-style-type: none"> • Alesina, Ozler, Roubini, and Swagel (1996) (+,*)
	in the Previous Period	<ul style="list-style-type: none"> • Easterly, et al. (1993) (+,_) • Alesina, et al. (1996) (+,*/_)
Health	Life expectancy	<ul style="list-style-type: none"> • Barro and Lee (1994) (+,*) • Bloom and Malaney (1998) (+,*) • Bloom and Sachs (1998) (+,*) • Bloom and Williamson (1998) (+,*) • Hamoudi and Sachs (1999) (+,*) • Gallup et al. (2000) (+,*)
	Change in Malaria Infection Rate	<ul style="list-style-type: none"> • Gallup, Mellinger and Sachs (2000).
	Adult Survival Rate	<ul style="list-style-type: none"> • Bhargava et al. (2001)
Industrial Structure	% Small and Medium Enterprises	<ul style="list-style-type: none"> • Beck, et al. (2003) (+,_)
	Ease of entry and exit	<ul style="list-style-type: none"> • Beck, et al. (2003) (+,*)
Inequality	Democratic Countries	<ul style="list-style-type: none"> • Persson and Tabellini (1994) (-,*)
	Non-Democratic Countries	<ul style="list-style-type: none"> • Persson and Tabellini (1994) (+,_)
	Overall	<ul style="list-style-type: none"> • Alesina and Rodrik (1994) (-,*) • Forbes (2000) (+,*) • Knowles (2001) (-,*)
Inflation	Growth	<ul style="list-style-type: none"> • Kormendi and Meguire (1985) (-,*)
	Level	<ul style="list-style-type: none"> • Levine and Renelt (1992) (-,not robust) • Levine and Zervos (1993) (? ,not robust) • Barro (1997) (-,*) (in the range above 15%) • Bruno and Easterly (1998) (-,*) • Motley (1998) (-,*)

		<ul style="list-style-type: none"> • Li and Zou (2002) (-,*)
	Variability	<ul style="list-style-type: none"> • Levine and Renelt (1992) (-,not robust) • Fischer (1993) (-,*) • Barro (1997) (+,_) • Sala-i-Martin (1997a,b) (?,_)
Infrastructure Proxies		<ul style="list-style-type: none"> • Hulten (1996) (+,*) • Easterly and Levine (1997a) (+,*) • Esfahani and Ramirez (2003) (+,*)
Initial Income		<ul style="list-style-type: none"> • Kormendi and Meguire (1985) (-,*) • Barro (1991) (-,*) • Sachs and Warner (1995) (-,*) • Harrison (1996) (?,_) • Barro (1997) (-,*) • Easterly and Levine (1997a)
Investment Ratio		<ul style="list-style-type: none"> • Barro (1991) (+,*) • Barro and Lee (1994) (+,*) • Sachs and Warner (1995) (+,*) • Barro (1996) (+,_) • Caselli, et al. (1996) (+,*) • Barro (1997) (+,_)
Investment Type	Equipment or Fixed Capital	<ul style="list-style-type: none"> • DeLong and Summers (1993) (+,*) • Blomstrom, et al. (1996) (-,_) • Sala-i-Martin (1997a,b) (+,*)
	Non-Equipment	<ul style="list-style-type: none"> • DeLong and Summers (1991) (+,*)
Labor	Productivity Growth	<ul style="list-style-type: none"> • Lichtenberg (1992) (+,*)
	Productivity Quality	<ul style="list-style-type: none"> • Hanushek and Kimko (2000) (+,*)
	Labor Force Part. Rate	<ul style="list-style-type: none"> • Blomstrom, et al. (1996) (+,*)
Luck	External Debt Dummy	<ul style="list-style-type: none"> • Easterly, et al. (1993) (-,_)
	External Transfers	<ul style="list-style-type: none"> • Easterly, et al. (1993) (mixed,_)

	Improvement in Terms of Trade	<ul style="list-style-type: none"> • Easterly, et al. (1993) (+,*) • Fischer (1993) (+,*) • Barro (1996) (+,*) • Caselli, et al. (1996) (+,*) • Barro (1997) (+,*) • Blattman, et al. (2003) (+,*) (
Money Growth		<ul style="list-style-type: none"> • Kormendi and Meguire (1985) (+,_)
Neighboring Countries' Education Proxies, Initial Incomes, Investment Ratios and Population Growth Rates		<ul style="list-style-type: none"> • Ciccone (1996) (*)
Political Instability Proxies		<ul style="list-style-type: none"> • Barro (1991) (-,*) • Barro and Lee (1994) (-,*) • Sachs and Warner (1995) (-,_) • Alesina, et al. (1996) (-,*) • Caselli, et al. (1996) (-,*) • Easterly and Levine (1997a) (-,*)
Political Rights and Civil Liberties Indices	Civil Liberties	<ul style="list-style-type: none"> • Kormendi and Meguire (1985) (+,_) • Levine and Renelt (1992) (? ,not robust) • Barro and Lee (1994) (-,*)
	Overall	<ul style="list-style-type: none"> • Sachs and Warner (1995) (+,*)
	Political Rights	<ul style="list-style-type: none"> • Barro (1991) (?,_) • Barro and Lee (1994) (+,*) • Sala-i-Martin (1997a,b) (+,*)
Political Institutions	Constraints on Executive	<ul style="list-style-type: none"> • Acemoglu, et al. (2001) (+,*)
	Judicial Independence	<ul style="list-style-type: none"> • Feld and Voigt (2003) (+,*)
Property Rights	ICRG index	<ul style="list-style-type: none"> • Knack (1999) (+,*)
	Expropriation Risk	<ul style="list-style-type: none"> • Acemoglu, et al. (2001) (+,*) • Macarthur and Sachs (2001) (+,*)
Population	Density	<ul style="list-style-type: none"> • Sachs and Warner (1995) (+,_)
	Growth	<ul style="list-style-type: none"> • Kormendi and Meguire (1985) (-,*)

		<ul style="list-style-type: none"> • Levine and Renelt (1992) (-,not robust) • Mankiw, et al. (1992) (-,*) • Barro and Lee (1994) (+,_) • Kelley and Schmidt (1995) (-,*) • Bloom and Sachs (1998) (-,*)
Price Distortions	Consumption Price	<ul style="list-style-type: none"> • Easterly (1993) (+,_) • Harrison (1996) (-,*)
	Investment Price	<ul style="list-style-type: none"> • Barro (1991) (-,*) • Easterly (1993) (-,*)
Price Levels	Consumption Price	<ul style="list-style-type: none"> • Easterly (1993) (+,_)
	Investment Price	<ul style="list-style-type: none"> • Easterly (1993) (-,*) • Sachs and Warner (1995) (-,*)
Real Exchange Rate	Black Market Premium	<ul style="list-style-type: none"> • Levine and Renelt (1992) (-,not robust) • Barro and Lee (1994) (-,*) • Barro (1996) (-,*) • Harrison (1996) (-,*) • Easterly and Levine (1997a) (-,*) • Sala-i-Martin (1997a,b) (-,*)
	Distortions	<ul style="list-style-type: none"> • Dollar (1992) (-,*) • Easterly (1993) (-,_) • Harrison (1996) (-,_) • Sala-i-Martin (1997a,b) (-,*) • Acemoglu, et al. (2002) (-,_)
	Variability	<ul style="list-style-type: none"> • Dollar (1992) (-,*)
Regional Effects	Absolute Latitude	<ul style="list-style-type: none"> • Barro (1996) (+,*)
	East Asia Dummy	<ul style="list-style-type: none"> • Barro and Lee (1994) (+,_) • Barro (1997) (+,_)
	Former Spanish Colonies Dummy	<ul style="list-style-type: none"> • Barro (1996) (-,*)
	Latin America Dummy	<ul style="list-style-type: none"> • Barro (1991) (-,*) • Barro and Lee (1994) (-,*) • Barro (1997) (-,_) • Easterly and Levine (1997a) (-,*) • Sala-i-Martin (1997a,b) (-,*)

	Sub-Saharan Africa Dummy	<ul style="list-style-type: none"> • Barro (1991) (-,*) • Barro and Lee (1994) (-,*) • Barro (1997) (-,_) • Easterly and Levine (1997a) (-,*) • Sala-i-Martin (1997a,b) (-,*)
Religion	Buddhist	<ul style="list-style-type: none"> • Barro (1996) (+,*)
	Catholic	<ul style="list-style-type: none"> • Sala-i-Martin (1997a,b) (-,*) • Masters and Sachs (2001) (+,*)
	Confucian	<ul style="list-style-type: none"> • Barro (1996) (+,*)
	Muslim	<ul style="list-style-type: none"> • Barro (1996) (+,*) • Sala-i-Martin (1997) (+,*) • Masters and Sachs (2001) (+,_)
	Protestant	<ul style="list-style-type: none"> • Barro (1996) (+,*) • Sala-i-Martin (1997) (-,*) • Masters and Sachs (2001) (+,*)
	Religious belief	<ul style="list-style-type: none"> • Barro and McCleary (2003) (+,*)
	Attendance	<ul style="list-style-type: none"> • Barro and McCleary (2003) (-,*)
Rule of Law Indices		<ul style="list-style-type: none"> • Barro (1996) (+,*) • Acemoglu, et al. (2001) (+,*) • Easterly and Levine (2001) (-,*) • Dollar and Kraay (2003) (+,_) • Alcala and Ciccone (2004) (+,_) • Rodrik et al. (2004) (+,*)
Scale Effects	Total Area	<ul style="list-style-type: none"> • Barro and Lee (1993) • Sala-i-Martin (1997a,b) (?,_)
	Total Labor force	<ul style="list-style-type: none"> • Barro and Lee (1993) • Sala-i-Martin (1997a,b) (?,_)
Social Capital and Related	Social “Infrastructure”	<ul style="list-style-type: none"> • Hall and Jones (1999) (+,*)
	Citizen Satisfaction with Government	<ul style="list-style-type: none"> • Helliwell and Putnam (2000) (+,*) (within Italy)
	Civic Participation	<ul style="list-style-type: none"> • Helliwell (1996) (,_) (within Asia)

		<ul style="list-style-type: none"> • Knack and Keefer (1997) (+,*)
	Groups – as defined by Putnam (1993)	<ul style="list-style-type: none"> • Keefer and Knack (1997) (-,_)
	Groups - as defined by Olson (1982)	<ul style="list-style-type: none"> • Keefer and Knack (1997) (+,_)
	Institutional Performance	<ul style="list-style-type: none"> • Helliwell and Putnam (2000) (+,*) (Italy)
	Civic Community (index of Participation newspaper readership, political behavior)	<ul style="list-style-type: none"> • Helliwell and Putnam (2000) (+,*) (Italy)
	Trust	<ul style="list-style-type: none"> • Granato, et al. (1996) (+, *) • Helliwell (1996) (,_) (Asia) • Knack and Keefer (1997) (+,*) • La Porta et al (1997) (+, *) • Beugelsdijk and van Schalk (2001) (,_) • Zak and Knack (2001) (+,*)
	Social Development Index	<ul style="list-style-type: none"> • Temple and Johnson (1998)
	Extent of Mass Communication	<ul style="list-style-type: none"> • Temple and Johnson (1998)
	Kinship	<ul style="list-style-type: none"> • Temple and Johnson (1998)
	Mobility	<ul style="list-style-type: none"> • Temple and Johnson (1998)
	Middle Class	<ul style="list-style-type: none"> • Temple and Johnson (1998)
	Outlook	<ul style="list-style-type: none"> • Temple and Johnson (1998)
	Social capital (WVS)	<ul style="list-style-type: none"> • Rupasingha, Goetz, and Freshwater (2000) (+,*)
	Social capital (WVS)	<ul style="list-style-type: none"> • Whiteley (2000) (+,*)
	Social Achievement Norm	<ul style="list-style-type: none"> • Granato, et al. (1996b) (+,*) • Swank (1996) (-,*)
	Capability	<ul style="list-style-type: none"> • Temple and Johnson (1998) (+,*)
Trade Policy	Import Penetration	<ul style="list-style-type: none"> • Levine and Renelt (1992) (? not)

		robust)
Indices	Leamer's Intervention Index	<ul style="list-style-type: none"> Levine and Renelt (1992) (-,not robust)
	Years-Open 1950-1990	<ul style="list-style-type: none"> Sachs and Warner (1996) (+,*) Sala-i-Martin (1997a,b) (+,*)
	Openness Indices (growth)	<ul style="list-style-type: none"> Harrison (1996) (+,*)
	Openness Indices (level)	<ul style="list-style-type: none"> Levine and Renelt (1992) (? ,not robust) Sachs and Warner (1995) (+,*) Harrison (1996) (+,*) Wacziarg and Welch (2003) (+,*)
	Outward Orientation	<ul style="list-style-type: none"> Levine and Renelt (1992) (? ,not robust) Sala-i-Martin (1997a,b) (? ,_)
	Tariff	<ul style="list-style-type: none"> Barro and Lee (1994) (-,_) Sala-i-Martin (1997a,b) (? ,_)
Trade Statistics	Fraction of Export/Import/Total-Trade in GDP	<ul style="list-style-type: none"> Levine and Renelt (1992) (+,not robust) Easterly and Levine (1997a) (? ,_) Frankel and Romer (1999) (+,*) Dollar and Kraay (2003) (+,_) Alcala and Ciccone (2004) (+,*) Rodrik et al. (2004) (+,_)
	Fraction of Primary Products in Total Exports	<ul style="list-style-type: none"> Sachs and Warner (1996) (-,*) Sala-i-Martin (1997) (-,*)
	Growth in Export-GDP Ratio	<ul style="list-style-type: none"> Feder (1982) (+,*) Kormendi and Meguire (1985) (+,*) 20+ studies others
	FDI inflows relative to GDP	<ul style="list-style-type: none"> Blomstrom, et al. (1996)
	Machinery and Equipment Import	<ul style="list-style-type: none"> Romer (1993) (+,*)
Volatility of Shocks	Growth Innovations	<ul style="list-style-type: none"> Kormendi and Meguire (1985) (-,*) Ramey and Ramey (1995) (-,*)

	Monetary Shock	<ul style="list-style-type: none"> • Kormendi and Meguire (1985) (-,*)
War	Casualties per Capita	<ul style="list-style-type: none"> • Easterly, et al. (1993) (-,_)
	Dummy	<ul style="list-style-type: none"> • Barro and Lee (1994) (-,_) • Easterly and Levine (1997a) (?,_) • Sala-i-Martin (1997a,b) (-,*)
	Duration	<ul style="list-style-type: none"> • Barro and Lee (1994) (+,_)

.Appendix 3: Instruments Variables for Solow Growth Determinants

Variable	Instrument	Study
GDP growth	Rainfall variation	Miguel, Satyanath, and Sergenti (2003)
GDP – initial	Lagged values	Barro and Sala-i-Martin (2004)
GDP – initial (per capital stock)	Newsprint consumption, and number of radios	Romer (1990)
GDP – initial	Log population initial and trade measure	Bosworth and Collins (2003)
Human Capital	Natural Disasters	Toya, Skidmore, and Robertson (2003)
Investment - Equipment	Equipment prices, WCR survey variables, national savings rates	DeLong and Summers, (1991)
Investment - Education	Age demographics (16) and lagged capital	Cook (2002b)
Investment - Education	Age demographic variables	Higgins (1998)
Investment - Education	Average level	Beaudry, Collard, and Green (2002)
Investment	Initial values of investment/GDP, population growth and GDP	Cho (1996)
Investment	Lagged investment, lagged output, lagged inflation, trade/GDP and gov spend/GDP	Bond, Leblebicioglu, and Schiantarelli (2004).
Investment	Initial investment in sub-period, average savings rate in sub-period,	Beaudry, Collard, and Green (2002)
Population Growth	Initial values of investment/GDP,	Cho (1996)

	population growth and GDP	
Population Growth	Average population growth over sub-period	Beaudry, Collard, and Green (2002)
Neoclassical convergence RHS variables	Civilian fatalities as %of population (and squared), Number of months of occupation by German forces, and number of months of land battles in country	Cook (2002a)

Appendix 4: Instruments Variables for non-Solow Growth Determinants

Variable		Instrument	Study
Capital market imperfections		Degree of insider trading	Bekaert, Harvey, and Lundblad (2001)
Capital Controls		Lagged Values	McKenzie (2001)
Capital Controls		Lagged Values	Grilli and Milesi-Ferretii (1995)
Corruption		Ethnolinguistic Fractionalization	Mauro (1995)
Coups		All variables (some lagged)	Londregan and Poole (1990)
Defense variables		Initial levels of investment, openness, military expenditure and GDP per capita	Guaresma and Reitschuler (2003).
Democracy		Various	Tavares and Wacziarg (2001)
Demography - Urban concentration		Lagged values	Henderson (2000)
Economic freedom		Lagged values	Lundström (2002)
Education	Male and Female level	Religion and civil liberty measures	Dollar and Gatti. (1999)
	Changes in attainment and female/male ratio of change	Change in total fertility rate, educational spend/GDP, initial fertility level	Klasen (2002).
	Change and level	Kyriacou schooling data	Krueger and Lindahl (2001).
Enterprise Size		Legal origin, resource endowments, religious composition, ethnic diversity, and others	Beck, Demirguc-Kunt, and Levine (2003)

Finance	Development	Legal origins and initial income	Demetriades and Law (2004).
	Competition	Legal origin	Claessens and Laeven (2003).
	Various indicators	Initial values of same	King and Levine (1993).
	Depth	“Legal origin” and lagged versions of all explanatory variables	Levine, Loayza, and Beck (2000)
	Depth	Consumption, GDP, and others	Levine and Zervos (1998).
	Depth	Lagged versions of all explanatory variables	Loayza and Ranciere (2002)
	Various “factors”	Wide variety of initial values of regressors and initial inflation	Rousseau and Sylla (2001).
	Depth	Initial values of inflation and financial depth	Rousseau and Wachtel (2002)
Gini Coefficient		Number of municipal townships in 1962, share of labor force in manufacturing in 1990, percentage of revenue from intergovernmental transfers in 1962	Alesina and La Ferrara (2002)
Government Change		Lagged government change and variable reflected composition change in the executive without a government change	Alesina, Ozler, Roubini, and Swagel (1996)
Government Expenditure and Taxation		Various	Agell, Ohlsson, and Thoursie (2003)

Health	Change in Malaria Infection rate	Six variable for % land coverage of type of forest and desert	Gallup, Mellinger, and Sachs (2000).
	Expenditure	Physicians, visits, dialysis, insurance coverage, alcohol, over 65, beds	Rivera and Currais (1999)
	Inflation	Lagged explanatory variables	Li and Zou (2002).
Inflation		Initial values of inflation and financial depth	Rousseau and Wachtel (2002)
Infrastructure		Lagged values	Esfahani and Ramirez (2003).
Institutions	Various	Settler mortality rate	Acemoglu, Johnson, and Robinson (2002)
	Various	Historically determined component of current institutional quality	Acemoglu, Johnson, and Robinson (2002).
	Various	Geographically determined component of trade as fraction of GDP AND linguistic origins	Alcalá and Ciccone (2004).
	Various and trust	Lagged values	Keefer and Knack (1997).
	Quality	Mortality rates and initial income)	Demetriades and Law (2004).
Manufacturing Exports		Lagged values	Calderón, Chong, and Zanforlin (2001).
Religiosity		Presence of state religion, regulation of religion, indicator of religious pluralism, and others	Barro and McCleary (2003)

Social Infrastructure		State antiquity	Bockstette, Chanda, and Putterman (2003)
Social Infrastructure		Distance from equator, fraction speaking primary European language, fraction speaking English, Frankel and Romer's log predicted trade share	Hall and Jones (1999).
Stock markets		Lagged stock market activity	Harris (1997)
Technology Gap (first difference)		Lagged (second difference)	Hultberg, Nadiri, and Sickles (2003)
Trade	As Share of GDP	Geographically determined component of trade as fraction of GDP AND linguistic origins	Alcalá and Ciccone (2001).
	Policy indices	Lagged values and others unreported by author	Edwards (1998).
	Policy indices	Lagged values	Amable (2000)
		Geographically determined component of trade as fraction of GDP	Frankel and Romer (1996,1999)
Various - Log initial GDP, broad money to GDP, gov expenditure to GDP		Lagged values	Rousseau (2002)

References

- Abramowitz, M., (1986), "Catching Up, Forging Ahead and Falling Behind," *Journal of Economic History*, 46, 385-406.
- Acemoglu, D., S. Johnson, and J. Robinson, (2001), "The Colonial Origins of Comparative Development: An Empirical Investigation," *American Economic Review*, 91, 5, 1369-1401.
- Acemoglu, D., S. Johnson, and J. Robinson, (2002), "Reversal of Fortune: Geography and Institutions in the Making of the Modern World Income Distribution," *Quarterly Journal of Economics*, 117, 4, 1231-1294.
- Acemoglu, D., S. Johnson, and J. Robinson, (2004), "Institutions as the Fundamental Cause of Long-run Growth," *National Bureau of Economic Research Working Paper no.10481*.
- Acemoglu, D., S. Johnson, J. Robinson, and Y. Thaicharoen, (2003), "Institutional Causes, Macroeconomic Symptoms: Volatility, Crises and Growth," *Journal of Monetary Economics*, 50, 1, 49-123.
- Acemoglu, D. and F. Zilibotti, (2001), "Productivity Differences," *Quarterly Journal of Economics*, 116, 563-606.
- Agell, J., H. Ohlsson, and P. Thoursie, (2003), "Growth Effects of Government Expenditure and Taxation in Rich Countries: A Comment," mimeo, Stockholm University.
- Agénor, P.-R., (2004), "Macroeconomic Adjustment and the Poor: Analytical Issues and Cross-Country Evidence," *Journal of Economic Surveys*, 18, 3, 351-408.
- Ahn, S. and P. Schmidt, (1995), "Efficient Estimation of Models for Dynamic Panel Data," *Journal of Econometrics*, 68, 5-27.
- Aizenman, J. and R. Glick, (2003), "Military Expenditure, Threats and Growth," *National Bureau of Economic Research Working Paper no. 9618*.
- Akerlof, G., (1997), "Social Distance and Economic Decisions," *Econometrica*, 65, 5, 1005-1027.
- Alcala, F. and A. Ciccone, (2004), "Trade and Productivity," *Quarterly Journal of Economics*, 119, 2, 613-646.
- Alesina, A., A. Devleeschauwer, W. Easterly, S. Kurlat, and R. Wacziarg, (2003), "Fractionalization," *Journal of Economic Growth*, 8, 2, 155-194.

Alesina, A. and E. La Ferrara, (2002), "Who Trusts Others?," *Journal of Public Economics*, 85, 207-235.

Alesina, A., S. Ozler, N. Roubini, and P. Swagel, (1996), "Political Instability and Economic Growth," *Journal of Economic Growth*, 1, 2, 189-211.

Alesina, A. and D. Rodrik, (1994), "Distributive Politics and Economic Growth," *Quarterly Journal of Economics*, 109, 2, 465-490.

Amable, B., (2000), "International Specialisation and Economic Growth," mimeo, University of Lille II.

Anderson, G., (2003), "Making Inferences About the Polarization, Welfare, and Poverty of Nations: A Study of 101 Countries 1970-1995," mimeo, University of Toronto and forthcoming, *Journal of Applied Econometrics*.

Anderson, G., (2004), "Toward an Empirical Analysis of Polarization," *Journal of Econometrics*, 122, 1, 1-26.

Anderson, G. and Y. Ge, (2004), "A New Approach to Convergence: City Types and "Complete" Convergence of Post-Reform Chinese Urban Income Distributions," mimeo, University of Toronto.

Andrade, E., M. Laurini, R. Madalozzo, and P. Valls Pereira, (2004), "Convergence Clubs among Brazilian Municipalities," *Economics Letters*, 83, 179-84.

Andres, J. and A. Lamo, (1995), "Dynamics of the Income Distribution Across OECD Countries," *London School of Economics, Centre for Economic Performance Discussion Paper no. 252*.

Anselin, L., (2001), "Spatial Econometrics," in *A Companion to Theoretical Econometrics*, B. Baltagi, ed., Oxford: Blackwell.

Arellano, M., (2003), *Panel Data Econometrics*, Oxford: Oxford University Press.

Arellano, M. and S. Bond, (1991), "Some Tests of Specification for Panel Data: Monte Carlo Evidence and an Application to Employment Equations," *Review of Economic Studies*, 58, 2, 277-97.

Arellano, M. and O. Bover, (1995), "Another Look at the Instrumental-Variable Estimation of Error-Components Models," *Journal of Econometrics*, 68, 29-51.

Azariadis, C. and A. Drazen, (1990), "Threshold Externalities in Economic Development," *Quarterly Journal of Economics*, 105, 2, 501-526.

Azariadis, C. and J. Stachurski, (2003), "A Forward Projection of the Cross-Country Income Distribution," *Institute of Economic Research, Kyoto University, Discussion Paper No. 570*.

Baltagi, B., S. Song, and W. Koh, (2003), "Testing Panel Data Regression Models with Spatial Error Correlation," *Journal of Econometrics*, 117, 1, 123-50.

Bandyopadhyay, S., (2002), "Polarisation, Stratification, and Convergence Clubs: Some Dynamics and Explanations of Unequal Economic Growth Across Indian States," mimeo, London School of Economics.

Banerjee, A. and E. Duflo, (2003), "Inequality and Growth: What Can the Data Say?," *Journal of Economic Growth*, 8, 3, 267-300.

Barro, R., (1991), "Economic Growth in a Cross Section of Countries," *Quarterly Journal of Economics*, 106, 2, 407-43.

Barro, R., (1996), "Democracy and Growth," *Journal of Economic Growth*, 1, 1, 1-27.

Barro, R., (1997), *Determinants of Economic Growth*, Cambridge: MIT Press.

Barro, R. and J.-W. Lee, (1994), "Sources of Economic Growth (with commentary)," *Carnegie-Rochester Conference Series on Public Policy*, 40, 1-57.

Barro, R., N. G. Mankiw, and X. Sala-i-Martin, (1995), "Capital Mobility in Neoclassical Models of Growth," *American Economic Review*, 85, 1, 103-115.

Barro, R. and R. McCleary, (2003), "Religion and Economic Growth Across Countries," *American Sociological Review*, 68, 5, 760-781.

Barro, R. and X. Sala-i-Martin, (1991), "Convergence Across States and Regions," *Brookings Papers on Economic Activity*, 1, 107-158.

Barro, R. and X. Sala-i-Martin, (1992), "Convergence," *Journal of Political Economy*, 100, 223-51.

Barro, R. and X. Sala-i-Martin, (1995), *Economic Growth*, New York: McGraw-Hill.

Barro, R. and X. Sala-i-Martin, (1997), "Technological Diffusion, Convergence, and Growth," *Journal of Economic Growth*, 2, 1-26.

Barro, R. and X. Sala-i-Martin, (2004), *Economic Growth*, Second edition, Cambridge: MIT Press.

Baumol, W., (1986), "Productivity Growth, Convergence, and Welfare: What the Long-run Data Show," *American Economic Review*, 76, 5, 1072-85.

- Beaudry, P., F. Collard, and D. Green, (2002), "Decomposing the Twin-Peaks in the World Distribution of Output-per-Worker," *National Bureau of Economic Research Working Paper no. 9240*.
- Beck, T., A. Demirguc-Kunt, and R. Levine, (2003), "Small and Medium Enterprises, Growth, and Poverty: Cross Country Evidence," *World Bank Research Policy Paper 3178*.
- Beck, T. and R. Levine, (2004), "Stock Market, Banks, and Growth: Panel Evidence," *Journal of Banking and Finance*, 28, 3, 423-442.
- Bekaert, G., C. Harvey, and C. Lundblad, (2001), "Does Financial Liberalization Spur Growth?," *National Bureau of Economic Research Working Paper no. 8245*.
- Ben-David, D., (1993), "Equalizing Exchange: Trade Liberalization and Income Convergence," *Quarterly Journal of Economics*, 108, 653-679.
- Ben-David, D., (1996), "Trade and Convergence Among Countries," *Journal of International Economics*, 40, 3/4, 279-298.
- Bernard, A. and S. Durlauf, (1995), "Convergence in International Output," *Journal of Applied Econometrics*, 10, 2, 97-108.
- Bernard, A. and S. Durlauf, (1996), "Interpreting Tests of the Convergence Hypothesis," *Journal of Econometrics*, 71, 1-2, 161-73.
- Bernard, A. and C. Jones, (1996), "Comparing Apples to Oranges: Productivity Convergence and Measurement across Industries and Countries," *American Economic Review*, 86, 5, 1216-1238.
- Berthelemy, J. and A. Varoudakis, (1996), "Economic Growth, Convergence Clubs, and the Role of Financial Development," *Oxford Economic Papers*, 48, 300-328.
- Besley, T. and R. Burgess, (2000), "Land Reform, Poverty Reduction, and Growth: Evidence from India," *Quarterly Journal of Economics*, 115, 2, 389-430.
- Besley, T. and R. Burgess, (2002), "The Political Economy of Government Responsiveness: Theory and Evidence from India," *Quarterly Journal of Economics*, 117, 4, 1415-1451.
- Besley, T. and R. Burgess, (2004), "Can Labor Regulation Hinder Economic Performance? Evidence from India," *Quarterly Journal of Economics*, 119, 1, 91-134.
- Beugelsdijk, S. and T. van Schalk, (2001), "Social Capital and Regional Economic Growth," mimeo, Tilburg University.

Bhargava, A., D. Jamison, L. Lau, and C. Murray, (2001), "Modeling the Effects of Health on Economic Growth," *Journal of Health Economics*, 20, 3, 423-440.

Bianchi, M., (1997), "Testing for Convergence: Evidence from Nonparametric Multimodality Tests," *Journal of Applied Econometrics*, 12, 4, 393-409.

Bils, M. and P. Klenow, (2000), "Does Schooling Cause Growth?," *American Economic Review*, 90, 5, 1160-1183.

Binder, M. and S. Brock, (2004), "A Re-Examination of Determinants of Economic Growth Using Simultaneous Equation Dynamic Panel Data Models," mimeo, Johannes Goethe University, Frankfurt.

Binder, M. and M. H. Pesaran, (1999), "Stochastic Growth Models and their Econometric Implications," *Journal of Economic Growth*, 4, 139-183.

Binder, M. and M. H. Pesaran, (2001), "Life Cycle Consumption Under Social Interactions," *Journal of Economic Dynamics and Control*, 25, 1-2, 35-83.

Blattman, C., J. Hwang, and J. Williamson, (2003), "The Terms of Trade and Economic Growth in the Periphery 1870-1983," *National Bureau of Economic Research Working Paper no. 9940*.

Bliss, C., (1999), "Galton's Fallacy and Economic Convergence," *Oxford Economic Papers*, 51, 4-14.

Bliss, C., (2000), "Galton's Fallacy and Economic Convergence: A Reply to Cannon and Duck," *Oxford Economic Papers*, 52, 420-422.

Blomstrom, M., R. Lipsey, and M. Zejan, (1996), "Is Fixed Investment the Key to Growth?," *Quarterly Journal of Economics*, February, 111, 1, 269-276.

Blonigen, B. and M. Wang, (2004), "Inappropriate Pooling of Wealthy and Poor Countries in Empirical FDI Studies," *National Bureau of Economic Research Working Paper no. 10378*.

Bloom, D., D. Canning, and J. Sevilla, (2002), "Health, Worker Productivity, and Economic Growth," mimeo, Harvard University.

Bloom, D., D. Canning, and J. Sevilla, (2003), "Geography and Poverty Traps," *Journal of Economic Growth*, 8, 355-378.

Bloom, D. and P. Malaney, (1998), "Macroeconomic Consequences of the Russian Mortality Crisis," *World Development*, 26, 11, 2073-2085.

Bloom, D. and J. Sachs, (1998), "Geography, Demography, and Economic Growth in Africa," mimeo, Harvard Institute for International Development.

Bloom, D. and J. Williamson, (1998), "Demographic Transitions and Economic Miracles in Emerging Asia," *World Bank Economic Review*, 12, 3, 419-55.

Blundell, R. and S. Bond, (1998), "Initial Conditions and Moment Restrictions in Dynamic Panel Data Models," *Journal of Econometrics*, 87, 1, 115-43.

Bockstette, V., A. Chanda, and L. Putterman, (2003), "States and Markets: The Advantages of an Early Start," *Journal of Economic Growth*, 7, 347-369.

Bond, S., (2002), "Dynamic Panel Data Models: A Guide to Micro Data Methods and Practicem," *Portugese Economic Journal*, 1, 141-162.

Bond, S., A. Hoeffler, and J. Temple, (2001), "GMM Estimation of Empirical Growth Models," *Centre for Economic Policy Research Discussion Paper no. 3048*.

Bond, S., A. Leblebicioglu, and F. Schiantarelli, (2004), "Capital Accumulation and Growth: A New Look at the Empirical Evidence," *Nuffield College, Oxford, Working Paper no. 2004-W8*.

Bosworth, B. and S. Collins, (2003), "The Empirics of Growth: An Update (with discussion)," *Brookings Papers on Economic Activity*, 2, 113-206.

Breiman, L., J. Friedman, R. Olshen and C. Stone, (1984), *Classification and Regression Trees*, Redwood City: Wadsworth Publishing.

Breusch, T. and A. Pagan, (1980), "The Lagrange Multiplier Test and Its Application to Model Specifications in Econometrics," *Review of Economic Studies*, 47, 239-253.

Brock, W. and S. Durlauf, (2001a), "Growth Empirics and Reality," *World Bank Economic Review*, 15, 2, 229-272.

Brock, W. and S. Durlauf, (2001b), "Interactions-Based Models," in *Handbook of Econometrics 5*, J. Heckman and E. Leamer, eds., Amsterdam: North-Holland.

Brock, W., S. Durlauf, and K. West, (2003), "Policy Evaluation in Uncertain Economic Environments (with discussion)," *Brookings Papers on Economic Activity*, 1, 235-322.

Bruno, M. and W. Easterly, (1998), "Inflation Crises and Long-Run Growth," *Journal of Monetary Economics*, 41, 1, 3-26.

Bulli, S., (2001), "Distribution Dynamics and Cross-Country Convergence: New Evidence," *Scottish Journal of Political Economy*, 48, 226-243.

- Bun, M. and J. Kiviet, (2001), "The Accuracy of Inference in Small Samples of Dynamic Panel Data Models," *Tinbergen Institute Discussion Paper no. 2001-006/4*.
- Calderón, C., A. Chong and L. Zanforlin, (2001), "On the Non-Linearities Between Exports of Manufactures and Economic Growth," *Journal of Applied Economics*, 4, 279-311.
- Campos, N. and J. Nugent, (2002), "Who is Afraid of Political Instability?," *Journal of Development Economics*, 67, 157-172.
- Cannon, E. and N. Duck, (2000), "Galton's Fallacy and Economic Convergence," *Oxford Economic Papers*, 53, 415-419.
- Canova, F., (2004), "Testing for Convergence Clubs in Income Per Capita: A Predictive Density Approach," *International Economic Review*, 45, 1, 49-77.
- Canova, F., and A. Marcet, (1995), "The Poor Stay Poor: Non-Convergence Across Countries and Regions," *Centre for Economic Policy Research Discussion Paper 1265*.
- Carlino, G. and L. Mills, (1993), "Are U.S. Regional Incomes Converging?: A Time Series Analysis," *Journal of Monetary Economics*, 32, 2, 335-346.
- Caselli, F. and W. J. Coleman, (2003), "The World Technology Frontier," mimeo, Harvard University.
- Caselli, F., G. Esquivel, and F. Lefort, (1996), "Reopening the Convergence Debate: A New Look at Cross Country Growth Empirics," *Journal of Economic Growth*, 1, 3, 363-89.
- Cashin, P., (1995), "Economic Growth and Convergence Across the Seven Colonies of Australasia: 1861-1991," *Economic Record*, 71, 132-144.
- Cashin, P. and R. Sahay, (1996), "Regional Economic Growth and Convergence in India," *Finance and Development*, 33, 49-52.
- Chesher, A., (1984), "Testing for Neglected Heterogeneity," *Econometrica*, 52, 4, 865-72.
- Cho, D., (1996), "An Alternative Interpretation of Conditional Convergence Results," *Journal of Money, Credit and Banking*, 28, 1, 669-681.
- Ciccone, A., (1996), "Externalities and Interdependent Growth: Theory and Evidence," mimeo, UC-Berkeley.
- Claessens, S. and L. Laeven, (2003), "Competition in the Financial Sector and Growth: A Cross Country Perspective," mimeo, University of Amsterdam.

Cohen, D., (1996), "Tests of the Convergence Hypothesis: Some Further Results," *Journal of Economic Growth*, 1, 3, 351-361.

Collier, P. and J. Gunning, (1999a), "Why Has Africa Grown Slowly?," *Journal of Economic Perspectives*, 13, 3, 3-22.

Collier, P. and J. Gunning, (1999b), "Explaining African Economic Performance," *Journal of Economic Literature*, 37, 1, 64-111.

Conley, T., (1999), "GMM Estimation with Cross-Section Dependence," *Journal of Econometrics*, 92, 1-45.

Conley, T. and E. Ligon, (2002), "Economic Distance and Long-run Growth," *Journal of Economic Growth*, 7, 2, 157-187.

Conley, T. and G. Topa, (2002), "Socio-Economic Distance and Spatial Patterns in Unemployment," *Journal of Applied Econometrics*, 17, 4, 303-327.

Cook, D., (2002a), "World War II and Convergence," *Review of Economics and Statistics*, 84, 1, 131-138.

Cook, D., (2002b), "Education and Growth: Instrumental Variables Estimates," mimeo, Hong Kong University of Science and Technology.

Corrado, L., R. Martin, and M. Weeks, (2004), "Identifying and Interpreting Regional Convergence Clusters Across Europe," *Economic Journal*, forthcoming.

Coulombe, S. and F. Lee, (1995), "Convergence Across Canadian Provinces, 1961 to 1991," *Canadian Journal of Economics*, 28, 886-898.

Dagenais, M., and D. Dagenais, (1997), "Higher Moment Estimators for Linear Regression Models with Errors in the Variables," *Journal of Econometrics*, 76, 1-2, 193-221.

Davidson, R. and J. MacKinnon, (1993), *Estimation and Inference in Econometrics*, Oxford: Oxford University Press.

De Gregorio, J. and P. Guidotti, (1995), "Financial Development and Economic Growth," *World Development*, 23, 3, 433-448.

Deidda, L. and B. Fattouh, (2002), "Non-Linearity Between Finance and Growth," *Economics Letters*, 74, 339-345.

DeLong, J. B., (1988), "Productivity Growth, Convergence, and Welfare: Comment," *American Economic Review*, 78, 5, 1138-1154.

DeLong, J. B. and L. Summers, (1991), "Equipment Investment and Economic Growth," *Quarterly Journal of Economics*, 106, 2, 445-502.

DeLong, J. B. and L. Summers, (1993), "How Strongly do Developing Economies Benefit from Equipment Investment?," *Journal of Monetary Economics*, 32, 3, 395-415.

Demetriades, P. and S. Law, (2004), "Finance, Institutions and Economic Growth," *University of Leicester Working Paper 04/5*.

Desdoigts, A., (1999), "Patterns of Economic Development and the Formation of Clubs," *Journal of Economic Growth*, 4, 3, 305-330.

Diebold, F. and A. Inoue, (2001), "Long Memory and Regime Switching," *Journal of Econometrics*, 105, 1, 131-159.

Dollar, D., (1992), "Outward-Oriented Developing Economies Really Do Grow More Rapidly: Evidence from 95 LDCs, 1976-85," *Economic Development and Cultural Change*, 40, 523-544.

Dollar, D. and R. Gatti, (1999), "Gender Inequality, Income, and Growth: Are Good Times Good for Women?," mimeo, World Bank.

Dollar, D. and A. Kraay, (2002), "Growth is Good for the Poor," *Journal of Economic Growth*, 7, 3, 195-225.

Dollar, D. and A. Kraay, (2003), "Institutions, Trade and Growth: Revisiting the Evidence," *Journal of Monetary Economics*, 50, 1, 133-162.

Doppelhofer, G., R. Miller, and X. Sala-i-Martin, (2004), "Determinants of Long-Term Growth: A Bayesian Averaging of Classical Estimates (BACE) Approach," *American Economic Review*, 94, 4, 813-835.

Dowrick, S., (2004), "De-Linearising the Neo-Classical Convergence Model," in *Economic Growth and Macrodynamics: Recent Developments in Economic Theory*, S. Turnovsky, S. Dowrick, and R. Pitchford, eds., New York: Cambridge University Press.

Dowrick, S. and J. Quiggin, (1997), "True Measures of GDP and Convergence," *American Economic Review*, 87, 1, 41-64.

Dowrick, S. and M. Rogers, (2002), "Classical and Technological Convergence: Beyond the Solow-Swan Growth Model," *Oxford Economic Papers*, 54, 369-385.

Draper, D., (1995), "Assessment and Propagation of Model Uncertainty," *Journal of the Royal Statistical Society, series B*, 57, 45-70.

Draper, D., J. Hodges, C. Mallows, and D. Pregibon, (1993), "Exchangeability and Data Analysis (with discussion)," *Journal of the Royal Statistical Society, series A*, 156, 9-37.

Driscoll, J. and A. Kraay, (1998), "Consistent Covariance Matrix Estimation with Spatially Dependent Panel Data," *Review of Economics and Statistics*, 80, 4, 549-560.

Duffy, J. and C. Papageorgiou, (2000), "A Cross-Country Empirical Investigation of the Aggregate Production Function Specification," *Journal of Economic Growth*, 5, 87-120.

Durlauf, S., (1996), "Controversy on the Convergence and Divergence of Growth Rates," *Economic Journal*, 106, 1016-1018.

Durlauf, S., (2001), "Manifesto for a Growth Econometrics," *Journal of Econometrics*, 100, 1, 65-69.

Durlauf, S., (2002), "On the Empirics of Social Capital," *Economic Journal*, 112, 459-479.

Durlauf, S. and P. Johnson, (1995), "Multiple Regimes and Cross Country Growth Behaviour," *Journal of Applied Econometrics*, 10, 4, 365-84.

Durlauf, S., A. Kourtellos, and A. Minkin, (2001), "The Local Solow Growth Model," *European Economic Review*, 45, 4-6, 928-40.

Durlauf, S. and D. Quah, (1999), "The New Empirics of Economic Growth," in *Handbook of Macroeconomics*, J. Taylor and M. Woodford, eds., Amsterdam: North Holland.

Easterly, W., (1993), "How Much Do Distortions Affect Growth?," *Journal of Monetary Economics*, 32, 2, 187-212.

Easterly, W., (1994), "Economic Stagnation, Fixed Factors, and Policy Thresholds," *Journal of Monetary Economics*, 33, 3, 525-57.

Easterly, W., (1996), "When is Stabilization Expansionary?," *Economic Policy*, 22, 67-98.

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

Easterly, W., M. Kremer, L. Pritchett, and L. Summers, (1993), "Good Policy or Good Luck? Country Growth Performance and Temporary Shocks," *Journal of Monetary Economics*, 32, 459-483.

Easterly, W. and R. Levine, (1997a), "Africa's Growth Tragedy: Policies and Ethnic Divisions," *Quarterly Journal of Economics*, 112, 4, 1203-50.

Easterly, W. and R. Levine, (1997b), "Troubles with the Neighbours: Africa's Problem, Africa's Opportunity," *Journal of African Economies*, 7, 1, 120-42.

Easterly, W. and R. Levine, (2001), "It's Not Factor Accumulation: Stylized Facts and Growth Models," *World Bank Economic Review*, 15, 177-219.

Eaton, J. and S. Kortum, (1999), "International Technology Diffusion: Theory and Measurement," *International Economic Review*, 40, 3, 537-570.

Eaton, J. and S. Kortum, (2001), "Trade in Capital Goods," *European Economic Review*, 45 7, 1195-1235.

Edwards, S., (1998), "Openness, Productivity and Growth: What Do We Really Know?," *Economic Journal*, 108, 383-98.

Edwards, S. and I. Magendzo, (2003), "Strict Dollarization and Economic Performance: An Empirical Investigation," *International Journal of Finance and Economics*, 8, 4, 351-363.

Eichengreen, B. and D. Leblang, (2003), "Capital Account Liberalization and Growth: Was Mr. Mahathir Right?," *International Journal of Finance and Economics*, 8, 3, 205-224.

Eicker, F., (1967), "Limit Theorems for Regressions with Unequal and Dependent Errors," in *Proceedings of the Fifth Berkeley Symposium on Mathematical Statistics and Probability*, Vol. 1, Berkeley: University of California.

Esfahani, H. and M. Ramirez, (2003), "Institutions, Infrastructure, and Economic Growth," *Journal of Development Economics*, 70, 2, 443-477.

Evans, P., (1996), "Using Cross-Country Variances to Evaluate Growth Theories," *Journal of Economic Dynamics and Control*, 20, 1027-1049.

Evans, P., (1997), "How Fast Do Economies Converge?," *Review of Economics and Statistics*, 79, 2, 219-225.

Evans, P., (1998), "Using Panel Data to Evaluate Growth Theories," *International Economic Review*, 39, 2, 295-306.

Feder, G., (1982), "On Exports and Economic Growth," *Journal of Development Economics*, 12, 1, 59-74.

Feld, L. and S. Voigt, (2003), "Economic Growth and Judicial Independence: Cross Country Evidence Using a New Set of Indicators," *European Journal of Political Economy*, 19, 3, 497-527.

Fernandez, C., E. Ley and M. Steel, (2001a), "Model Uncertainty in Cross-Country Growth Regressions," *Journal of Applied Econometrics*, 16, 5, 563-76.

Fernandez, C., E. Ley and M. Steel, (2001b) "Benchmark Priors for Bayesian Model Averaging," *Journal of Econometrics*, 100, 2, 381-427.

Feyrer, J., (2003), "Convergence By Parts," mimeo, Dartmouth College.

Fiaschi, D. and A. Lavezzi, (2004), "Distribution Dynamics and Nonlinear Growth," *Journal of Economic Growth*, 8, 379-401.

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

Forbes, K., (2000), "A Reassessment of the Relationship between Inequality and Growth," *American Economic Review*, 90, 4, 869-87.

Frankel, J., (2003), "Discussion," *Brookings Papers on Economic Activity*, 2, 189-199.

Frankel, J. and D. Romer, (1996), "Trade and Growth: An Empirical Investigation," *National Bureau of Economic Research Working Paper no. 5476*.

Frankel, J. and D. Romer, (1999), "Does Trade Cause Growth?," *American Economic Review*, 89, 3, 379-399.

Frankel, J., D. Romer, and T. Cyrus, (1996), "Trade and Growth in East Asian Countries: Cause and Effect?," *National Bureau of Economic Research Working Paper no. 5732*.

Friedman, J., (1987), "Exploratory Projection Pursuit," *Journal of the American Statistical Association*, 82, 249-266.

Friedman, J. and J. Tukey, (1974), "A Projection Pursuit Algorithm for Exploratory Data Analysis," *IEEE Transactions on Computers*, C23, 881-890.

Friedman, M., (1992), "Do Old Fallacies Ever Die?," *Journal of Economic Literature*, 30, 2129-2132.

Galor, O., (1996), "Convergence? Inferences from Theoretical Models," *Economic Journal*, 106, 1056-1069.

Gallup, J., A. Mellinger, and J. Sachs, (2000), "The Economic Burden of Malaria," *Harvard University, Center for International Development Working Paper No. 1*.

Giavazzi, F. and G. Tabellini, (2004), "Economic and Political Liberalizations," *Centre for Economic Policy Research Discussion Paper no. 4579*.

Goetz, S. and D. Hu, (1996), "Economic Growth and Human Capital Accumulation: Simultaneity and Expanded Convergence Tests," *Economic Letters*, 51, 355-362.

Graham, B. and J. Temple, (2003), "Rich Nations, Poor Nations: How Much Can Multiple Equilibria Explain?," mimeo, University of Bristol.

Granato, J., R. Inglehart, and D. Leblang, (1996), "The Effect of Cultural Values on Economic Development: Theory, Hypotheses, and Some Empirical Tests," *American Journal of Political Science*, 40, 3, 607-31.

Granger, C. W. J., (1980), "Long-Memory Relationships and the Aggregation of Dynamic Models," *Journal of Econometrics*, 14, 227-238.

Greasley, D. and L. Oxley, (1997), "Time-Series Tests of the Convergence Hypothesis: Some Positive Results," *Economics Letters*, 56, 143-147.

Grier, K. and G. Tullock, (1989), "An Empirical Analysis of Cross National Economic Growth, 1951-80," *Journal of Monetary Economics*, 24, 2, 259-76.

Grilli, V. and G. Milesi-Ferretii, (1995), "Economic Effects and Structural Determinants of Capital Controls," *International Monetary Fund Working Paper WP/95/31*.

Guaresma, J. and G. Reitschuler, (2003), "Guns or Butter?' Revisited: Robustness and Nonlinearity Issues in the Defense-Growth Nexus," mimeo, University of Vienna.

Hahn, J., (1999), "How Informative Is the Initial Condition in the Dynamic Panel Model with Fixed Effects?," *Journal of Econometrics*, 93, 2, 309-26.

Hahn, J., J. Hausman, and G. Kuersteiner, (2001), "Bias Corrected Instrumental Variables Estimation for Dynamic Panel Models with Fixed Effects," mimeo, Massachusetts Institute for Technology.

Hall, A., (1987), "The Information Matrix Test for the Linear Model," *Review of Economic Studies*, 54, 2, 257-63.

Hall, R., and C. Jones, (1999), "Why Do Some Countries Produce So Much More Output Per Worker Than Others?," *Quarterly Journal of Economics*, 114, 1, 83-116.

Hall, S., D. Robertson, and M. Wickens, (1997), "Measuring Economic Convergence," *International Journal of Finance and Economics*, 2, 131-143.

Hamoudi, A. and J. Sachs, (2000), "Economic Consequences of Health Status: A Review of the Evidence," *Harvard University, CID Working Paper No. 30*.

Hansen, B., (2000), "Sample Splitting and Threshold Estimation," *Econometrica*, 68, 3, 575-603.

- Hanushek, E. and D. Kimko, (2000), "Schooling, Labor-Force Quality, and the Growth of Nations," *American Economic Review*, 90, 5, 1184-1208.
- Harberger, A., (1987), "Comment," in *Macroeconomics Annual 1987*, S. Fischer, ed., Cambridge: MIT Press.
- Harris, D., (1997), "Stock Markets and Development: A Re-Assessment," *European Economic Review*, 41, 1, 139-146.
- Harrison, A., (1996), "Openness and Growth: A Time-Series, Cross-Country Analysis for Developing Countries," *Journal of Development Economics*, 48, 2, 419-447.
- Hausmann, R., L. Pritchett and D. Rodrik, (2004), "Growth Accelerations," *Centre for Economic Policy Research Discussion Paper no. 4538*.
- Helliwell, J., (1996), "Economic Growth and Social Capital in Asia," in *The Asia Pacific Region in the Global Economy: A Canadian Perspective*, R. Harris, ed., Calgary: University of Calgary Press.
- Helliwell, J. and R. Putnam, (2000), "Economic Growth and Social Capital in Italy," in *Social Capital: A Multifaceted Perspective*, P. Dasgupta and I. Seragilden, eds., Washington DC: World Bank.
- Henderson, D. and R. Russell, (2004), "Human Capital and Convergence: A Production Frontier Approach," mimeo, SUNY Binghamton and forthcoming, *International Economic Review*.
- Henderson, J. V., (2000), "The Effects of Urban Concentration on Economic Growth," *National Bureau of Economic Research Working Paper no. 7503*.
- Hendry, D., (1995), *Dynamic Econometrics*, New York: Oxford University Press.
- Hendry, D. and H.-M. Krolzig, (2004), "We Ran One Regression," mimeo, Oxford University.
- Hendry, D. and H.-M. Krolzig, (2005), "The Properties of Automatic Gets Modelling," *Economic Journal*, forthcoming.
- Henry, P., (2000), "Do Stock Market Liberalizations Cause Investment Booms?," *Journal of Financial Economics*, 58, 1-2, 301-334.
- Henry, P., (2003), "Capital Account Liberalization, the Cost of Capital, and Economic Growth," *American Economic Review*, 93, 2, 91-96.
- Heston, A., R. Summers and B. Aten, (2002), "Penn World Table Version 6.1," *Center for International Comparisons at the University of Pennsylvania (CICUP)*.

Higgins, M., (1998), "Demography, National Savings and International Capital Flows," *International Economic Review*, 39, 343-369.

Hobijn, B. and Franses, P. H., (2000), "Asymptotically Perfect and Relative Convergence of Productivity," *Journal of Applied Econometrics*, 15, 59-81.

Hoeffler, A., (2002), "The Augmented Solow Model and the African Growth Debate," *Oxford Bulletin of Economics and Statistics*, 64, 2, 135-158.

Holtz-Eakin, D., W. Newey, and H. Rosen, (1988), "Estimating Vector Autoregressions with Panel Data," *Econometrica*, 56, 6, 1371-1395.

Hoover, K. and S. Perez, (2004), "Truth and Robustness in Cross-Country Growth Regressions," *Oxford Bulletin of Economics and Statistics*, forthcoming.

Howitt, P., (2000), "Endogenous Growth and Cross-Country Income Differences," *American Economic Review*, 90, 4, 829-846.

Hultberg, P., M. Nadiri, and R. Sickles, (2003), "Cross-country Catch-up in the Manufacturing Sector: Impacts of Heterogeneity on Convergence and Technology adoption," mimeo, University of Wyoming.

Hulten, C., (1996), "Infrastructure Capital and Economic Growth: How Well You Use It May Be More Important Than How Much You Have," *National Bureau of Economic Research Working Paper no. 5847*.

Islam, N., (1995), "Growth Empirics: A Panel Data Approach," *Quarterly Journal of Economics*, 110, 4, 1127-70.

Islam, N. (1998), "Growth Empirics: A Panel Data Approach-A Reply," *Quarterly Journal of Economics*, 113, 325-329.

Islam, N., (2003), "What Have We Learned from the Convergence Debate?," *Journal of Economic Surveys*, 17, 309-362.

Johnson, P., (2000), "A Nonparametric Analysis of Income Convergence Across the US States," *Economics Letters*, 69, 219-223.

Johnson, P., (2004), "A Continuous State Space Approach to 'Convergence by Parts'," mimeo, Vassar College, and forthcoming, *Economic Letters*.

Johnson, P. and L. Takeyama, (2001), "Initial Conditions and Economic Growth in the US States," *European Economic Review*, 45, 4-6, 919-27.

Jones, C., (1995), "Time Series Tests of Endogenous Growth Models," *Quarterly Journal of Economics*, 110, 2, 495-525.

- Jones, C., (1997), "Convergence Revisited," *Journal of Economic Growth*, 2, 2, 131-53.
- Jones, L. and R. Manuelli, (1990), "A Convex Model of Equilibrium Growth: Theory and Policy Implications," *Journal of Political Economy*, 98, 5, 1008-1038.
- Judson, R. and A. Owen, (1999), "Estimating Dynamic Panel Data Models: A Guide for Macroeconomists," *Economics Letters*, 65, 9-15.
- Kalaitzidakis, P., T. Mamuneas, and T. Stengos, (2000), "A Non-linear Sensitivity Analysis of Cross Country Growth Regressions," *Canadian Journal of Economics*, 33, 3, 604-17.
- Kaufmann, D., A. Kraay, and M. Mastruzzi, (2003), "Governance Matters III: Governance Indicators for 1996-2002," mimeo, World Bank.
- Kaufmann, D., A. Kraay, and P. Zoido-Lobaton, (1999a), "Aggregating Governance Indicators," *World Bank Policy Research Department Working Paper No. 2195*.
- Kaufmann, D., A. Kraay, and P. Zoido-Lobaton, (1999b), "Governance Matters," *World Bank Policy Research Department Working Paper No. 2196*.
- Keefer, P. and S. Knack, (1997), "Why Don't Poor Countries Catch Up? A Cross-National Test of an Institutional Explanation," *Economic Inquiry*, 35, 3, 590-602.
- Kelley, A. and R. Schmidt, (1995), "Aggregate Population and Economic Growth Correlations: The Role of the Components of Demographic Change," *Demography*, 32, 543-555.
- Kelly, M., (1992), "On Endogenous Growth with Productivity Shocks," *Journal of Monetary Economics*, 30, 1, 47-56.
- Kelly, T., (1997), "Public Expenditures and Growth," *Journal of Development Studies*, 34, 1, 60-84.
- King, R. and R. Levine, (1993), "Finance and Growth: Schumpeter Might be Right," *Quarterly Journal of Economics*, 108, 3, 717-737.
- Kiviet, J., (1995), "On Bias, Inconsistency, and Efficiency of Various Estimators in Dynamic Panel Data Models," *Journal of Econometrics*, 68, 1, 53-78
- Kiviet, J., (1999), "Expectations of Expansions for Estimators in a Dynamic Panel Data Model: Some Results for Weakly Exogenous Regressors," in *Analysis of Panels and Limited Dependent Variable Models: In Honour of G. S. Maddala, C. Hsiao, ed.*, Cambridge: Cambridge University Press.

- Klasen, S., (2002), "Lower Schooling for Girls, Slower Growth for All? Cross-Country Evidence on the Effect of Gender Inequality in Education on Economic Development," *World Bank Economic Review*, 16, 3, 345-373.
- Klenow, P. and A. Rodriguez-Clare, (1997a), "The Neoclassical Revival in Growth Economics: Has it Gone Too Far?," in *Macroeconomics Annual 1997*, B. Bernanke and J. Rotemberg, eds., Cambridge: MIT Press.
- Klenow, P. and A. Rodriguez-Clare, (1997b), "Economic Growth: A Review Essay," *Journal of Monetary Economics*, 40, 597-617.
- Klepper, S., (1988), "Regression Diagnostics for the Classical Errors-in-Variables Model," *Journal of Econometrics*, 37, 225-250.
- Klepper, S. and E. Leamer, (1984), "Consistent Sets of Estimates for Regressions with Errors in All Variables," *Econometrica*, 52, 1, 163-83.
- Knack, S., (1999), "Social Capital, Growth and Poverty: A Survey of Cross Country Evidence," *World Bank, Social Capital Initiative Working Paper No. 7*.
- Knack, S. and P. Keefer, (1995), "Institutions and Economic Performance: Cross Country Tests Using Alternative Institutional Measures," *Economics and Politics*, 7, 3, 207-27.
- Knack, S. and Keefer, P., (1997), "Does Social Capital Have an Economic Payoff? A Cross-Country Investigation," *Quarterly Journal of Economics*, 112, 4, 1252-1288.
- Knight, M., N. Loayza, and D. Villaneuva, (1993), "Testing the Neoclassical Growth Model," *IMF Staff Papers*, 40, 512-541.
- Knowles, S., (2001), "Inequality and Economic Growth: The Empirical Relationship Reconsidered in the Light of Comparable Data," *CREDIT Research Paper 01/03*.
- Knowles, S. and P. Owen, (1995), "Health Capital and Cross-Country Variation in Income per Capita in the Mankiw-Romer-Weil Model," *Economics Letters*, 48, 1, 99-106.
- Kocherlakota, N. and K.-M. Yi, (1997), "Is There Endogenous Long-Run Growth? Evidence from the United States and the United Kingdom," *Journal of Money, Credit and Banking*, 29, 2, 235-262.
- Kormendi, R. and P. Meguire, (1985), "Macroeconomic Determinants of Growth: Cross Country Evidence," *Journal of Monetary Economics*, 16, 2, 141-63.
- Kourtellos, A., (2003a), "Modeling Parameter Heterogeneity in Cross-Country Growth Regression Models," mimeo, University of Cyprus.

Kourtellos, A., (2003b), "A Projection Pursuit Approach to Cross-Country Growth Data," mimeo, University of Cyprus.

Krasker, W. and J. Pratt, (1986), "Bounding the Effects of Proxy Variables on Regression Coefficients," *Econometrica*, 54, 3, 641-55

Krasker, W. and J. Pratt, (1987), "Bounding the Effects of Proxy Variables on Instrumental Variables Coefficients," *Journal of Econometrics*, 35, 2/3, 233-52.

Kremer, M., A. Onatski, and J. Stock, (2001), "Searching for Prosperity," *Carnegie-Rochester Conference Series on Public Policy*, 55, 275-303.

Krueger, A. and M. Lindahl, (2001), "Education for Growth: Why and for Whom," *Journal of Economic Literature*, 39, 4, 1101-1136.

Lamo, A., (2000), "On Convergence Empirics: Some Evidence for Spanish Regions," *Investigaciones Economicas*, 24, 681-707.

Landes, D., (1998), *The Wealth and Poverty of Nations: Why Some are So Rich and Some So Poor*, New York: W. W. Norton and Company.

LaPorta, R., F. Lopez-de-Silanes, A. Shleifer, and R. Vishny, (1997), "Trust in Large Organizations," *American Economic Review*, 87, 2, 333-8.

Leamer, E., (1983), "Let's Take the Con Out of Econometrics," *American Economic Review*, 73, 1, 31-43.

Leamer, E. and H. Leonard, (1983), "Reporting the Fragility of Regression Estimates," *Review of Economics and Statistics*, 65, 2, 306-17.

Lee, K., M. Pesaran, and R. Smith, (1997), "Growth and Convergence in Multi Country Empirical Stochastic Solow Model," *Journal of Applied Econometrics*, 12, 4, 357-92.

Lee, K., M. Pesaran, and R. Smith, (1998), "Growth Empirics: A Panel Data Approach: A Comment," *Quarterly Journal of Economics*, 113, 1, 319-23.

Levine, R., N. Loayza, and T. Beck, (2000), "Financial Intermediation and Growth: Causality and Causes," *Journal of Monetary Economics*, 46, 1, 31-77.

Levine, R. and D. Renelt, (1991), "Cross-Country Studies of Growth and Policy: Methodological, Conceptual, and Statistical Problems." *World Bank PRE Working Paper no. 608*.

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

Levine, R. and S. Zervos, (1993), "What We Have Learned About Policy and Growth from Cross-Country Regressions," *American Economic Review*, 83, 2, 426-430.

Levine, R. and S. Zervos, (1998), "Stock Markets, Banks and Economic Growth," *American Economic Review*, 88, 3, 537-558.

Li, H. and H.-F. Zou, (2002), "Inflation, Growth, and Income Distribution: A Cross Country Study," *Annals of Economics and Finance*, 3, 1, 85-101.

Li, Q. and D. Papell, (1999), "Convergence of International Output: Time Series Evidence for 16 Countries," *International Review of Economics and Finance*, 8, 267-280.

Lichtenberg, F., (1992), "R&D Investment and International Productivity Differences," *National Bureau of Economic Research Working Paper no. 4161*.

Liu, Z. and T. Stengos, (1999), "Non-Linearities in Cross Country Growth Regressions: A Semiparametric Approach," *Journal of Applied Econometrics*, 14, 5, 527-38.

Loayza, N. and R. Ranciere, (2002), "Financial Development, Financial Fragility, and Growth," *CESifo Working Paper Series no. 684*.

Loewy, M. and D. Papell, (1996), "Are U.S. Regional Incomes Converging? Some Further Evidence," *Journal of Monetary Economics*, 38, 3, 587-598.

Loh, W.-Y., (2002), "Regression Trees with Unbiased Variable Selection and Interaction Detection," *Statistica Sinica*, 12, 361-386.

Londregan, J. and K. Poole, (1990), "Poverty, the Coup Trap, and the Seizure of Executive Power," *World Politics*, 4, 2, 151-183.

Long, J. and L. Ervin, (2000), "Using Heteroscedasticity Consistent Standard Errors in the Linear Regression Model," *American Statistician*, 54, 3, 217-224.

Lucas, R., (1988), "On the Mechanics of Economic Development," *Journal of Monetary Economics*, 22, 1, 3-42.

Lundström, S., (2002), "On Institutions, Economic Growth, and the Environment," mimeo, Goteborg University.

Maasoumi, E., J. Racine, and T. Stengos, (2003), "Growth and Convergence: A Profile of Distribution Dynamics and Mobility," mimeo, Southern Methodist University.

MacKinnon, J. and H. White, (1985), "Some Heteroskedasticity-Consistent Covariance Matrix Estimators with Improved Finite Sample Properties," *Journal of Econometrics*, 29, 305-325.

Maddala, G. and S. Wu, (2000), "Cross-Country Growth Regressions: Problems of Heterogeneity, Stability, and Interpretation," *Applied Economics*, 32, 635-642.

Maddison, A., (1982), *Phases of Capitalist Development*. New York: Oxford University Press.

Maddison, A., (1989), *The World Economy in the 20th Century*, OECD: Paris.

Mamuneas, T., A. Savvides, and T. Stengos, (2004), "Economic Development and Return to Human Capital: A Smooth Coefficient Semiparametric Approach," mimeo, University of Guelph and forthcoming, *Journal of Applied Econometrics*.

Mankiw, N. G., (1995), "The Growth of Nations," *Brookings Papers on Economic Activity*, 1, 275-310.

Mankiw, N. G., D. Romer, and D. Weil, (1992), "A Contribution to the Empirics of Economic Growth," *Quarterly Journal of Economics*, 107, 2, 407-37.

Manski, C., (1993), "Identification of Endogenous Social Effects: The Reflection Problem," *Review of Economic Studies*, 60, 3, 531-42.

Masanjala, W. and C. Papageorgiou, (2004), "Rough and Lonely Road to Prosperity: A Reexamination of the Sources of Growth in Africa Using Bayesian Model Averaging," mimeo, Louisiana State University.

Masters, W. and M. McMillan, (2001), "Climate and Scale in Economic Growth," *Journal of Economic Growth*, 6, 3, 167-186.

Masters, W. and J. Sachs, (2001), "Climate and Development," mimeo, Purdue University.

Mauro, P., (1995), "Corruption and Growth," *Quarterly Journal of Economics*, 110, 3, 681-713.

McArthur, J. and J. Sachs, (2001), "Institutions and Geography: Comment on Acemoglu, Johnson and Robinson," *National Bureau of Economic Research Working Paper no. 8114*.

McCarthy, D., H. Wolf and Y. Wu, (2000), "The Growth Costs of Malaria," *National Bureau of Economic Research Working Paper no. 7541*.

McKenzie, D., (2001), "The Impact of Capital Controls on Growth Convergence," *Journal of Economic Development*, 26, 1, 1-24.

- Michelacci, C. and P. Zaffaroni, (2000), "Fractional (Beta) Convergence," *Journal of Monetary Economics*, 45, 129-153.
- Miguel, E., S. Satyanath, and E. Segenti, (2003), "Economic Shocks and Civil Conflict: An Instrumental Variables Approach," mimeo, UC Berkeley.
- Minier, J., (1998), "Democracy and Growth: Alternative Approaches," *Journal of Economic Growth*, 3, 3, 241-266.
- Mokyr, J., (1992), *Lever of Riches: Technological Creativity and Economic Progress*, Princeton: Princeton University Press.
- Motley, B., (1998). "Growth and Inflation: A Cross-Country Study." *FRBSF Economic Review*, 1.
- Murphy, K., A. Shleifer, and R. Vishny, (1991), "The Allocation of Talent: Implications for Growth," *Quarterly Journal of Economics*, 106, 2, 503-530.
- Nelson, M. and R. Singh, (1994), "The Deficit-Growth Connection: Some Recent Evidence from Developing Countries," *Economic Development and Cultural Change*, 42, 167-191.
- Nerlove, M., (1999), "Properties of Alternative Estimators of Dynamic Panel Models: An Empirical Analysis of Cross-Country Data for the Study of Economic Growth," in *Analysis of Panels and Limited Dependent Variable Models: In Honour of G. S. Maddala*, C. Hsiao, ed., Cambridge: Cambridge University Press.
- Nerlove, M., (2000), "Growth Rate Convergence, Fact or Artifact? An Essay on Panel Data Econometrics," in *Panel Data Econometrics: Future Directions: Papers in Honour of Professor Pietro Balestra*, E. Ronchetti, ed., Amsterdam: North-Holland.
- Nickell, S., (1981), "Biases in Dynamic Models with Fixed Effects," *Econometrica*, 49, 6, 1417-1426.
- Odedokun, M., (1996), "Alternative Econometric Approaches for Analysing the Role of the Financial Sector in Economic Growth: Time Series Evidence from LDCs," *Journal of Development Economics* 50, 1, 119-146.
- Olson, M., (1982), *The Rise and Decline of Nations*, New Haven: Yale University Press.
- Paap, R. and H. van Dijk, (1998), "Distribution and Mobility of Wealth of Nations," *European Economic Review*, 42, 7, 1269-93.
- Papageorgiou, C., (2002), "Trade as a Threshold Variable for Multiple Regimes," *Economics Letters*, 71, 1, 85-91.

- Papageorgiou, C. and W. Masanjala, (2004), "The Solow Model with CES Technology: Nonlinearities with Parameter Heterogeneity," *Journal of Applied Econometrics*, 19, 2, 171-201.
- Perron, P., (1989), "The Great Crash, the Oil Price Shock, and the Unit Root Hypothesis," *Econometrica*, 57, 6, 1361-1401.
- Persson, J., (1997), "Convergence Across Swedish Counties, 1911-1993," *European Economic Review*, 41, 1835-1852.
- Persson, T. and G. Tabellini, (1994), "Is Inequality Harmful for Growth?," *American Economic Review*, 84, 3, 600-621
- Persson, T. and G. Tabellini, (2003), *The Economic Effects of Constitutions*, Cambridge: MIT Press.
- Pesaran, M. H., (2004a), "A Pair-Wise Approach to Testing for Output and Growth Convergence," mimeo, University of Cambridge.
- Pesaran, M. H., (2004b), "General Diagnostic Tests for Cross-Section Dependence in Panels," mimeo, University of Cambridge.
- Pesaran, M. H., Y. Shin, and R. Smith, (1999), "Pooled Mean Group Estimation of Dynamic Heterogeneous Panels," *Journal of the American Statistical Association*, 94, 446, 621-34.
- Pesaran, M. H. and R. Smith, (1995), "Estimating Long-run Relationships from Dynamic Heterogeneous Panels," *Journal of Econometrics*, 68, 1, 79-113.
- Phillips, P. and D. Sul, (2002), "Dynamic Panel Estimation and Homogeneity Testing Under Cross-Section Dependence," *Cowles Foundation Discussion Paper no. 1362*.
- Phillips, P. and D. Sul, (2003), "The Elusive Empirical Shadow of Growth Convergence," *Cowles Foundation Discussion Paper no. 1398*.
- Prescott, E., (1998), "Needed: A Theory of Total Factor Productivity," *International Economic Review*, 39, 525-551.
- Pritchett, L., (1997), "Divergence, Big Time," *Journal of Economic Perspectives*, 11, 3, 3-17.
- Pritchett, L., (2000a), "Understanding Patterns of Economic Growth: Searching for Hills among Plateaus, Mountains, and Plains," *World Bank Economic Review*, 14, 2, 221-50.
- Pritchett, L., (2000b), "The Tyranny of Concepts: CUDIE (Cumulated, Depreciated, Investment Effort) Is Not Capital," *Journal of Economic Growth*, 5, 4, 361-84.

Putnam, R., R. Leonardi and R. Nanetti, (1993), *Making Democracy Work*, Princeton: Princeton University Press.

Rupasingha, A., S. Goetz, and D. Freshwater, (2000), "Social Capital and Economic Growth: A County-Level Analysis," *Journal of Agricultural and Applied Economics*, 32, 3, 565-72.

Quah, D., (1993a), "Galton's Fallacy and Tests of the Convergence Hypothesis," *Scandinavian Journal of Economics*, 95, 427-443.

Quah, D., (1993b), "Empirical Cross-Section Dynamics in Economic Growth," *European Economic Review*, 37, 2-3, 426-34.

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

Quah, D., (1996b), "Empirics for Economic Growth and Convergence," *European Economic Review*, 40, 6, 1353-75.

Quah, D., (1996c), "Convergence Empirics Across Economies with (Some) Capital Mobility," *Journal of Economic Growth*, 1, 1, 95-124.

Quah, D., (1997), "Empirics for Growth and Distribution: Stratification, Polarization, and Convergence Clubs," *Journal of Economic Growth*, 2, 1, 27-59.

Quah, D., (2001), "Searching for Prosperity: A Comment," *Carnegie-Rochester Conference Series on Public Policy*, 55, 305-19.

Ram, R., (1999), "Financial Development and Economic Growth," *The Journal of Development Studies*, 27, 2, 151-167.

Ramey, G. and V. Ramey, (1995), "Cross-Country Evidence on the Link Between Volatility and Growth," *American Economic Review*, 85, 5, 1138-1151.

Reichlin, L., (1999), "Discussion of 'Convergence as Distribution Dynamics', by Danny Quah," in *Market Integration, Regionalism, and the Global Economy*, R. Baldwin, D. Cohen, A. Sapir, and A. Venables, eds., Cambridge: Cambridge University Press.

Rivera, B. and L. Currais, (1999), "Economic Growth and Health: Direct Impact or Reverse Causation?," *Applied Economics Letters*, 6, 11, 761-64.

Robertson, D and J. Symons, (1992), "Some Strange Properties of Panel Data Estimators," *Journal of Applied Econometrics*, 7, 2, 175-89.

Rodriguez, F. and D. Rodrik, (2001), "Trade Policy and Economic Growth: A User's Guide," in *Macroeconomics Annual 2000*, B. Bernanke and K. Rogoff, eds., Cambridge: MIT Press.

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

Rodrik, D. (ed.), (2003), *In Search of Prosperity: Analytic Narratives on Economic Growth*, Princeton: Princeton University Press.

Rodrik, D., A. Subramanian, and F. Trebbi, (2004), "Institutions Rule: The Primacy of Institutions Over Geography and Integration in Economic Development," *Journal of Economic Growth*, 9, 2, 131-165.

Romer, D., (2001), *Advanced Macroeconomics*, New York: McGraw-Hill.

Romer, P., (1986), "Increasing Returns and Long-run Growth," *Journal of Political Economy*, 94, 5, 1002-1037.

Romer, P., (1990), "Human Capital and Growth: Theory and Evidence," *Carnegie-Rochester Series on Public Policy*, 32, 251-286.

Romer, P., (1993), "Idea Gaps and Object Gaps in Economic Development," *Journal of Monetary Economics*, 32, 3, 543-573.

Roubini, N. and X. Sala-i-Martin, (1992), "Financial Repression and Economic Growth," *Journal of Development Economics*, 39, 5-30.

Roubini, N. and X. Sala-i-Martin, (1992), "Financial Repression and Economic Performance: Historical Evidence from Five Industrialized Countries," *Journal of Money, Credit and Banking*, 30, 4, 657-678.

Rousseau, P., (2002), "Historical Perspectives on Financial Development and Economic Growth," *Review Federal Reserve Bank of St. Louis*, 84, 4.

Rousseau, P. and R. Sylla, (2001), "Financial Systems, Economic Growth, and Globalization," mimeo, Vanderbilt University.

Rousseau, P. and P. Wachtel, (2002), "Inflation Thresholds and the Finance-Growth Nexus," *Journal of International Money and Finance*, 21, 6, 77-793.

Sachs, J., (2003), "Institutions Don't Rule: Direct Effects of Geography on Per Capita Income," *National Bureau of Economic Research Working Paper no. 9490*.

Sachs, J. and A. Warner, (1995), "Economic Reform and the Process of Global Integration (with discussion)," *Brookings Papers on Economic Activity*, 1, 1-118.

Sachs, J and A. Warner, (1996), "Natural Resource Abundance and Economic Growth," *National Bureau of Economic Research Working Paper no. 5398*.

Sala-i-Martin, X., (1991), "Growth, Macroeconomics, and Development: Comments," in *Macroeconomics Annual 1991*, O. Blanchard and S. Fischer, eds., Cambridge: MIT Press.

Sala-i-Martin, X., (1996a), "The Classical Approach to Convergence Analysis," *Economic Journal*, 106, 1019-1036.

Sala-i-Martin, X., (1996b), "Regional Cohesion: Evidence and Theories of Regional Growth and Convergence," *European Economic Review*, 40, 1325-1352.

Sala-i-Martin, X., (1997a), "I Just Ran 4 Million Regressions," *National Bureau of Economic Research Working Paper no. 6252*.

Sala-i-Martin, X., (1997b), "I Just Ran 2 Million Regressions." *American Economic Review*, 87, 2, 178-83.

Shioji, E., (2001a), "Composition Effect of Migration and Regional Growth in Japan," *Journal of the Japanese and International Economies*, 15, 29-49.

Shioji, E., (2001b), "Public Capital and Economic Growth: A Convergence Approach," *Journal of Economic Growth*, 6, 3, 205-227.

Solow, R., (1956), "A Contribution to the Theory of Economic Growth," *Quarterly Journal of Economics*, 70, 1, 65-94.

Solow, R., (1994), "Perspectives on Growth Theory," *Journal of Economic Perspectives*, 8, 45-54.

Summers, R. and A. Heston, (1988), "A New Set of International Comparisons of Real Product and Price Levels Estimates for 130 Countries, 1950-1985," *Review of Income and Wealth*, 34, 1-25.

Summers, R. and A. Heston, (1991), "The Penn World Table (Mark 5): An Expanded Set of International Comparisons, 1950-1988," *Quarterly Journal of Economics*, 106, 2, 327-368.

Swan, T., (1956), "Economic Growth and Capital Accumulation," *Economic Record*, 32, 334-361.

Swank, D., (1996), "Culture, Institutions, and Economic Growth," *American Journal of Political Science*, 40, 660-79.

Swartz, S. and R. Welsch, (1986), "Applications of Bounded-Influence and Diagnostic Methods in Energy Modeling," in *Model Reliability*, D. Belsley, and E. Kuh, eds., Cambridge: MIT Press.

Tan, C. M., (2004), "No One True Path to Development: Uncovering the Interplay Between Geography, Institutions, and Ethnic Fractionalization in Economic Development," mimeo, Tufts University.

Tavares, J. and R. Wacziarg, (2001), "How Democracy Affects Growth," *European Economic Review*, 45, 8, 1341-78.

Taylor, C., (1998), *Socrates*, New York: Oxford University Press.

Temple, J., (1998), "Robustness Tests of the Augmented Solow Model," *Journal of Applied Econometrics*, 13, 4, 361-75.

Temple, J., (1999), "The New Growth Evidence," *Journal of Economic Literature*, 37, 1, 112-56.

Temple, J., (2000a), "Inflation and Growth: Stories Short and Tall," *Journal of Economic Surveys*, 14, 4, 395-426.

Temple, J., (2000b), "Growth Regressions and What the Textbooks Don't Tell You," *Bulletin of Economic Research*, 52, 3, 181-205.

Temple, J., (2003), "The Long-run Implications of Growth Theories," *Journal of Economic Surveys*, 17, 3, 497-510.

Temple, J. and P. Johnson, (1998), "Social Capability and Economic Growth," *Quarterly Journal of Economics*, 113, 3, 965-90.

Toya, H., M. Skidmore, and R. Robertson, (2003), "Why are Estimates of Human Capital's Contribution to Growth So Small," mimeo, Nagoya City University.

Wacziarg, R., (2002), "Review of Easterly's The Elusive Quest for Growth," *Journal of Economic Literature*, 40, 3, 907-18.

Wacziarg, R. and K. Welch, (2003), "Trade Liberalization and Growth: New Evidence," *National Bureau of Economic Research Working Paper no. 10152*.

Warner, A., (1992), "Did the Debt Crisis Cause the Investment Crisis?," *Quarterly Journal of Economics*, 107, 4, 1161-86.

Welsch, H., (2003), "Corruption, Growth and the Environment: A Cross Country Analysis," mimeo, German Institute for Economic Research.

White, H., (1980), "A Heteroskedastic-Consistent Covariance Matrix Estimator and a Direct Test for Heteroskedasticity," *Econometrica*, 48, 817-838.

Whiteley, P. (2000), "Economic Growth and Social Capital," *Political Studies*, 48, 443-466.

Zak, P. and S. Knack, (2001), "Trust and Growth," *Economic Journal*, 111, 295-321.

Zietz, J., (2001), "Heteroskedasticity and Neglected Parameter Heterogeneity," *Oxford Bulletin of Economics and Statistics*, 63, 2, 263-73.