Journal of Economic Literature
Vol. XXXVII (March 1999), pp. 112–156

The New Growth Evidence

JONATHAN TEMPLE[1]

1. *Introduction*

WHY HAVE SOME countries grown rich while others remain poor? It is hard to think of a more fundamental question for economists to answer. According to the definition used by the World Bank's 1996 *World Development Report,* over 4.5 billion of the world's 5.6 billion people live in developing countries, and so a better understanding of what generates economic growth could make a huge contribution to human welfare. Yet macroeconomists have traditionally shown little interest in the gulf between rich and poor. The study of growth at the aggregate level has often been something of a backwater, relegated to a brief last chapter in mainstream textbooks, and rarely taken on by anyone outside development economics.

There are at least two reasons for this state of affairs. One is that, until recently, lack of data made it hard to compare income levels across a wide range of countries. A second reason is more subtle, and turns on a common interpretation of the early theoretical literature. The theoretical work of the mid-1950s suggested that growth was ultimately driven by technical advances, and it was widely felt that these advances would be resistant to further analysis. Only rarely did anyone consider that even these early models, with their simple assumptions about the exogeneity of technical progress, might be used to help understand the wide international variation in levels of income and growth rates.

Thirty years later, macroeconomists returned to growth issues, spurred by the availability of the Summers-Heston data set, and by the work of William Baumol (1986), Robert Lucas (1988) and Paul Romer (1986). Although it was new growth theory that initially drew the most attention, the last ten years have also seen an outpouring of empirical work intended to explain post-1960 growth experiences. This work, the "new growth evidence," is the subject of what follows.

A common view of this research is that, somewhat ironically, it has mostly demonstrated the explanatory power of the original 1950s-style neoclassical models. I believe that this interpretation misses a great deal, and one aim of this essay is to argue for a different view. Another aim, perhaps no less central, is to examine why this crucial area of economic investigation is regarded with such widespread distrust. Empirical work on growth has often been controversial, and there is a widespread feeling that growth theory and

[1] Hertford College, and Institute of Economics and Statistics, Oxford University. I am grateful to Philippe Aghion, Janine Aron, Tony Atkinson, John McMillan, John Muellbauer, Steve Nickell, Steve Redding, and the referees for helpful contributions. I would especially like to thank John Pencavel, who played a major role in shaping the paper. The ESRC and Hertford College, Oxford provided financial support.

econometrics are best kept apart. Their offspring, cross-country regressions, are not greatly loved by either parent.

It is not difficult to see why this might be the case. The early papers present some easy targets, and basic points are often misunderstood. However, some frequent criticisms are easily answered. Too many people have dismissed the prospects for cross-country research, ignoring recent improvements and greater awareness of econometric issues. I will argue that we have already learned some interesting things.

The focus of the paper is predominantly the cross-country empirical work carried out by macroeconomists. This work has three distinguishing features. First, researchers have often tried to integrate developing and developed countries in a single empirical framework; as we shall see, this endeavour is not without its problems. Second, the research makes intensive use of the cross-section variation in growth rates and other variables. Finally, the research questions are often inspired at least in part by recent growth theory. There is a renewed emphasis on human capital, and to a lesser extent, research and development (R&D), as important variables in explaining differing growth experiences.

Work sharing these features has been used to address a wide variety of questions. Among the many issues of interest, I choose to emphasize six:

(Q1) How is the world income distribution evolving?

(Q2) Do countries converge to steady state paths and, if so, how quickly?

(Q3) How rapidly do returns to inputs like physical capital diminish?

(Q4) Are poor countries poor mainly because they lack inputs, or because of technology differences?

(Q5) Why do growth rates differ over long periods?

(Q6) What happens in the long run?

The first question covers such widely-discussed issues as whether poor countries are catching up with the rich. I will discuss possible answers in Section 2. To place the discussion in context, I also discuss the purpose and nature of the Summers-Heston data set, and some subtleties in measuring output levels and growth rates.

In answering the other five questions, the new growth evidence makes use of a variety of methods, many of them controversial. Much of the first half of the survey is taken up by a discussion of the possible approaches. Section 3 contrasts cross-country growth regressions with the older methods, historical case studies and growth accounting, and builds the case for using the cross-country variation in the data. Next, I turn to the many econometric problems with growth regressions (Section 4) and to approaches that may overcome some of these problems (Section 5). An underlying argument will be that the use of panels is often the best way forward, though not without problems of its own.

The second half of the survey is more directly concerned with what we have learned from this research. Section 6 gathers together some of the issues often discussed under the heading "convergence." It makes good sense to treat these issues together, as answers to one question are usually relevant to the next. For instance, measurement of the rate of convergence is often used to assess the extent of diminishing returns, which then feeds into the debate about the role of technology differences, which then affects our view of growth differentials and the likely nature of long-run growth.

For practical ends, the main question to answer is (Q5): why do growth rates differ over long periods? Our knowledge is incomplete, but a synthesis of recent contributions is provided in Sections 7 and 8. Although growth differentials have been a central focus of the literature, it is clear that there is more to explain, and issues needing further study are the subject of Section 9. Finally, the conclusions give a personal view of the most reliable findings.

One aspect of my approach should already be clear. Since the recent literature has partly been spurred by theoretical developments, where relevant I will relate findings to growth models. However, I will not take the attitude that distinguishing between specific growth models should be the main aim of the empirical literature. Anything that might contribute to better policy-making should be of interest, and the coverage here reflects that.

2. *Preliminaries*

In thinking about growth, it helps to begin by trying to establish some stylized facts. This section draws on the work of many researchers to highlight some of the most interesting regularities in the data. One advantage of this approach is that important measurement issues can be introduced in what should be a relatively painless way.

2.1 *Income Disparities*

The most striking aspect of the world income distribution is the very large and persistent disparities in per capita income. However, one has to be rather careful in making such comparisons. The obvious method is to value each country's quantities of final goods and services at domestic prices, and then convert these figures into a common monetary unit using the relevant exchange rates.

However, observant travellers know that a given sum of, say, dollars will buy very different amounts of goods in different countries. In other words, there are often large and systematic departures of exchange rates from "purchasing power parity" (PPP). Instead of using exchange rates, incomes should be converted using special currency indexes which are calculated so that one unit will purchase the same bundle of goods across countries.

The calculation and use of these indexes, called purchasing power parities or PPPs, is essential for accurate cross-country comparisons of real incomes and expenditures. To give some idea of the difference this makes, consider a comparison between the incomes of the US and India. If we convert Indian per capita GDP into US dollars at the official exchange rate, this tends to suggest that India's average income is just 2 percent of the USA's. If we use PPPs, India's relative position is improved by a factor of two and a half, so that average income is in fact 5 percent of the USA's. It turns out that comparison using exchange rates tends to overstate the magnitude of income disparities.[2]

The construction of a world table of national accounts with figures that are comparable across space and time, based on PPPs, is clearly a difficult and research-intensive exercise. It relies on obtaining price data for a wide range of goods, and then devising suitable aggregation procedures to obtain a national PPP. The United Nations International Comparison Project (ICP), launched in the late 1960s, was designed to make such comparisons possible. Over 90

[2] Another way of saying this is that price levels are higher in richer countries. For a very clear survey of possible explanations, see Kenneth Rogoff (1996).

TABLE 1
GDP RELATIVE TO THE US, SELECTED COUNTRIES, 1960 AND 1990

	Population 1990	Proportion of US GDP per capita (at current prices) 1960	1990
Zaire	36m	0.05	0.02
Nigeria	96m	0.05	0.05
India	850m	0.07	0.07
China	1134m	0.05	0.07
Bangladesh	108m	0.09	0.08
Pakistan	112m	0.07	0.08
The Philippines	61m	0.12	0.10
Indonesia	178m	0.06	0.11
Iran	56m	0.23	0.16
South Africa	38m	0.21	0.18
Thailand	56m	0.10	0.20
Brazil	149m	0.18	0.22
Argentina	32m	0.44	0.25
Mexico	81m	0.28	0.32
Korea	43m	0.09	0.38
Japan	124m	0.30	0.81

Note: Most recent figures for Zaire from 1989.

countries have participated in benchmark studies, which are then used to derive an aggregate PPP for each participating economy.

These estimates have then been combined with national accounts data to form the Penn World Table (PWT). This table, often known as the Summers-Heston data set, has been used by empirical growth researchers since the mid-1980s. The data set has been discussed most recently by Alan Heston and Robert Summers (1996), and is described in more detail in Summers and Heston (1988, 1991).[3]

Table 1 displays some useful information calculated from version 5.6 of the Penn World Table. For a selection of countries with large populations that might have been considered developing in 1960, the table shows their incomes relative to the US for 1960 and 1990, ranked by the 1990 figure. The table makes clear that, even after making PPP adjustments, there are enormously large disparities in average living standards. Zaire's per capita income in 1990 was less than a fortieth of the USA's, while the huge populations of India and China have average incomes rather less than a tenth of the USA's.

Remember, too, that these figures are the *average* income per head: many in Zaire have much less than 2 percent of the average US income to live on. Supplementing the raw figures with other pieces of evidence gives some idea of the human cost. The typical African mother has only a 30 percent chance of seeing all her children survive to age 5. The life expectancy of a person born in Sub-Saharan Africa in 1980 is just 48 years (William Easterly and Ross Levine 1997a). In the

[3] For a more detailed exposition of the methods of the ICP, including an assessment of its achievement, see the article by one of its originators, Irving Kravis (1984).

TABLE 2
GROWTH MIRACLES AND DISASTERS, 1960–90
ANNUAL GROWTH RATES OF OUTPUT PER WORKER

Miracles	Growth	Disasters	Growth
Korea	6.1	Ghana	–0.3
Botswana	5.9	Venezuela	–0.5
Hong Kong	5.8	Mozambique	–0.7
Taiwan	5.8	Nicaragua	–0.7
Singapore	5.4	Mauritania	–0.8
Japan	5.2	Zambia	–0.8
Malta	4.8	Mali	–1.0
Cyprus	4.4	Madagascar	–1.3
Seychelles	4.4	Chad	–1.7
Lesotho	4.4	Guyana	–2.1

Note: Figures for Botswana and Malta based on 1960–89.

circumstances, it is not surprising that we should want to know as much as possible about how countries can grow more quickly.

We can extract some other useful information from Table 1. One interesting fact is that most of these large developing countries roughly maintained their positions relative to the USA between 1960 and 1990. Since the USA's per capita income grew at around 2 percent a year over this period, this indicates that these countries have grown at a similar rate. There does not appear to be an absolute poverty trap.

2.2 *Growth Miracles and Disasters*

Although there is considerable persistence in relative positions, Table 1 also hints at the possible variety of growth experience. There have been "growth disasters," countries in which per capita income has fallen since 1960, and "growth miracles" like Japan and Korea, countries which have risen rapidly up the world income distribution.

One of the first questions we might ask is whether there is a systematic tendency for poor countries to grow faster than rich ones, which would allow them to catch up. Figure 1 shows the growth of real GDP per worker over 1960–90 against its initial level, a familiar diagram in the growth literature. If countries are converging, one would expect to see a negative slope.[4] In fact, like Table 1, the figure shows that there is no general tendency for countries to converge to a common level of per capita income.

Some interesting patterns emerge from Figure 1. Within the group of countries that were relatively well off by 1960, those with 1960 GDP per worker greater than $10,000 in 1985 "international dollars," there has been some convergence.[5] Within the group of poorer countries, there has been a greater variety of experience: some have done very well and others very badly.

Table 2 lists the ten fastest and the ten slowest growers among the countries in the Summers-Heston data set, using data on output per worker. The table demonstrates a regional pattern familiar to growth researchers: many of the fastest growing countries are in East Asia, many of the slowest in Sub-Saharan Africa. The strength of this pattern suggests that growth rates are quite persistent over time, but this is misleading. For the sample as a whole, the correlation between growth in GDP per worker over 1960–75 and that over 1975–90 is just 0.17. Relatively few countries have done consistently well, and averaging over long periods tends to obscure the episodic nature of growth, particularly in Sub-Saharan Africa. Frequently countries have done well for short periods, only for growth

[4] Xavier Sala-i-Martin (1996) discusses convergence concepts in more depth.

[5] The "international dollar" is the currency unit used in the Penn World Table.

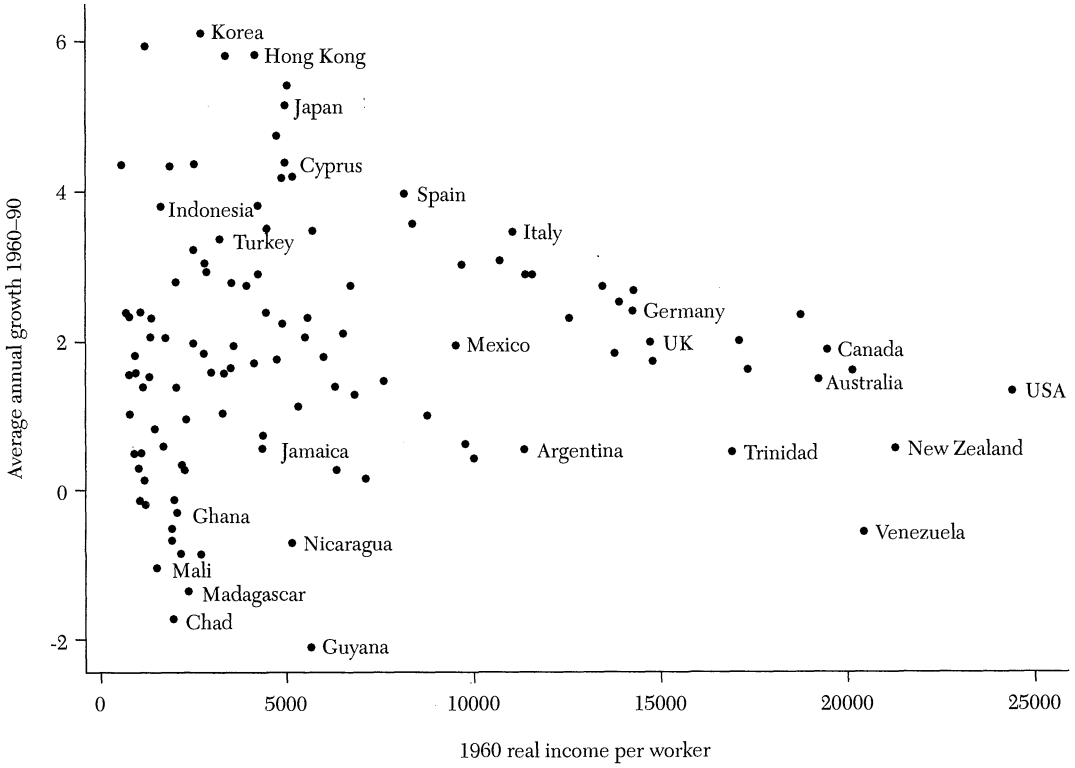

Figure 1. Growth and initial income, 1960–90

to collapse later on. This has important implications for growth studies, which I will return to later on.

Overall, the figure and tables suggest that many of the most important poor countries do not seem to be catching up to the USA's level of income. Instead, countries have roughly maintained their place within the world income distribution over the last thirty years or so, with little tendency for reduced income dispersion, and perhaps even some divergence.[6] The exception to this stability is some of the countries that were middle income around 1960; some of these have joined the relatively well off. Most of the countries in Sub-Saharan Africa have done very badly.

These stylized facts might lead one to think about theories that emphasize relative development traps and multiple equilibria, but it is always worth remembering that there are often simpler explanations, not least low levels of factor accumulation among some countries. To distinguish between theories we clearly need more sophisticated methods. Sections 3, 4, and 5 will consider the pros and cons of those usually adopted, before Section 6 returns to convergence issues in depth.

Before reviewing more complex empirical work, however, it is worth assessing two of its foundations. These are

[6] Space precludes a more detailed picture, which can be found in the recent work of Charles Jones (1997a,b) and Danny Quah (1993, 1997). The presentation here follows most of the literature in analysing post-1960 experience; for a longer view see Lant Pritchett (1997).

the measurement of output and the calculation of growth rates. Both involve subtleties that some growth researchers have missed. Lack of space precludes a full treatment, but the coverage here at least points towards useful further reading.

2.3 *The Quality of the Output Data*

Nobody disputes that there are very large income disparities, but the accuracy of precise comparisons is often questioned. Given the difficulty of the task facing the compilers of the PWT, it is not surprising that worries have sometimes been expressed about some of the methods used and the reliability of the final data (for instance, Robin Marris 1984).

For those countries that have not participated in ICP benchmarks, and in which the PWT estimates are based on extrapolations, the quality of the data is likely to be particularly low.[7] Also worrying is the quality of the underlying national accounts data for certain developing countries. Heston (1994) quotes confidence intervals for estimated growth rates that are wide enough to be somewhat dispiriting.

More recently, Heston and Summers (1996) have pointed to two major concerns. One is that data on capital service lives is hard to come by, which makes it difficult to arrive at accurate figures for net investment and capital stocks. A second problem is the quality of available data on labor force participation and working hours. Heston and Summers write that participation rates "vary enough to make GDP per capita a very unsatisfactory proxy for GDP per worker" (1996, p. 23).[8] As an example, Japan's GDP per capita is around 80 percent of the USA's, according to the PWT, but its GDP per worker is just 60 percent.

Growth theories are essentially concerned with output per worker hour, but for developing countries in particular, the available data on worker hours are weak. Holger Wolf (1994) discusses the implications for the convergence literature, and argues that the problem should not be neglected. It is clear that we need either better data on working hours and participation or, at the very least, a framework for relating them to more easily observable variables.

Other problems include the likelihood that measured output per worker hour will vary with changes in the age structure of a country's population (Michael Sarel 1995) and with price distortions (Jagdish Bhagwati and Bent Hansen 1972). Finally, the presence of a sizeable nonmarket sector may lead to the output of developing countries being understated.

2.4 *Measuring Growth*

The Summers-Heston data set was primarily intended as a means of comparing income levels. It is widely used to derive growth rates as well, but there is noticeable confusion about whether or not this is a good idea. Some researchers are unaware that growth rates obtained from the Penn World Table will usually differ from those implicit in countries' own national accounts.

Both sets of growth rates are weighted averages of growth in the components of GDP, but the PWT weights are the real GDP shares of the components—based on an average of international prices—rather than the nominal shares used in domestic national accounts. If a country's price structure changes much between the initial and final years, the growth rate

[7] Summers and Heston use a grading system to highlight the worst cases.

[8] In an earlier article, Heston and Summers (1988) note the low rank correlation between GDP per capita and GDP per worker hour for 12 OECD countries.

based on domestic prices will differ from the Summers-Heston growth rate.

Daniel Nuxoll (1994) argues that using domestic prices to measure growth rates is more reliable, because the domestic prices are the ones reflecting the trade-offs that agents actually face. He concludes that "probably the ideal is to use Penn World Table numbers for levels and the usual national-accounts data for growth rates" (Nuxoll 1994, p. 1434). In general, the decision to use domestic or international prices depends on the hypothesis being tested. As Nuxoll points out, too little thought is given to this question at present.

Once one has decided on a set of output measures, the next question is how to calculate the growth rate. The usual method uses only initial and final output. Since either of these may be some distance from the trend path of output, it may well be preferable to use the least squares growth rate, obtained by regressing the whole of the log output series on a constant and a time trend. This should be more robust to short-run instability, such as business cycle effects.[9]

Talking about measuring "the trend growth rate" presupposes that the trend is deterministic rather than stochastic. For the most part, growth researchers have been content to assume that output is trend stationary. One justification might be this. When unit root tests are used country by country, the presence of a unit root is typically rejected for some of them. It seems unlikely that the process generating output is so different across countries that shocks have a permanent effect in some and only a temporary one in others.[10]

[9] It might be useful to go a step further, by including dynamics in the regression. For more on growth rate measurement see Eugene Canjels and Mark Watson (1997) and Nanak Kakwani (1997).

[10] The use of panel data unit root tests should be increasingly informative on this kind of question.

3. *Cross-Country Research*

Now that we have briefly covered data issues, and gained some impression of recent growth experiences, it is time to dig a little deeper, and start to identify reasons why those experiences have been so widely divergent. The ultimate aim of this literature is to understand why growth rates differ, and which policy measures will be effective in raising growth.

In this section, I turn to consideration of the methods that might be used to address these questions. The most popular method has been cross-section growth regressions, combining data from the Penn World Table with other variables. Despite the popularity of this endeavour, many believe that it is fruitless, partly because there are not likely to be any general answers. The appropriate research questions and policies will depend on a country's particular situation, for instance whether a country is a technological leader, or a developing country trying to catch up.

Historical studies are likely to be far more sensitive to these issues, and I first consider their particular merits as a way of understanding growth. This is followed by a comparison of growth accounting and cross-country regressions. In keeping with my underlying argument that careful growth regressions can play a valuable role, I then review various possible specifications. One or two problems are highlighted in passing, but detailed consideration of the econometric problems will be postponed until Section 4.

3.1 *The Role of Historical Studies*

For many people, statistical research on growth seems rather cruder and less informative than historical case studies. Certainly it is important to remember that growth regressions will never

offer a complete account of the growth process, and that historical analyses must have an important complementary role. Historians bring to bear a much deeper conception of the social, political, institutional and technological sources of growth than theoretical and empirical economists are usually able to incorporate in formal models.

Equally, the limitations of historical treatments should not be ignored. One of the pioneers of growth research, Simon Kuznets, argued that the isolated study of just one or two countries provides only a partial view (Kuznets 1966, p. 32). It is also worth quoting another influential economic historian, Alexander Gerschenkron:

> The historians' contribution consists in pointing at *potentially* relevant factors and at *potentially* significant contributions among them which could not be easily perceived within a more limited sphere of experience. (Gerschenkron 1952, p. 4, emphasis in original)

This seems to me exactly right: historians can usefully point to particular factors that others are likely to miss. An unspoken corollary is that statistical work, perhaps using cross-section variation, is often necessary to quantify the importance of the "potentially relevant" factors. Another way of saying the same thing is that case studies are a useful means of generating hypotheses, but econometrics is needed to test the validity of generalizations (Angus Deaton and Ron Miller 1996). The two methods are complementary in other ways: Gustav Ranis (1984) points out that the "average pattern" is only the beginning of wisdom, but knowledge of this pattern can inform one's choices of the countries to study in greater detail.

The problems with statistical work that are discussed below might lead one to disillusion, and a general preference for history. Before becoming too enthusiastic, it is worth remembering that econometric problems such as endogeneity and omitted variables have close equivalents in historical studies, usually with even less hope of a solution. In the conclusions, I consider the role of historical case studies as one of several possible ways forward.

3.2 *Growth Accounting versus Growth Regressions*

In some contexts, another alternative to cross-country regressions is growth accounting. When investigating why growth rates differ, a common starting point is to think about the relative contributions of growth in inputs, and growth in efficiency or total factor productivity (TFP). Then one might think about separately modelling input growth and TFP growth. Traditionally, though, growth accountants have only made progress on the first issue. A familiar response is that estimates of TFP growth are frequently no more than a measure of our ignorance. They bring us little closer to an understanding of why TFP growth rates might differ across time and space.

It is important to understand that one aim of running growth regressions is to use the cross-country variation in TFP growth to understand its determinants. However, the TFP growth rates are derived somewhat differently from growth accounting. The cross-country variation in the data is used to estimate the parameters of an "average" production function, instead of imposing these parameters based on factor shares or micro evidence.

If we are to understand the determinants of TFP growth, use of its variation across disparate units (countries, regions) seems pretty much essential. Whatever method of deriving TFP growth one prefers, at some point one will need to think about cross-section or

panel data econometrics. In this respect, the approaches used in the new growth evidence are simply reflecting the inevitable. The idea of returning to single-country growth accounting as the main mode of investigation is not an enticing one, especially when one is confronted with the sparse data of developing countries.

Where one may depart from the recent literature is in the derivation of TFP growth. Should the elasticities of output to inputs be imposed, or estimated? Nearly all of the work surveyed below implicitly estimates them, by including factor accumulation in growth regressions. One reason lies in recent theory. When the parameters are imposed, typically one assumes perfect competition, constant returns to scale, and the absence of externalities. All three assumptions have been questioned, often convincingly, by new growth theorists.

Yet the remainder of this paper will reveal many difficulties in estimating technology parameters. If we are willing to ignore externalities to physical and human capital, which are quite possibly negligible, one might well ask why we choose estimation. We could set these parameters based on micro data or factor shares, and thus avoid all the problems connected with estimating production functions. In time, this may turn out to be the best method of obtaining comparable TFP growth figures for a large number of countries. The point remains that cross-country studies will still be needed to understand the determinants of TFP growth.

Finally, one might ask whether the initial decomposition into input and TFP contributions is always a useful one. Researchers may well be right to attribute a high degree of the cross-country variation in growth rates to the variation of inputs, like physical capital investment. In a sense, however, this does not get us very far. It merely pushes the demanding question down a rung, from explaining growth to explaining investment. As I will discuss in Section 9, we need better explanations of the cross-country variation in investment ratios.

In practice, whether or not we seek to decompose output growth into the contributions of inputs and TFP should depend on the research question. Often we are interested in the overall growth effect of a policy outcome like inflation or the budget deficit; whether or not it acts through factor accumulation or TFP growth is of secondary importance, at least initially. Hence the approach of some recent papers, though perhaps too few, is to use a regression model which simply relates growth to policy outcomes, using a model sufficiently general that it should account for most influences on factor accumulation as well as on TFP. This approach often has much to recommend it.

3.3 *Specifying Growth Regressions*

Overall, then, cross-country empirical work will continue to form a useful complement to more traditional approaches, depending on the particular questions that are being asked. Cross-country work allows differences in productivity growth to be explained, unlike conventional growth accounting. If well done, this work can help identify the relative contributions of different influences more precisely than historical studies. Given that there is a clear place for cross-country work, in this section I turn to the specification of growth regressions.

First, though, a warning is necessary. Many of the issues require technical discussion, and this discussion presumes a good knowledge of econometrics and an eagerness to get to grips

with methodological issues. Those readers who want only a view of the forest, without worrying too much about the trees, should cover the remainder of this section only briefly, and then skip to the summary of research findings, which begins with Section 6.

Before then, I will discuss several approaches in turn, starting with the framework of N. Gregory Mankiw, David Romer and David N. Weil (1992), henceforth MRW. The exposition should clarify some of the underlying ideas that almost all growth regressions hold in common. It can also be used to understand the usual selection of variables in equations that are not explicitly derived from a theoretical model, and this less formal approach I will discuss in Section 3.4.

Hence the framework introduced by MRW is a useful starting point.[11] They start from a Cobb-Douglas production function with constant returns to scale,

$$Y = K^{\alpha}H^{\beta}(AL)^{1-\alpha-\beta} \qquad (1)$$

where K is physical capital, H is human capital, L is labour supply and A is an index of technical efficiency. MRW assume that investment rates in physical and human capital are constant at s_k and s_h respectively, and that both types of capital depreciate at a common rate δ. Technical efficiency grows at the same exogenous rate g across countries, while the labour force grows at differing rates n. This is clearly just the Solow model augmented with human capital and an assumption that countries share the same rate of efficiency growth. The initial level of efficiency, $A(0)$, is assumed to vary randomly across countries (due to local factors like climate) and this can be used to justify the error term. Mankiw, Romer and Weil show, by approximating around the steady state, that growth in this model is given by:

$$\ln\frac{Y(t)}{L(t)} - \ln\frac{Y(0)}{L(0)} = \theta\ln A(0) + gt + \theta\frac{\alpha}{1-\alpha-\beta}\ln s_k + \theta\frac{\beta}{1-\alpha-\beta}\ln s_h - \theta\frac{\alpha+\beta}{1-\alpha-\beta}\ln(n+g+\delta) - \theta\ln\frac{Y(0)}{L(0)} + \varepsilon \qquad (2)$$

where $\theta = 1 - e^{-\lambda t}$ and λ is the rate of convergence to a country's steady state, a measure of how fast countries attain their long-run equilibrium path.[12] The measure is defined by

$$\frac{d\ln y(t)}{dt} = \lambda[\ln y^* - \ln y(t)]$$

and the convergence rate is related to other variables by

$$\lambda = (n+g+\delta)(1-\alpha-\beta) \qquad (3)$$

In implementing the equation empirically, there are one or two subtleties. In practice investment rates are not constant, so MRW average them over the period. Another point to note is that n and δ are likely to vary across countries, and so too will the rate of convergence. Conventional estimation of the model ignores this subtlety, and we shall return to this later on.[13]

Why does initial income affect growth in the above equation? The negative coefficient implies that, if we take two countries with the same rates of investment and the same level of efficiency, the poorer one will grow more quickly for a transitional period.

[11] More general treatments, particularly useful for specialists, can be found in Steven Durlauf and Quah (1998) and Jeffrey Sachs and Andrew M. Warner (1997).

[12] How good is the approximation used to derive these results? David Romer (1996, p. 23) argues that the approximation is likely to be fairly reliable for the changes in parameter values typically considered.

[13] Note also that the convergence equation is obtained by taking a Taylor series approximation around a deterministic steady state. Michael Binder and M. Hashem Pesaran (1996) argue that, when growth is stochastic, this approach to deriving convergence paths is potentially misleading.

The reason for these "transitional dynamics" is that the relatively poor economy must have lower stocks of physical and human capital. Hence the marginal product of extra capital is higher in this economy, and for a given rate of investment its growth will be faster.

One consequence is that if we run a regression which controls for the determinants of steady states, like investment ratios, then initial income will take a negative sign. This is the "conditional convergence" result. The convergence may be conditional, in that it is only apparent when we take into account the variation across countries in steady state levels of income. Hence the result does not imply that poorer countries will catch up with rich ones. Indeed, the Solow model is perfectly compatible with income divergence: anything that drives apart investment rates in rich and poor countries will tend to lead to increased income dispersion.[14]

Given the role of transitional dynamics, it is widely agreed that growth regressions should control for the steady state level of income. Note, though, an important consequence: in principle the level of technology A should be included in the regression. Since this variable is unobserved, it has to be omitted. As with any other omitted variable problem, the other parameter estimates are biased if one or more regressors are correlated with the level of technology. In practice, countries that are relatively less efficient are also likely to have lower investment rates, and one can easily imagine further correlations with other right-hand-side variables.

In the absence of a suitable proxy for A, the only way to obtain consistent estimates of a conditional convergence regression is to use panel data methods. Since initial efficiency is an omitted variable that is constant over time, it can be treated as a fixed effect, and the time dimension of a panel used to eliminate its influence. The advantages of panel data approaches will be discussed in Section 5.

Finally, note that we could also use the MRW approach to explain the variation of income levels, rather than growth rates. Robert Hall and Jones (1997) have argued that this may be a more natural research question, since ultimately we are interested in growth rates mainly because of their impact on levels. A crucial drawback is that, in the likely absence of good instruments, we cannot explain the cross-section variation in income using any variables which are endogenous to the level of income, like the development of physical infrastructure or the financial system. Hence 'levels accounting' tends to be a little short of implications for policy, and growth regressions retain the advantage that they can be used to study a wider range of variables.

3.4 *Informal Growth Regressions*

MRW's work provides a simple theoretical framework for growth regressions. The most common approach, though, is simply to use a more or less *ad hoc* regression, driven in its specification mainly by previous results in the literature. Typically such a regression will include the investment ratio, initial income, and measures of policy outcomes like school enrollment and the black market exchange rate premium.[15]

[14] It is also worth noting that conditional convergence does not imply a rejection of "Ak" models of growth. On this see Narayana Kocherlakota and Kei-Mu Yi (1995) and Charles Leung and Quah (1996).

[15] Regressions of this type are sometimes known as "Barro regressions," after Robert Barro (1991). Precursors that were less influential, but are similar in spirit, include Sherman Robinson (1971) and Roger Kormendi and Philip Meguire (1985).

Since these regressions usually include the investment ratio and initial income, they can be interpreted in terms of the MRW model. This interpretation is rarely made explicit, and there are few attempts to calculate technology parameters from the coefficient estimates. Yet the informal growth regressions can be seen as a crude extension of MRW, in that they attempt to account simultaneously for input growth and variation in TFP growth. Hence g is seen as a function of observable variables, $g(X)$. In practice the extension is not perfect, since the term $\ln(n + g(X) + \delta)$ is difficult to estimate even when using nonlinear methods.

More importantly, the extension is also imperfect because the additional X variables, like financial depth or inflation, may be correlated with initial efficiency. If countries that are relatively inefficient also tend to have higher inflation rates, inflation will take a negative sign when entered into (2), even if it has no long-run effect on output. It is also possible that the supposedly puzzling continent dummies often used in convergence regressions just proxy for differences in initial efficiency. Hence, although simple aggregate models are always dubious, some important insights are neglected in the absence of a formal theoretical derivation.

Going back to the production function derivation of MRW, another closely related point should be noted. When a variable enters an informal growth regression, it is not clear whether it affects the long-run growth rate, the steady state level of income, or both. Disentangling the two is rarely possible, and the problem is not unique to this method. It is not as devastating as it may first appear. A large effect on the steady state level of income may be as important in practical terms as a growth effect, and of just as much relevance for policy. I will return to this issue in the conclusions.

3.5 *Cross-Country Growth Accounting*

The cross-country growth literature has sometimes been unnecessarily ambitious in trying to infer whether variables act through factor accumulation or TFP growth. Yet this remains a question of great interest to many. In this section, I consider other ways of learning about technology parameters through estimation. The central idea is that we can avoid the omission of a term in initial efficiency by making use of capital stock data. One way of doing this is to use the initial capital-output ratio in place of initial income on the right-hand-side of the regression (2), as suggested by Jess Benhabib and Jordi Gali (1995).

However, if we have data on capital stocks, a more straightforward approach is that adopted by Benhabib and Mark Spiegel (1994) among others. It might be called cross-country growth accounting, or growth accounting with externalities. The method starts from the framework of traditional single country growth accounting, but with two differences. The output elasticities are estimated rather than imposed; and part of the model may be designed to capture the cross-country variation in TFP growth. Unlike the MRW and Benhabib-Gali approaches, no approximations are needed.

Remember that in the conditional convergence regression, initial income is used to control for the transitional dynamics induced by factor accumulation. If, however, we have data on factor inputs, it is not difficult to set up a regression in which the change of output is directly regressed on changes in these inputs. Taking the simplest example, we can take logs of a Cobb-Douglas production function and difference it to obtain:

$$\Delta \log Y = \Delta \log A + \alpha \Delta \log K + \beta \Delta \log H + \theta \Delta \log L + \varepsilon \quad (4)$$

where the notation is the same as before. This provides direct estimates of factor shares, and there is no term in initial efficiency. Instead, in estimating the model we should replace the unobservable $\Delta \log A$ by some function of observables, $g(X)$. Otherwise, the estimates of factor shares will be biased.[16] In any case, it should be clear by now that building a good model for $g(X)$ is often a motivation for this kind of study.

One problem with estimating (4) is well known from the microeconomic literature on production functions. Since the factor inputs are decision variables, agents may respond to shocks by altering inputs of, say, physical capital. In that case, the regressors will be correlated with the error term, and estimates of the parameters will be inconsistent. Various solutions to this are possible, as outlined by Zvi Griliches and Jacques Mairesse (1995). Perhaps the method most likely to find favor is using lagged levels of the factor inputs as instruments, given the likely absence of other suitable variables.

It is worth remembering that these difficulties apply equally to the other possible specifications of growth regressions. The framework advocated here has a key advantage: there is no longer a danger of spurious correlations driven by the omission of initial efficiency. A second advantage is that it can be implemented in a panel without the additional complexities of dynamic panel data models, provided that TFP growth is unrelated to initial income.

As it stands, the method will recover only the average output-capital elasticity across countries, and so TFP growth will be mismeasured for individual countries. One promising elaboration is that of Gary Koop, Jacek Osiewalski, and Mark Steel (1995). They estimate stochastic production frontiers which vary across regional country groups; the method allows them to decompose growth into input changes, efficiency change, and technical progress. One might be sceptical about the details of this decomposition, but this research is innovative in deriving rates of productivity growth on a comparable basis for an unusually wide range of countries.

One remaining disadvantage with all these approaches is that they require information on the initial stocks of physical and human capital. For many developing countries, investment series are not available before the late 1950s, and hence any estimate of a stock for this period is likely to be no more than an educated guess. The implication is that these techniques will become more useful as longer spans of data become available.

4. *Problems*

Whatever empirical framework is adopted, there are usually substantial problems in estimating and interpreting growth regressions. The difficulty of unobserved fixed effects is just one. In this section I consider other frequent concerns, starting with the objection that very different countries are unlikely to be drawn from a common surface, as multiple regression assumes. Section 4.2 discusses outliers, while Section 4.3 examines model uncertainty, drawing on influential work by Levine and David Renelt (1992). Then the problems of measurement error, endogeneity and error correlation are also discussed.[17]

[16] Benhabib and Boyan Jovanovic (1991) and Benhabib and Spiegel (1994) analyse the likely directions of the biases.

[17] Heteroscedasticity I leave aside, partly because it is better understood. Most researchers use

The underlying aim is to ensure that the reader gains an impression of why some results are fragile, and how this fragility might be overcome in future research. Again, I should emphasize that readers anxious for a summary of findings should now skip to Section 6, perhaps returning to this material on a second reading.

4.1 *Parameter Heterogeneity*

I start with a frequent objection to empirical work on economic growth, one which some mistakenly seem to regard as devastating. Countries differing widely in social, political and institutional characteristics are unlikely to fall on a common surface. One well known comment along these lines is that of Arnold 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?

There is a cheap answer to this, and a more thoughtful one. The cheap answer is that any statistical modelling requires assumptions about parameter constancy, and we can see how closely the countries fall on a common surface simply by looking at a regression R^2. This is typically about 0.5 in a cross-section growth regression, suggesting that these regressions do have some explanatory power, despite the heterogeneity difficulty and, one might add, measurement error and omitted variables. Recent regressions, such as those in Sachs and Warner (1997), can explain more than 80 percent of the variation in growth rates.

A more thoughtful answer is to acknowledge that problems of parameter heterogeneity are likely to be more severe in the cross-section growth context than in other areas of economics (Levine and Renelt 1991). It is easy to see, for instance, that the coefficient on the investment ratio is likely to be lower in war-torn and unstable countries than in peaceful ones. It is possible that parameters vary across countries in ways which render conventional estimates inconsistent.

Several papers have provided strong evidence for widespread heterogeneity. Particularly important in this respect is the work of Durlauf and Paul A. Johnson (1995), who use regression trees to identify multiple regimes, those country groupings across which parameters differ widely. Together with Durlauf and Quah (1998), they have emphasized that linear regressions assuming common parameters are an inappropriate way to investigate new growth theories. Such regressions will often fail to distinguish recent models with multiple equilibria from the traditional Solow-Swan framework.

One approach that has been advocated instead is to set up a transition matrix, showing the probabilities of moving between particular income ranges in a given time period. Quah (1993, 1997) argues that this kind of study is preferable to growth regressions in describing the evolution of the world income distribution. The transition matrices form a very useful summary of income dynamics, but are less effective at revealing causal mechanisms, at least as yet.

Given the existence of heterogeneity, a key question for traditional methods is how much they tell us about parameter *averages*. Given that the purpose of cross-country empirical work is often to arrive at generalisations about growth, the averages are important. It turns out

heteroscedastic-consistent standard errors. Perhaps, though, alternative methods like weighted least squares might be preferable. It should also be remembered that where heteroscedasticity is observed, it may reflect neglected parameter heterogeneity.

that when a dynamic panel data model is used, as is frequent in the growth literature, the estimate of the average effect will be inconsistent, even when the length of the time series tends to infinity. The reasoning is that if the regressors are serially correlated and one neglects to model parameter heterogeneity, this will induce serial correlation in the disturbance. In a model with a lagged dependent variable the outcome will be inconsistent estimates (Pesaran and Ron Smith 1995).

In contrast, cross-section regressions can produce consistent estimates of the average long-run coefficients even in a dynamic model, provided that the varying coefficients differ randomly and are distributed independently of the regressors. This does not mean that cross-section regressions are necessarily the best technique: we have already noted the omitted variable bias inherent in conditional convergence regressions. It does, however, reinforce the case for using static panel data models of the kind proposed in Section 3.5.

Unfortunately, it will be difficult to make much progress on questions of parameter heterogeneity until we have data over a longer time period. When more data becomes available, it may be possible to use panel data models with stochastic parameters. At present, more imaginative use of methods for detecting heterogeneity, such as robust estimation, regression trees, sample splits, interaction terms, and dummy variables, would be welcome, especially when the use of panel data allows more degrees of freedom. Recent research has been far more careful about this. Harberger had a point.

4.2 *Outliers*

Once we acknowledge the existence of parameter heterogeneity, we must also acknowledge that our regression model is only a crude approximation to reality. This means that it is likely to fit some observations particularly badly. A problem immediately arises here: observations that are "unrepresentative," due to parameter heterogeneity, omitted variables or measurement error, can act as influential outliers or leverage points.[18]

Unfortunately, attempts to address this difficulty are often rather token. Sometimes observations are dropped one at a time, or single-case diagnostics like Cook's distance are used; but these methods are well known, at least to statisticians, to be inadequate. Single-case diagnostics are likely to miss groups of outliers (the "masking" effect) and may even be misled into identifying representative observations as outlying (the "swamping" effect).

A better alternative is provided by robust regression. As discussed by Peter Rousseeuw and Annick Leroy (1987), robust estimation procedures allow researchers to characterise the most coherent part of the data set, and use the resulting parameter estimates to identify possible outliers. These estimators have rarely been applied in the growth literature, even though they have considerable advantages over least squares, at least at the exploratory stage. The identification of possible outliers will not only help render generalizations more robust, but will also highlight countries with atypical growth experiences, ones that are particularly likely to reward further study.

4.3 *Model Uncertainty*

A central difficulty with empirical growth studies has been widely

[18] As an illustration, in the context of growth regressions analysing equipment investment, Brad De Long and Lawrence Summers (1991) discuss the large effect that Botswana and Zambia have on the coefficient estimates and their precision.

acknowledged by practitioners since the work of Levine and Renelt (1992). Many variables have been found to be significant in growth regressions; nearly as many have been found to be "fragile," in the sense that their statistical significance disappears when a different group of right-hand-side variables is selected.

Problems of this kind have been given some discussion in the statistics literature. The difficulty is that several models may all seem reasonable, but lead to different conclusions about the parameters of interest. In these circumstances, presenting the results of a single model is misleading. It ignores the researcher's inevitable uncertainty about the form of the model, and so leads the reader to underestimate the uncertainty actually present. Edward Leamer (1983, 1985) was one of the first to emphasise the difficulty, and it is a variant of his proposed solution, "extreme bounds analysis," which underlies the work of Levine and Renelt.

Since their paper, several writers have used the same framework to demonstrate that a correlation is robust to changes in specification. This is an advance, but not without problems of its own. Several of the regressors used by Levine and Renelt are almost certainly endogenous. More fundamentally, the robustness study which simultaneously addresses fixed effects, measurement error, endogeneity, outliers and model uncertainty, is yet to be written, and one suspects that this will remain the case for some time. Hence the finding that a variable is robust to changes in the model specification is not sufficient for valid inference.

It is also worth pointing out that robustness is not a necessary condition for useful information. One danger with the Levine and Renelt approach is that too much attention is paid to the statistical significance of variables. A worryingly common mistake is to conclude that "there is no effect of variable X" simply because its effect is imprecisely measured; one suspects that Type II errors are frequent.[19] This is not a mistake of the original paper, which makes clear that a lack of robustness should often spur further investigation into causality and inter-relationships. Finding that a result is fragile to a particular conditioning variable could in itself be valuable information.

Another illustration of the dangers of taking the Levine-Renelt analysis at face value is useful. It is plausible that many of the variables in growth regressions are symptoms of deeper problems or "syndromes": high inflation reflects bad macroeconomic policy, while a high black market premium may reflect a mixture of bad short-run policy and inward orientation. There are many possible indicators for each syndrome, which will tend to render one another insignificant when included at the same time. If one treats individual variables as symptoms of underlying problems, finding that some are "fragile" in the Levine-Renelt sense does not tell us much about their potential usefulness (Pritchett 1998). One promising way forward is to combine indicators using latent variable methods, and then examine the robustness of these overall measures.

4.4 *Endogeneity*

One of the most frequently expressed concerns about work in the growth

[19] Having said that, the issues in statistical inference raised by growth regressions are interesting, since we are often looking at the population rather than a sample from it. Even when the regressors are considered to be stochastic, arguably inference should proceed conditional on their realized values.

literature is the probable endogeneity of some regressors. To avoid simultaneity concerns, researchers often make use of initial values, for instance regressing growth over 1960–85 on the 1960 secondary school enrollment rate. This is not quite as watertight as researchers seem to think: even if the endogeneity problem is solved, perhaps some omitted variable, like the political regime, affects both growth and the initial level of variables like schooling.

Early papers often used the average investment ratio as a regressor. It is easy to construct arguments that causality could go from growth to investment, as well as vice versa.[20] Recent work uses instrumental variables to avoid these problems (Barro and Sala-i-Martin 1995; Barro 1997; and Francesco Caselli, Gerald Esquivel, and Fernando Lefort 1996). There is a tendency for the coefficient on the investment ratio to fall, as one might expect. Incidentally, this suggests that some of the earlier results in the literature may understate the effect of variables like human capital that are positively correlated with investment and growth.

In general there is a shortage of good instruments. So many variables could be used to explain growth that it is difficult to find variables that are not only highly correlated with the endogenous variables but can also plausibly be excluded from the regression. One solution is to use a panel, and employ lags of the endogenous variables as instruments. Unfortunately their exogeneity is not always clear. There may be long delays in the effects of, say, human capital accumulation. Although serial correlation tests can be used to examine the identifying assumptions, it should be remembered that these are based on the estimated residuals, not the true disturbances.

This will be a particular difficulty when researchers estimate conditional convergence regressions but do not take steps to eliminate the effect of the initial level of technology. Since this level is persistent but omitted, there is likely to be serial correlation in the error term for this reason alone. Serial correlation tests may not detect this. The variables in the growth equation and the lagged values used as instruments will both be correlated with the level of technology, and so little of this level will remain in the estimated residuals. Hence the instruments will appear to be valid, when a serial correlation test on the true residuals would show them not to be.[21]

To avoid endogeneity problems, the best solution may be to use panel data to estimate the cross-country growth accounting framework proposed above, still using lags as instruments. Admittedly the use of capital stock data will introduce measurement error. It should, though, be easier to find instruments uncorrelated with this error than it is to find instruments uncorrelated with the unobservable level of technology.

Another fruitful line of research might be to adapt empirical growth models to allow explicitly for the possibility of regressors that are endogenous to the growth rate or the level of income. In an innovative paper De Long (1997) extends a Solow model by adding equations relating population growth and the investment share to the level of per capita income, and he shows how this affects the interpretation of

[20] For instance, it seems to be a fairly robust stylized fact that growth raises saving (Christopher Carroll and Weil 1994). With imperfect capital mobility, this will induce a correlation between investment and growth even in the absence of any independent causal role for investment.

[21] The arguments in this and the next paragraph are due to Christopher Sims (1996).

parameter estimates. A useful insight is that increases in total factor productivity, in raising income, can also reduce population growth and the relative price of capital goods, which then feed through to further increases in the capital-labour ratio and per capita income.

4.5 *Measurement Error*

As discussed earlier, the Penn World Table used in cross-section growth regressions is inevitably flawed in some important respects. Researchers also frequently make use of additional data, such as that on income inequality, for which problems of mismeasurement are even more severe. Yet few studies attempt to assess sensitivity to measurement error, let alone correct for it. In part, this may be due to the common misperception that measurement errors simply bias coefficients towards zero. When there are several badly measured variables, or the errors depart from classical assumptions, then biases may go in either direction.

That said, some regard the presence of measurement error as pretty much fatal for the whole exercise, and to hold this view is surely to go too far. Those instinctively suspicious of the literature sometimes fall back on measurement error as an excuse to avoid engaging with it. Yet some results hold using different measures, samples and specifications. A more useful reaction to the problem of measurement error is to try and identify which variables are particularly badly measured, and hence isolate the findings that are least reliable. This would also help researchers direct their efforts towards the variables that are relatively well measured.

Occasionally sensitivity to measurement error is explicitly assessed, as in Torsten Persson and Guido Tabellini (1994) and Temple (1998a). Multivariate reverse regression and classical method-of-moments estimators are two particularly useful techniques. Another possibility is the use of instrumental variables, although I have noted above the shortage of plausible candidates. When a panel is used and the measurement error is not thought to be persistent, one can use lags of the variables as instruments in GMM estimation, as Caselli, Esquivel, and Lefort (1996) discuss.

4.6 *Error Correlation and Regional Spillovers*

Several researchers have suggested that the disturbances in cross-section growth regressions may not be independently distributed. Although this can often be seen as an omitted variable problem, there are common shocks like the climate which will always be difficult to model, and could lead to correlation in the disturbances. Although De Long and Summers (1991) found little evidence for spatial correlation, it is often the case that regional dummies add substantially to a growth regression's explanatory power.

These dummies partly reflect the fact that failure has been concentrated in sub-Saharan Africa, and success in East Asia. It is also interesting that large growth fluctuations seem to have been sychronized across many of the major Latin American countries over the last three decades. Observations like these have led several researchers to think about spillovers between geographical neighbors (Alberto Ades and Hak B. Chua 1997; Easterly and Levine 1998; Ramon Moreno and Bharat Trehan 1997).

Adjustment for spatial correlation raises formidable statistical problems, and the identification of genuine spillovers is problematic (Charles Manski 1993). The spillover effects that have been found may simply reflect omitted

variables. As it stands, these questions probably need to be investigated further. Without more evidence that the disturbances are independent, the standard errors in most growth regressions should be treated with a certain degree of mistrust.

5. *Panel Data and Time Series Methods*

Having absorbed the potential difficulties of cross-country studies, some readers will be tempted to dismiss the usefulness of this style of research altogether. I think that would be a mistake, and I now complete the discussion of methods by explaining why. Techniques that make more use of time series variation in the data might yet overcome many of the objections often raised to cross-country research. My argument will be that although the various approaches are complementary, panel data studies will increasingly offer the best way forward for many questions of interest, especially as longer spans of data become available.

This conclusion may be controversial, especially since it is becoming fashionable to argue that calibration techniques can overcome some of the difficulties of the cross-country growth literature.[22] One argument along these lines, already encountered, is that we should select the parameters of the aggregate production function based on micro evidence.

As pointed out earlier, this kind of approach ignores externalities. If we are willing to do that, we do not need to assume a common aggregate production function at all, but could probably derive comparable TFP growth rates across countries using data on factor shares or micro evidence. However, theory gives little guidance on a theoretical specification for TFP growth, or the micro parameters that could be used to calibrate it. Hence the econometric issues raised here will need to be addressed whether or not calibration is used.

5.1 *Panel Data Methods*

The advantages of using panel data techniques to study growth are several. First, and most fundamentally, they allow one to control for omitted variables that are persistent over time. As we have seen, conditional convergence regressions rarely give unbiased estimates, because variations in technical efficiency across countries are likely to be correlated with the regressors. By moving to a panel data framework, we can at least control for unobserved heterogeneity in the initial level of efficiency.

Another advantage is that several lags of the regressors can be used as instruments where required, thus alleviating measurement error and endogeneity biases. A highly promising approach is that of Caselli, Esquivel, and Lefort (1996), who use GMM to estimate a dynamic panel data model. The growth model is first differenced to eliminate the effect of initial efficiency, and then lags of the variables used as instruments. Their work suggests that existing cross-section estimates of convergence rates are severely biased, and so demonstrates the potential usefulness of this technique.

However, this approach is still in its infancy and there remain reasons to be sceptical about even the latest results. The particular GMM estimator used by Caselli, Esquivel, and Lefort is likely to perform badly when the variables are highly persistent, because then lagged levels are weak instruments for first differences. More generally, the finite

[22] Unfortunately, shortage of space precludes a full discussion. David Romer (1996) provides an interesting critique of calibrating short-run macro models, much of which applies in this context as well.

sample properties of most dynamic panel data estimators are not yet well understood. Different estimators do well in different circumstances, so as yet it is difficult to advise on the best course for growth researchers.

There are some remaining worries about the use of fixed effects specifications. Too often researchers use fixed effects approaches to analyse the effects of variables that are fairly constant over time, or that will affect growth only with a long lag. Standard transformations like first differences or "within groups" are likely to exacerbate the problem of measurement errors, at least if these errors are not persistent. They typically lead to a large fall in precision, since in effect the between-country variation is thrown away. One might even prefer to ignore fixed effects when using something like a mean-square error criterion.

In other areas of economics, the results from fixed effects estimation are often found to be disappointing. Griliches and Mairesse (1995) argue that it may be better to proxy for omitted variables, since this leaves more identifying variance in the independent variables, and is informative in itself. One simple approach is to use carefully specified regional dummies. The justification for this is that much of the variation in efficiency levels occurs between rather than within continents, an argument suggested by the estimates in Koop et al. (1995). Along similar lines, one could use panel data to estimate growth regressions separately for different regional groupings.

Given the presence of cyclical effects, researchers face some difficult choices in selecting the time intervals over which to study growth. Whether one is best using annual data, or five- or ten-year averages to avoid business cycle effects, is a question that remains largely unsettled. If annual data is used, one must take great care in modelling the short-run dynamics, especially as some heterogeneity is inevitable. Most researchers have opted for five or ten year averages, but the latter in particular mean that one is left with little time series variation. It is clear that the prospects for rigorous growth research will greatly improve as longer spans of data become available. In the meantime, we need to investigate how well short time averages eliminate business cycle effects.

Despite the problems, the use of panel data methods, together with tests for parameter heterogeneity, can probably succesfully address many of the objections raised to cross-country empirical work on growth. These papers certainly present a harder target for the sceptics. At the very least, standard cross-section approaches should sometimes be supplemented with an assessment of sensitivity to a fixed effects specification.

However, my own view on the use of *dynamic* panel data methods is that they may introduce an unnecessary degree of sophistication. In conventional specifications, they are needed because of the presence of transitional dynamics in output. It may be much simpler to adopt cross-country growth accounting, either by estimating equations like (4) above or by using other data to impose the technology parameters. Panel data methods can then be used to study the determinants of TFP growth, whilst avoiding the extra complexities of dynamic models.

5.2 *Time Series Methods*

Some econometricians, primarily those from a time series background, argue that we should go further than using panels. They argue that standard cross-section methods throw away useful information, while panel data

methods make unjustifiable assumptions about parameter homogeneity. Instead, we should estimate parameters for countries individually, using separate time series regressions for each country.

One immediate problem is the quality of the data for developing countries. Many important variables, such as population growth and school enrollment, are often interpolated from just three or four census years. A second difficulty is the limited timespan of the available data; even when observations are available for each of thirty years, it is difficult to discern the long-run effect of variables like inflation. To prevent short-run business cycle effects driving apparent long-run correlations, long lags of the independent variables need to be included, and one starts to run into a degrees-of-freedom problem if the cross-section variation is ignored.

As it stands, growth researchers have only rarely confined themselves to the information in the time series properties of the data. Among the exceptions are work by Jones (1995a) and several studies of inflation and growth, all of which will be discussed later on. My own feeling is that time series econometricians are usually mistaken in believing that there is much to be gained by moving to the use of annual data, even in panels. Inevitably the short-run variation in growth rates is dominated by business cycle effects, not by changes in fundamental long-run growth prospects. Only long time averages of growth rates, compared in the cross-section or using a panel, can al- low us to address the determinants of long-run growth with any degree of confidence.

6. *Convergence and Technology: What Have We Learned?*

The preceding sections argue that important improvements can be made to the cross-country literature by more careful attention to econometric issues. However, surveys that conclude that all the valuable work lies ahead are never particularly useful, and are likely to gloss over some useful contributions. With that in mind, the second half of the survey reviews what we have learned so far.

In reviewing what we know, I start in this section with issues at the heart of the literature, those relating to convergence and technology. Section 7 will cover the contribution to per capita growth of the various inputs (physical and human capital, and research). Then Section 8 will broaden the survey to include wider influences, such as inequality, infrastructure, and macroeconomic management.

In the convergence literature, some confusion has been caused by the tendency to see various studies as being opposing answers to the same question. Instead, the different studies are best seen as complementary approaches to the six questions outlined in the introduction. Section 6.1 will discuss measurement of the rate at which countries converge to their respective steady states, and its implications. Sections 6.2 and 6.3 focus on the role of technology. Section 6.4 briefly considers the forces at work in convergence, paying especial attention to trade.

6.1 *Measuring the Rate of Convergence*

Researchers have been keenly interested in the proportionate rate at which regions and countries close the gap between their current positions and their respective steady states. It is not uncommon to read strong claims that this rate is remarkably stable, at around 2 percent a year (Barro and Sala-i-Martin 1992; Mankiw 1995; Sala-i-Martin 1996). The significance of this lies in its implication that returns to physical and

human capital diminish, but only very slowly.

This kind of story about convergence and the extent of diminishing returns, though popular, is not entirely reliable. The 2 percent estimates are taken from rather crude cross-section regressions, and there is no shortage of accompanying econometric problems. Fixed effects are either ignored or dismissed a priori. Sensitivity to measurement error is rarely explicitly discussed or assessed. The possibility of heterogeneity biases or outliers is not considered. More sophisticated panel data and time series studies suggest that these flaws are crucial to obtaining the 2 percent result, and recent estimates of the rate of convergence have varied between zero and 30 percent a year.[23]

The consensus now emerging is one of uncertainty. It is just not very easy to disentangle the convergence rate from other aspects of growth. Arguably this should not be surprising. Conditional convergence implies mean reversion, and so there is a close link between investigating convergence and testing for unit roots. We know from the persistence literature just how hard it is to arrive at definitive conclusions in this field.

Progress is likely to require more careful thought about fixed effects, parameter heterogeneity and the adequacy of unit root tests in short panels. The attendant danger is that researchers will see technical sophistication as an end in itself, and lose sight of the reasons for interest in conditional convergence. At present, too little thought is given to the value added in studying convergence rates. There are other, perhaps easier, ways of learning about the extent of diminishing returns.

[23] See Caselli, Esquival, and Lefort (1996), Paul Evans (1997), Nazrul Islam (1995), Kevin Lee, Pesaran, and Smith (1996) and Temple (1998a).

6.2 *Technology and International Differences in Growth Rates*

The debate about convergence rates is linked to another ongoing dispute, the importance of differences in technology across countries. At one extreme is the work of MRW, who claim that income and growth differences can be almost completely explained using a model in which technology is a public good, freely available to individuals in all countries. Differences in incomes originate largely in differences in steady state levels of human and physical capital.

Near the other extreme are most development economists, economic historians, and theorists, who see "idea gaps" as central to the problems of developing countries. Capital accumulation is often seen as a subsidiary part of the central task, which is adopting best-practice technology from abroad. Researchers in this tradition emphasize that many important ideas are protected or secret, and others can only be acquired by experience. An emphasis on the need to transfer ideas and expertise points to factors one instinctively thinks of as important, such as national culture and institutions.

Why, then, has the "public good" view of technology held such a sway over the recent empirical literature? One argument in its favour is that, at least ultimately, countries should have access to any technical knowledge available elsewhere. Hence, in the long run, the rate of technical progress will be the same worldwide, even if levels of productivity differ because of local factors like climate.

Unfortunately empirical researchers often jump from this view to a much more controversial position, without much explanation. In some hypothetical long-run equilibrium technical progress

will be the same across countries. This does not mean, however, that one is justified in assuming constant rates of TFP growth in any sample we observe. The key difficulty should be obvious. If developing countries often lag behind others in technical efficiency, there is no reason to expect their efficiency to grow at the same rate. Furthermore, we measure not technical progress but TFP growth, and this may be affected by such things as instability and war.

Yet the early assumption of similar technologies and TFP growth rates appeared quite successful. MRW draw attention to one particularly dramatic finding. They claim that around 80 percent of the international variation in per capita incomes can be explained using just three variables: population growth, and investment rates for physical and human capital. The corollary is that differences in technical efficiency can have only a small role to play in explaining cross-country income variation.

However, their approach makes two controversial assumptions, that investment rates are exogenous to the level of income and uncorrelated with efficiency. More recently, Peter J. Klenow and Andrés Rodríguez-Clare (1997) have drawn attention to another underlying problem with the MRW finding. The human capital variable that MRW use only captures variation in secondary schooling. Since it ignores primary schooling, it tends to exaggerate the variation in human capital across countries. When Klenow and Rodríguez-Clare correct for this, they find that MRW's model only explains around half the variation in incomes, leaving a central role for technology differences.

Consider some other difficulties with the MRW approach.[24] If one follows them in assuming that TFP has grown at a rate of 2 percent a year in all countries since 1960, what does one make of the many countries which have grown at less than 2 percent a year over this period? Perhaps these countries have received a large negative shock to technology. Otherwise, the logical implication of the model is that these countries are converging to their steady states *from above*. Any country which has grown at less than 2 percent a year must have exceeded its steady state capital stock in the early 1960s, and have been running it down subsequently. Perhaps this is right, but it seems unlikely.

Other problems abound. Economic miracles, such as Japan's post-war growth, are hard to explain by capital accumulation alone. In most frameworks, such a process would have to be accompanied by a steeply falling real interest rate, something that has not been observed. To paraphrase the words of Klenow and Rodríguez-Clare, the neoclassical revival in growth economics went too far. Understanding the reasons for differences in technical efficiency and TFP growth is essential to the empirical growth project.

6.3 *Technical Progress in the Long Run*

Although it is probably safe to dismiss the idea that TFP growth rates are presently the same across countries, there are some interesting issues surrounding rates of technical progress in the advanced industrial countries, connected to endogenous growth models. Theorists working in this field argue that the long run growth rate depends on the allocation of resources, and

[24] Papers relevant to this discussion include Andrew Bernard and Jones (1996b), Dongchul Cho and Stephen Graham (1996), Robert King and Sergio Rebelo (1993), Stephen Parente and Edward Prescott (1994), and Jonathan Eaton and Samuel Kortum (1995). For a broader treatment of technology and growth, see Jan Fagerberg (1994).

particularly the amount of resources devoted to research. At least in the early models, one implication is that even industrial countries will grow at different rates for long periods of time.

One useful piece of evidence here is the path of the cross-country variance of incomes. Evans (1996) starts from the observation that, if countries differ in their trend growth rates, the logarithms of their per capita incomes should wander away from each other over time. More precisely, it can be shown that the cross-country variance of log incomes will be integrated of order one around an upward quadratic trend. Evans presents evidence for thirteen industrial countries that the variance has not trended upward and that, if anything, the reverse has occurred.

This clearly corresponds to the earlier finding of convergence among developed countries. It tends to suggest that at least the advanced industrial countries are growing at the same rate over the long run. Evans writes that "either endogenous growth models are fundamentally flawed, or else the effects they predict must be relatively unimportant for the countries considered here."

Since his paper, some work has suggested that rates of technical progress may vary even across industrial countries.[25] It is also worth pointing out that, even if countries do grow at the same rate in the long run, new growth theories may still provide the explanation of the world growth rate. For empirical applications of endogenous growth theory, one is likely to need a model in which domestic research is accompanied by knowledge spillovers from other countries, as provided by Philippe Aghion and Peter Howitt (1998). There is certainly evidence for such spillovers. David Coe and Elhanan Helpman (1995) find large effects of foreign R&D on domestic total factor productivity, while Eaton and Kortum (1994) calibrate a model of international technology diffusion and find that, even for the USA, around half its productivity growth depends on foreign technology improvements. This suggests that the evidence for a common long-run growth rate can be seen as consistent with endogenous growth.

[25] See in particular Lee, Pesaran, and Smith (1997).

6.4 *What Drives Convergence?*

If we agree that there is more to income variation than differences in investment rates, it is interesting to ask why convergence might be taking place. If we allow for technology differences, the negative coefficient on initial income requires careful interpretation. It may indicate not just that countries are converging to their steady states through capital accumulation (as MRW proposed) but also that technology transfer is taking place.

In this case, convergence studies will be an especially poor way of learning about the extent of diminishing returns, and there is no longer a theoretical reason to expect the relation between growth and initial income to be linear. Several writers, most recently Easterly and Levine (1997), have modelled growth as a quadratic function of initial income. A common finding is that middle income countries have grown relatively fastest once controlling for policy differences.

Some studies have looked directly at convergence in TFP, particularly for the OECD. Given links through trade, one might expect to find TFP convergence between sectors in different countries, perhaps especially in manufacturing. In practice, there is only weak evidence that efficiency in OECD

manufacturing sectors converged in the 1970's, and some evidence of divergence in the 1980's. This suggests that the aggregate TFP convergence is driven by the performance of services.[26]

The link between trade and convergence is of great interest, partly because trade theory provides other mechanisms, like factor price equalisation, through which countries might converge. There is some suggestive evidence of decreasing income dispersion among countries linked by trade, which can be partly attributed to the closing of technology gaps.

Finally, it should be noted that trade links between countries may undermine many of the traditional beliefs about conditional convergence. Jaume Ventura (1997) points out that there is no necessary link between diminishing returns and conditional convergence in models of trading economies. If a weak form of factor price equalisation holds, the law of diminishing returns applies only to the world's stock of capital, and investment will be equally productive in each country. Hence one would not necessarily expect to find any link between growth and initial income, at least when controlling for investment. Thus Ventura's work reinforces the idea that the usual conditional convergence result may partly be driven by technology transfer.

7. *The Proximate Sources of Growth*

Overall, my reading of the literature is that growth differentials are partly explained by varying degrees of technology transfer, as well as variation in input growth. It is a useful first step, but a small one. It invites many questions, not least concerning the relative importance of particular inputs, and the role in growth of wider influences, including some that have long fascinated academics and policy-makers, like free trade and democracy.

In this section I will assess the role of various inputs, or what might be called the 'proximate' sources of growth: investments in physical and human capital, and in research and development. The next section will turn to the wider influences.

7.1 *Investment in Physical Capital*

As most undergraduates know, the Solow-Swan model taught us that long-run growth is likely to be independent of the investment rate, because the returns to accumulable factors are probably sufficiently diminishing for this to be true. The same students were often taught that no correlation between investment rates and growth is observable in the data. We now know that was wrong. There is a robust correlation between investment rates and growth, which sometimes survives the use of instrumental variables. A key point is that over a finite time period, this correlation is to be expected even in the Solow-Swan model, and its strength can be used to estimate the share of capital in the production function.

However, some argue that if the endogeneity of investment could be correctly dealt with, the effect of investment would be close to zero in developing countries, because much of it takes place in heavily distorted environments. As discussed in Section 4.4, there are several reasons for believing that endogeneity is important to the investment and growth results, and also several problems in overcoming its presence.

One interesting reason for believing

[26] See the papers by Bernard and Jones (1996a) and Steve Dowrick and Duc-Tho Nguyen (1989) for more on TFP convergence. On trade, see Dan Ben-David (1993) and Ben-David and A. K. M. Atiqur Rahman (1996).

that endogeneity is important is the literature's general ranking of estimated returns to different types of investment. For instance, the returns to foreign direct investment are often found to be extremely high, but it also seems likely that here the endogeneity problem is greatest: foreign direct investment will be pulled into countries already doing well, or expected to do well in the future. The usual ranking of returns to different types of investment tends to suggest that endogeneity is driving the estimates.[27]

The strongest result in the investment-growth literature is that the returns to physical capital are almost certainly diminishing, in agreement with the Solow-Swan growth model and most theoretical work since. This is the finding of both convergence regressions and cross-country growth accounting (Benhabib and Spiegel 1994, King and Levine 1994).

The finding that returns are diminishing does not rule out important externalities to investment. The work of De Long and Summers (1991) has drawn attention to the potential importance of investment in equipment, not without controversy along the way. A consensus has emerged that the correlation between equipment investment and growth appears to be weak in the OECD, but the debate over this has distracted attention from its role in developing countries, for which the findings of De Long and Summers appear to be rather robust.[28] It is difficult to explain away their results by simply referring to influential outliers or measurement error concerns, although endogeneity almost certainly accounts for some portion of the very high estimated returns.

[27] The arguments in this paragraph are due to Lant Pritchett.

[28] For more on equipment investment, see Temple (1998b) and the references there.

Given the apparent robustness of the results, attention should perhaps shift to their interpretation. Perhaps the returns to equipment investment are overestimated in the presence of industrialization. Another argument is that equipment investment is an important part of technology transfer, and the returns to it then appear to be high because other costs of technology transfer (technical adaptation, training, reorganization) are omitted. Both these points may have something to them, and there remains considerable scope for further work. It is not as easy as it once looked to dismiss the findings of De Long and Summers.

7.2 *Human Capital*

Harold Alderman et al. (1996) point out that developing country governments spend over $100 billion a year on education, health, and other human capital investments. It is clearly important to understand how these investments contribute to growth. There is a long tradition of microeconomic studies of the returns to schooling, but given recent emphasis on externalities to education, macroeconomic studies also have a role to play.

The MRW paper seemed to provide a broad-brush answer to questions about the role of human capital in growth. Using the proportion of the adult population enrolled in secondary school as a proxy for human capital investment, they argued that aggregate technology is simply described by a Cobb-Douglas production function taking human capital as one of its inputs (the augmented Solow model). An estimated exponent of 1/3 on human capital is, they claimed, consistent with both the estimated rate of convergence and sketchy evidence on the importance of human capital to the United States.

Papers since then, however, have

questioned whether the links between human capital and growth are quite so simple. We have already seen that the MRW analysis perhaps overstates the variation in output per head that can be explained by variation in human capital. Another problem to emerge is that changes in human capital appear to explain little of the variation in changes in output, casting doubt on the augmented Solow formulation.[29] This macroeconomic evidence conflicts with the finding of the micro literature that schooling has a significant return in terms of higher wages. The failure to discern this effect at the macro level is worrying.

The literature uses somewhat dubious proxies for aggregate human capital. The focus is almost exclusively on schooling rather than training. This is mainly due to data limitations, but less excusable is the attention frequently paid to school enrollment rates. These rates were initially regarded as one of the more robust and satisfactory variables in the growth literature, but it is worth remembering that they are likely to be positively correlated with initial efficiency, so the results could often be spurious. Certainly it has been much harder to find an effect of human capital in panel data studies, although it is also true that too few researchers think carefully about the specification. Rather optimistically, they tend to expect a change in school enrollments to raise growth almost instantly.

Equally importantly, there are some conceptual difficulties with the use of school enrollment data. Only rarely do these rates correspond well to the human capital variables highlighted in theoretical models. In many empirical growth papers it is not clear whether school enrollment rates are intended to represent a flow of investment in human capital, or its stock. In practice these rates may be a poor proxy for either, and given that data on average years of schooling in the population or labour force is now available, the continuing use of school enrollment figures has little to recommend it.[30]

An alternative approach is to construct indices of human capital based on the returns to schooling and experience indicated by micro studies. The basic idea is to combine schooling coefficients from Mincerian wage regressions with data on years of schooling, and then examine the contribution made to cross-country income differences by variation in this human capital measure. This approach is useful, but it is necessarily silent on the extent of human capital externalities, and tends to assume that returns to schooling are not driven by signalling effects.

As this discussion indicates, there is much work still to be done on the role of human capital. One problem is that the aggregate data seems to be too crude to ask some of the most interesting questions: for instance, what matters most, breadth of access or school quality? Other questions surround the role of health status. Variables like life expectancy are often used in growth regressions, but their role is never justified by a well-articulated theory. One interesting possibility is that good health status raises the return to education and training.

Finally, there is widespread agreement that human capital accumulation is not a sufficient condition for growth.

[29] Two important papers here are those by Benhabib and Spiegel (1994) and Pritchett (1996b).

[30] Those not convinced should consult Norman Gemmell (1996) and Pritchett (1996b). Eric Hanushek and Dongwook Kim (1995) consider another alternative to average years of schooling, an index of cognitive skills based on international test scores. Its determinants are investigated by Jong-Wha Lee and Barro (1997).

Most development economists are familiar with examples of countries that have expanded education only to see high subsequent unemployment rates and falling returns; Korea before 1960 might be one example. In other developing countries, wage employment is sparse and the highly educated have often found work only in the public sector, perhaps undermining the overall impact on growth. Hence a key challenge is to elicit the conditions in which expanding education is most beneficial.

7.3 *Research and Development*

Of the "proximate" sources of growth, the last to be considered is research and development. Since the modelling of industrial research lies at the heart of some of the most promising new growth theories, it is worth trying to clarify empirically the role of R&D, and I will take particular care here to relate such findings to recent growth models. I will follow convention and common sense in assuming that research is of greatest importance for advanced industrial economies.

As in the study of human capital, this is a field in which there is already a wealth of microeconomic evidence. Early studies found private rates of return to R&D as high as 30–50 percent for the USA in the 1950's and 1960's. There are some well-known problems in measuring the contribution of research to productivity growth. Overall, though, studies have concurred in finding high private returns. Estimates of R&D spillovers are necessarily flawed and subject to a variety of reservations, but the impression remains that social returns are even higher.[31]

Since empirical studies ignore such externalities as the "business stealing" effect, it might be thought that they typically overstate the social returns to research. This intuitive reasoning may be incorrect. It is possible that distortions like creative destruction and monopoly power affect the share of R&D in output, but do not affect the relationship between the social rate of return and the R&D share. Hence regressions based on the R&D share may recover the social rate of return reasonably accurately, and perhaps even understate it (Jones and John Williams 1997).

Although the microeconomic evidence suggests that theorists are right to emphasise R&D, some researchers have criticised the research-driven growth models on empirical grounds. Central here is the work of Jones (1995a,b). He points out that growth rates in OECD countries since World War II have not shown any persistent upwards trend in spite of policy changes: trade liberalisation, an increase in average years of schooling, increases in investment, and a substantial increase in R&D efforts.

This apparent constancy of long-run growth, in the face of structural changes, obviously calls into question the validity of research-driven growth models. Jones argues that such models are driven by a knife-edge assumption in the way that the benefits of research are related to the existing stock of knowledge. Although a more general relationship can explain the evidence, so can other modifications, not all of them ruling out research-driven models.

There is a well-known stylized fact that the productivity of research, at least as measured by realised patents, shows a long decline from the mid-1950's until the mid-1980's. This may indicate some systematic relationship between the knowledge stock and research benefits, or it may just represent an exogenous fall in research

[31] Two particularly useful references on research and productivity growth are Griliches (1979, 1992).

productivity. Certainly there is little direct evidence of diminishing returns to research. In general, one might expect the progress of knowledge to be uneven, depending as it does on the emergence of new "fishing grounds."

Since some of the underlying concepts (knowledge, ideas) are so hard to pin down, the validity of research-driven growth will be a difficult debate to resolve.[32] Taking sides is not as important as it might first seem. The model Jones proposes allows research to have large level effects, and this is sufficient for an interest in the policy effects of research incentives. It is always worth remembering that "the long run growth rate" is a theoretical abstraction, never observable in practice. Debates about whether or not policy affects it will distract us from questions that are ultimately of more practical importance.

8. *Wider Influences on Growth*

So far, we have considered three proximate sources of growth: physical and human capital accumulation, and spending on research and development. In this section, I widen the survey, to assess the contribution of macroeconomic evidence to our understanding of wider influences on growth. Sometimes these will work through the proximate sources, explaining the cross-country variation in input changes. They may also have a direct effect through total factor productivity, for instance by affecting technology transfer.

In examining these effects, I tend to focus on developing countries. Since institutions and the growth process in developing countries may be very different to those in countries already near the technological frontier, one should often be careful about extrapolating findings from the developing countries to the more developed, and vice versa.

The influences considered here include population growth, the financial sector, the macroeconomic environment, government spending, income distribution, and political and social arrangements. Most of the findings discussed in this section are from cross-country empirical work. Throughout, I concentrate on findings which seem to have survived the attentions of a variety of researchers, using different data sets and specifications, and where at least some thought has been given to problems such as endogeneity, outliers and model uncertainty.

Since I concentrate on synthesizing previous work, there are some important influences on growth that are not given the coverage they are perhaps due. Chief among these are microeconomic structure and the fine detail of institutions. Although the new growth evidence has covered trade policy and political institutions, there are other aspects which are harder to quantify. Few of the variables considered here would offer much insight into the experience of China or the former command economies, for example. It may well be that these issues are better addressed by case studies, or by statistical work within countries at the firm or sectoral level.

8.1 *Population Growth*

One area in which aggregate studies may make an especially useful contribution is the relation between population growth and economic performance.[33]

[32] In an ambitious and important paper, Ricardo Caballero and Adam Jaffe (1993) make some progress in identifying proxies for ideas and knowledge spillovers.

[33] The microeconomic evidence is often felt to be unsatisfactory because the variables of interest, like education and fertility, are endogenous within a household demand framework.

Most recent growth researchers have included population growth as a variable of subsidiary interest, perhaps in the manner suggested by MRW, and then noted a weak negative correlation between it and the growth of per capita income. This raises some endogeneity concerns, although typically one might think of causality running to population growth from the level of per capita income rather than its rate of growth.

Most economists, asked to suggest negative consequences of population growth, would probably argue that average human capital and capital-labour ratios are likely to fall. In the context of MRW's work, there is a negative effect of population growth on the steady state capital-labour ratio for a given investment rate, but it makes only a small contribution to the observed variation in levels of income and growth rates. There is some evidence that students in countries with higher population growth record lower achievement (Hanushek 1992). There is also a weak negative relation between population growth and changes in total factor productivity, perhaps something of a puzzle for theorists.

More generally, it is clear that in-depth studies are needed to address the links between population growth and macroeconomic outcomes. The work so far has established one key finding: the small negative effect of population growth on per capita growth is partly mediated through changes in labour force participation.[34] These shifts in the composition of the labor force may have only second-order welfare effects.

Some researchers have looked at the link from fertility rates to subsequent growth, sometimes finding a negative correlation. Again, one must think carefully about causality. If people perceive that incomes are likely to rise, and possibly the returns to human capital, they may decide to have fewer children. Another worry is that fertility decisions appear to be quite strongly correlated with various wider aspects of social development. It may be that high fertility rates, rather than having a direct negative effect, simply proxy for social arrangements that are relatively "traditional" in other respects.[35]

Overall, the popular belief that population growth is economically harmful is not yet well supported by statistical evidence, and overall the new growth evidence has done little to modify the conclusions of the influential survey by Allen Kelley (1988). He argues that some government policies are likely to exacerbate the negative effects of population growth, and this would be an interesting area to explore. It should also be remembered that population growth may have adverse consequences beyond those for per capita income, on such things as access to safe water, and environmental quality more generally.

8.2 *Trade and Growth*

Understanding the interaction between trade policies and macroeconomic performance has traditionally been one of the central concerns of development economics. Often it seems that microeconomic studies allow much sharper discrimination between hypotheses than the aggregate studies of trade and growth. However, researchers have long been aware that micro studies often miss the economy-wide resource allocation effects that may be central to understanding the effects of trade policy.

This awareness partly explains why

[34] The key papers here are James Brander and Dowrick (1994) and Pritchett (1996a).

[35] Temple and Johnson (1998) discuss the possible role of social factors. See also Section 8.8 below.

the first cross-country studies of openness and growth considerably pre-date much of the rest of the empirical growth literature. The early start does not seem to have conferred much advantage, and in this area as in others, the usual suspects—endogeneity, specification problems, outliers—are still at large. Since Sebastian Edwards (1993) covers the field in some detail, I will only briefly consider the issues.

One of the most important difficulties is finding some way of quantifying trade regimes. The measures which are most defensible on theoretical grounds, such as effective rates of protection, can be difficult to calculate for a sufficiently large number of countries. Typically, researchers fall back on simple proxies, such as trade shares in GDP or the black market exchange rate premium, meant to give some indication of openness. Sachs and Warner (1995a) create a binary dummy variable based on judging trade regimes against a set of such criteria. They find some striking results, including unconditional convergence among the group of open economies.

Causality is another key problem in the study of trade and growth. It is easy to think of ways in which fast growth may lead to, say, an increase in the trade share.[36] One interesting possibility is to use trade shares predicted by a gravity model as instruments for actual trade shares (Jeffrey Frankel, David Romer and Teresa Cyrus 1996). They find that this tends to raise the suggested importance of trade. However, the gravity model predicts high shares for most East Asian countries, so policies for outward orientation (reflected in actual trade shares above predicted) do not emerge as central except for Singapore and Malaysia.

These articles are exceptions, and if macroeconomic studies are to be genuinely useful, the measurement and simultaneity problems often need to be better addressed. One way of doing this will be through finding natural experiments. As an example, Ben-David (1993) documents the observed convergence effect of trade liberalization by the European Economic Community, and shows that it represents neither a continuation of a long-run trend nor a return to the pre-war extent of dispersion.

This kind of work may clarify the conditions under which openness is most beneficial. Recent endogenous growth models have emphasised that open economies, if they specialise according to comparative advantage, may do less well than under autarky. There are several qualifications to these arguments, especially before translating them into activist trade policy, but recent empirical work indicates that the ideas have some relevance for developing countries.[37] The gains from openness may be greatest for countries already specialising in manufacturing exports, like those in East Asia.

8.3 *Finance and Growth*

If cross-country work is going to be useful, it should sometimes succeed in shifting priors. One area in which this may have already happened is the empirical literature on finance and growth. Some prominent economists have tended to dismiss the role of financial factors, arguing (from rather little evidence) that financial development is simply a passive consequence of growth. There are now a number of papers

[36] Some researchers work with initial shares; a remaining problem is that these initial shares may be correlated with unmeasured aspects of domestic policy.

[37] See for instance the evidence of Sachs and Warner (1995b) that natural resource abundance works against long-run growth.

looking at the effect on growth of initial financial development, including both banking system and stock market variables. Taken together, these studies make those traditional arguments harder to justify.

The survey by Levine (1997) covers this area in some detail, so here the issues are only briefly addressed. To be useful, regressions of growth on financial systems must take into account the endogeneity of financial development. Most studies, including the influential work of King and Levine (1993), have done this. Hence perhaps the major worry with these results is the fixed effects difficulty discussed earlier.

As in the study of trade effects, causality problems can sometimes be addressed by natural experiments. In particular, it will be essential to investigate the consequences of financial liberalizations across developing countries, including the rise of stock markets. Work in Levine and Sara Zervos (1997) suggests that active stock markets play a role in subsequent growth, and since they examine the value of trades while controlling for capitalization, as well as looking at turnover, their re- sults are unlikely to reflect only the forward-looking nature of stock prices.

8.4 *Short-Run Macroeconomics*

One of the most controversial areas is the link between short-run macroeconomic management and long-run growth. It is not difficult to present regressions which show significant correlations, of the expected signs, between the growth rate and variables such as budget deficits, inflation, and real exchange rate instability. It is rather harder to isolate any particular policy variable and demonstrate that it has a robust correlation with growth, regardless of endogeneity concerns and the selection of other variables. A common conclusion from this literature is that although "policy matters" we do not yet have any clear idea which elements of policy are crucial.

Given the importance of this information for policy-makers, our lack of progress on this question is disappointing. Yet it may be an inevitable consequence of cross-country studies. The central problem is that things tend to go wrong together: high inflation is accompanied by political instability, exchange rate volatility, and so on, and disentangling the various contributions is not easy. Given that timing evidence would be very useful, this is one field in which a very strong case for historical case studies, in preference to regressions, can be made.

It is also worth pointing out that many of these growth-and-policy studies control for the accumulation of physical capital. It may be that the most important and robust links between policy and growth are mediated through investment. The focus in the policy literature could perhaps shift to explaining cross-country differences in investment.

The links between inflation and growth are particularly controversial. Most studies have looked at the cross-section but it is increasingly clear that such results can be distorted by just one or two high-inflation outliers. Michael Bruno and Easterly (1998) chose instead to take a short run approach, and examine the path of output during high inflation crises. They found that high inflation crises are associated with output losses, but that output returns to the same long-run growth path once inflation has been reduced. This may be the reason for the weak inflation and growth relation in the cross-section.

From a methodological point of view, some of the most interesting studies are

those making use of the time series properties of inflation and output. The strategy is to emphasize that inflation is highly persistent, to the extent that it may be I(1). This finding implies that, should inflation affect growth, some shocks to the growth rate will be permanent, a result that is rarely found. Clearly strong conclusions here are built on a somewhat shaky foundation: the finding that inflation may have a unit root inevitably uses tests with low power.

Arguably more reliable is the research on short-run output volatility and long-run growth. Usually economists have reserved separate compartments for the analysis of business cycles and growth. Maybe that was wrong: some recent theories and evidence suggest that there are deep connections between the two. There appears to be a negative relationship between output volatility and growth in the OECD, and in samples that combine the OECD and developing countries (Garey Ramey and Valerie Ramey 1995). As yet, the interpretation of the findings is unclear, but this is a fascinating area for further theoretical and empirical work, especially on the links between uncertainty and growth.

8.5 *Government Size*

Among questions that are very important but likely to remain largely unsettled, that of government size and growth looms large. In political discussion it is common to hear claims that a high ratio of social security transfers to GDP and a high level of government consumption can be damaging to growth prospects. The evidence is not strong. Some researchers find a negative link between government consumption and growth, but overall studies disagree, and it would be wrong to argue that a correlation between small government and fast growth leaps out from the data.

In particular it is worth noting that government consumption is one of the variables whose correlation with growth is identified as fragile by Levine and Renelt (1992). As for the effect of social security transfers, Anthony Atkinson (1995) points out that the detailed structure of welfare state institutions is likely to be crucial, and cross-section studies will only succeed in obscuring the most important issues.[38]

Overall, this seems to be one of those fields in which stylized facts appear hard to come by, at least at the aggregate level. Microeconomic evidence on labour supply and investment responses to changes in tax rates is likely to be more fruitful. Perhaps the only promising macro approach is that of Hall and Jones (1997), who find that high government consumption lowers the level of income, and also point out that in their framework, endogeneity is likely to mean that the effect is understated.

8.6 *Government Spending on Infrastructure*

More progress has been made on the question of the composition of government spending. There has been intense debate surrounding the role of public capital in the growth of developed countries (Edward Gramlich 1994). This debate has a quieter counterpart in the cross-country literature. Although there is a lack of data on the quantity and quality of public infrastructure in developing countries, measures of telephone networks and electricity capacity have been found to have a significant effect on subsequent growth. Easterly and Rebelo (1993) find that the share of public investment in transport and

[38] Other important discussions are Easterly and Rebelo (1993) and Joel Slemrod (1995).

communication is robustly correlated with growth. Infrastructure spending appears to raise the social return to ongoing investment, rather than spurring further efforts.

There is an endogeneity problem here, and one that instrumental variables usually cannot solve. Infrastructure investments may be made in anticipation of fast growth, rather than playing a causal role. Even so, it seems easier to make headway on this issue than on the more controversial one of government size.

8.7 *Inequality*

One of the most active areas of cross-country research has been investigating the consequences of inequality for growth. Somewhat unusually for the growth literature, studies have tended to concur in finding a negative effect of high inequality on subsequent growth.[39] The evidence has not been accepted by all: some writers point out the concentration of richer countries at the lower end of the inequality spectrum, the poor quality of the distribution data, and the lack of robustness to fixed effects specifications.[40] At least, though, it has become extremely difficult to build a case that inequality is good for growth. This in itself represents a considerable advance.

Given the indications that inequality is harmful for growth, attention has moved on to the likely mechanisms. The best known line of argument is the political economy one, in which democratic governments are under more pressure to redistribute income in unequal societies, because of the relatively low income of the median voter. The argument has the advantage of clarifying a key point: just because high inequality lowers growth does not necessarily mean that one should redistribute. It may be redistribution which drives the correlation in the first place.

That said, the political economy arguments have not done particularly well empirically. There is not much evidence that fiscal policy is strongly related to inequality. Testing a central implication, that the inequality-growth relation differs between democracies and non-democracies, is made difficult by the fact that stable democracies are usually found only among the richer countries. Hence the literature seems to be moving away from the political economy line, towards an examination of the effects of inequality on fertility rates, investment in education, and political stability.

8.8 *Social and Political Factors*

Some of the most interesting thinking on economic growth is to be found on the borders of political science and sociology. Of the two areas, political factors have received more consideration in the empirical growth literature, probably because they lend themselves to measurement more easily, and because the lines of causation are better understood.

A common approach has been to relate growth to indices of civil, political, and economic rights. Studies of this kind tend to leave open the question of which type of rights are most important. Some have argued that political and economic rights are complementary, but there are counter-examples of autocratic regimes that have extended economic freedoms.

At present, economic freedom seems

[39] Roland Benabou (1996) surveys this work in depth. Roberto Perotti (1996) is another very good introduction, and the summary here relies heavily on his evidence and arguments.

[40] Note that fixed effects specifications may not be particularly useful here. Inequality is usually found to be fairly stable over time, and the lags in its effects may well be too long to be picked up by the panel data analyses that have been carried out.

to have a stronger observable link to growth than political rights. Researchers use indicators of property rights provided by country risk evaluators to potential foreign investors.[41] The indicators, which evaluate such things as the ease of enforcing contracts and the risk of expropriation, are found to have a relatively strong connection to growth. Other studies of institutional factors tend to find that, even when good institutions do not raise growth through technical progress, they often have an effect in encouraging investment.

There is an emerging consensus that the link between democracy and growth is difficult to isolate. Casual empiricism suggests a wide variety of experience under both authoritarian and democratic regimes. Some studies have found positive effects of political rights, but as many if not more fail to find a significant correlation. Barro (1997) suggests that this may be because the relation is nonlinear: extensions of political rights appear to be beneficial only up to a certain point. Ultimately there is not a great deal in the empirical literature to support the positions either of instinctive democrats or those who advocate "strong states" isolated from the populist and lobbying pressures sometimes associated with democratic freedoms.

It may be that the type of regime is far more important—whether, for instance, those running autocratic regimes are bent on self-interest and accumulating wealth, or whether they are oriented to national economic goals. In general, empirical assessment of political regimes clearly presents some formidable measurement problems. There are many conceptual and measurement traps just in quantifying democracy; since, however, the indices devised by different methods are often highly correlated, there does remain some hope for this research.[42]

Also interesting is the role of social and political instability. The excellent survey by Alberto Alesina and Perotti (1994) concludes that it is here, rather than in the question of democratic freedoms, that the most important political influences on growth are to be found. Simple proxies for political instability have been in use in cross-country regressions almost from the start. The key problem with most results is that growth and stability are likely to be jointly determined. However, papers that explicitly address this difficulty tend to confirm the idea that instability has a strong negative effect. One issue for further work will be to draw a clearer distinction between the effects of civil war and those of collective protest.

Also interesting is the impact of social arrangements on growth. Laymen are often willing to suggest that there are sometimes fundamental social and cultural barriers to growth, and there is support for this idea in Marx, Weber and Kuznets, among others. More recently, theoretical papers have started to appear that draw out the relation between social arrangements and various kinds of incentives. Some economists distrust an emphasis on "social capital" or "social capability" since, as Moses Abramovitz (1986) has pointed out, "no one knows just what it means or how to measure it."

Researchers are starting to make some progress on the measurement issue. One promising approach is to model social and political influences on growth as latent variables related to a

[41] Evaluations provided by private consultancies have been used to good effect by Stephen Knack and Philip Keefer (1995) and Paolo Mauro (1995).

[42] Kenneth Bollen (1990) discusses measurement issues in more detail.

variety of observable indicators. Despite some interest from development economists, simple techniques for data reduction like factor analysis and principal components have been largely ignored by recent growth researchers.[43] Their use seems to have a great deal of potential, and the renewed interest in social factors aligns well with recent theoretical work, reinforcing the case for further study.

9. *Where Next?*

In summarising the recent literature, I have often indicated areas in which there is much scope for further work, and the direction it might take. In this section, I will consider some broader questions about the future direction of the empirical growth literature. There seems to be general agreement that empirical testing of research-driven growth models, perhaps building on earlier research in applied microeconomics, has much to recommend it. There is far less consensus on the future status of cross-country empirical work.

9.1 *The Prospects for Different Methods*

The reasons for this lack of consensus should be clear by now. Even the most enthusiastic proponent of cross-country regressions must acknowledge that we are a daunting distance from the ultimate goal, a model with high explanatory power which indicates with precision the relative contributions of different influences. For the most part, researchers are satisfied with isolating relationships that are statistically significant and in some sense quantitatively 'strong'. Few attempt the more ambitious task of trying to place bounds on the likely contributions of a set of variables.

For the last few years, commentators have been claiming that the Summers-Heston data set has been mined to the point of exhaustion, and implying that there is nothing new to be learned from further work of this kind. From some published comments, it seems that "regression fatigue" has set in, so that hearts sink when yet another dubious growth regression is presented. Despite this pessimism, researchers in the field continue to show some ingenuity in finding new and interesting variables to combine with the data set. Recent studies have looked at, among other things, the growth effects of natural resource abundance, social factors, protection for intellectual property rights, the quality of infrastructure, and the composition of government spending.

The genuine worry is not that the core data set has been exhausted, but that this kind of relatively simple aggregate work may drive out other promising approaches, in a kind of Gresham's Law of Growth Studies. Clearly, there is a danger of this happening. Yet those who try to rule out cross-country empirical work rarely seem to have thought much about the problems of alternatives. There is a now routine call for more detailed studies of individual countries' experience, but we have had studies of this kind for years, sometimes without being able to draw many useful generalizations. In particular, it is almost impossible, from the experience of just a few countries, to quantify the effect of different variables on long-run growth.

Even so, historical analysis must be a major way forward. Amongst other things, historians may be able to identify some interesting natural experiments. There are important examples of influential historical growth studies,

[43] Studies making use of latent variables include Irma Adelman and Cynthia Taft Morris (1968), Eberhard Scholing and Vincenz Timmermann (1988), and Temple and Johnson (1998).

notably those of experiences with differing trade policies. There are areas which have been largely neglected by the new growth evidence, like the role of external debt, where the use of case studies is essential. Historical analyses may be particularly useful in identifying reasons for the instability of growth within countries; as noted earlier, the history of many developing countries has been marked by alternating booms and growth collapses (Pritchett 1998).

As the earlier quotation from Gerschenkron was meant to suggest, those with a knowledge of history can identify factors of potential relevance and generate interesting hypotheses, and thus the two approaches have much to learn from each other. So far there has been disappointingly little dialogue.[44]

Traditional growth accounting can also be useful for certain questions. These methods have been put to good use in describing the experience of small groups of countries, as in the work of Alwyn Young (1992, 1995) on East Asian growth. The next task will be to try and derive comparable TFP growth figures for a larger number of countries, a necessary first step in explaining the variation in productivity growth across countries.

9.2 *Open Research Questions*

There will continue to be a place for cross-country work. With new and more thoughtful methods, there is much left to learn. As pointed out earlier, we may have a better idea of how factor inputs influence growth, but our understanding of the international variation in factor accumulation is still weak. Put in more concrete terms, we know that physical capital investment has been important to growth in Singapore, but why has Singapore's investment ratio been so high? We have learnt something about the general question from Levine and Renelt (1992) and Barro and Sala-i-Martin (1995), among others, but one persistent difficulty is that explanations for differences in, say, investment ratios, tend to be rather *ad hoc*. Compared to growth regressions, there is even less consensus on the variables that are essential to explaining differences in rates of factor accumulation.

There is another issue that should be central to the literature, but which is rarely even acknowledged. Development economists have long distinguished between different types of growth, and their varying consequences for welfare. Again putting the question in more concrete terms, we know that East Asian countries have managed to combine fast growth with low inequality, but we lack a good understanding of how this type of growth has been achieved. This kind of question is a central challenge for the many development economists, economic historians, theorists and empirical macroeconomists currently working on growth.

Connected to this is the fundamental question of whether growth will raise welfare. Useful progress can be made here by relating indices of "human development" and quality of life to macroeconomic variables. The intuition of most economists is that there will be a strong connection between welfare indices and per capita income, but given the controversy surrounding certain policies, such as those for structural adjustment, there is plenty of room for further research. Easterly (1997) finds that the effects of growth on a wide range of indicators are surprisingly uneven.

[44] A notable exception is the book edited by Nicholas Crafts and Gianni Toniolo (1996). This represents perhaps the first major attempt to integrate the lessons of the new growth theory and evidence with the economic history of the advanced countries.

As well as classifying and analysing differing types of growth, development economists have sometimes found it useful to draw distinctions between types of countries. Ranis and John Fei (1988) argue that it is important to move away from characterizing the "average" developing country, and work towards a deeper understanding of differences. Along similar lines, Pritchett (1997) and Quah (1997) have emphasised the need to acknowledge heterogeneity, and move away from techniques based on "representative" economies.

Quah's work makes clear that it is possible to approach the data with more imaginative methods than are usually adopted in the cross-country literature. In general, careful thought about how countries and their experiences differ, drawing on historical studies, may suggest new and useful ways of examining the data. One example of a thoughtful approach is Easterly (1994). He points out that a surprisingly high number of countries have stagnated at some point since the 1960s, and attempts to construct an empirical model explaining stagnation.

Similarly, Rodrik (1998) seeks to explain why growth collapsed in many countries in the 1980s. Given the instability of growth, noted in Section 2, investigations of growth variation over time could have a high payoff. This is one area where case studies and panel data analyses may both have something to contribute, allowing identification of the key variables that determine whether or not growth can be sustained.

Another challenge will be to identify the empirical relevance of theoretical work on multiple equilibria, and examine the possibility of "development traps." At present, theorists justify their frequent emphasis on multiple equilibria with some rather rudimentary stylized facts; more work on deriving and testing implications would clearly be useful. At present, it is difficult to distinguish these models empirically from more orthodox explanations for underdevelopment.

Another issue that has not been sufficiently addressed in the macroeconomic literature is that of structural transformation. The study of sectoral shifts has a long history in development economics.[45] Few policy recommendations have been forthcoming, which has perhaps led some macroeconomists to neglect structural change. Given that many of the countries under study go through significant changes as they industrialize, it will sometimes be important to take this into account if conventional cross-country work is not to give misleading results. For instance, Temple and Hans-Joachim Voth (1998) argue that this problem may have occurred in the study of equipment investment.

9.3 *The Role of Aggregate Production Functions*

At present, there do remain some long-term concerns with the whole cross-country research endeavor. The mention of structural change leads naturally to questioning the relevance of aggregate production functions. Arguably the aggregate production function is the least satisfactory element of macroeconomics, yet many economists seem to regard this clumsy device as essential to an understanding of national income levels and growth rates.

One could debate the need to divide growth between the accumulation of inputs and improvements in technology. For some questions of interest, this division is unnecessary, and should

[45] This work is surveyed by Hollis Chenery (1988). A paper by Robinson (1971) estimates cross-country regressions allowing for structural change.

probably be avoided. As was emphasized earlier, if the focus of interest is a policy variable like inflation or the budget position, it will often be preferable to omit factor accumulation altogether and concentrate on the overall growth effect of policy measures. If a sufficiently complete model is used, this approach should be very useful. It is often the overall effect that is of principal interest and, as a side benefit, fewer variables need to be handled.

Sometimes, though, we would like to know whether a growth effect acts through factor accumulation or efficiency change. If answers using macro data require us to make incredible assumptions, this naturally suggests that we should be working at a more disaggregate level. Useful data sets of this kind are hard to come by for developing countries, and further problems lie in synthesising results for different industries and countries to get an overall picture of what drives growth. Yet there can be little doubt that this older approach, in which a branch of applied microeconomics meets development, is also an important way ahead.

10. *Conclusions*

One of the themes of this paper has been that cross-country research provides a useful complement to other investigative techniques, and for some questions of interest, it may even be the only way forward. I make this claim despite the scepticism that is possibly invited by the discussion above. Those who instinctively distrust the new growth evidence may well have discovered further support for their position.

It is certainly true that, taken as a whole, the growth literature can seem something of a disappointment. Conclusions that once seemed well established have been overturned. We have learned some useful things about the study of convergence, but some of the most influential papers take an approach to international technology diffusion which has little support from elsewhere. The most interesting findings are rarely convincing, while the more reliable ones hold few surprises.

What then have we learned? I return to the six questions posed in the introduction, and answer them as follows. Poor countries are not catching up with the rich, and to some extent the international income distribution is becoming polarized. Countries do converge to their own steady states, but at an uncertain rate. One reason for this uncertainty is that countries catch up by adopting technologies from abroad, as well as by investing in physical capital and education. It is easy to envisage a hypothetical long-run equilibrium in which countries grow at the same rate, but over the last thirty years, rates of efficiency growth have almost certainly varied widely.

How compatible is this view with existing growth models? The Solow-Swan model is almost certainly correct in assuming that returns to physical capital are diminishing. However, it would be a long way off the mark to assert that the new growth evidence only succeeds in demonstrating the explanatory power of the original neoclassical growth model. A wide variety of variables have been shown to affect growth over thirty year time spans, often through systematic differences in rates of efficiency growth across countries. These differences are left unexplained by the Solow-Swan model, but are surely important in understanding the variation in experience over the last thirty years.

Another implication should be noted. Either growth is endogenous, or it is exogenous and level effects are large. Given the presence of large level

effects, distinguishing between exogenous and endogenous growth models is not as pressing as it might seem. The important point is that policy can have a major impact on a country's level of welfare. As pointed out earlier, the debate on whether policy affects the long-run growth rate or just the steady state level of income is almost impossible to resolve, and not much of practical importance will turn on it.

Arcane discussions about the validity of endogenous growth theories are likely to frustrate the policymaker in search of knowledge that can be put to good use. Those forming policies would ideally like to find a set of instructions that is short, clear and easy to implement. It is a foolhardy researcher who responds to this desire by condensing many books and articles into a one paragraph summary, but somebody perhaps ought to try. A quick overview of why growth rates differ is a useful way to close.

A key reason why growth rates differ across countries is that macroeconomic stability differs across countries. This effect partly acts through capital investment, and equipment investment may have a special role. As yet the growth benefits of education are imprecisely measured. The social returns to R&D are high, and even if the long-run growth rate is independent of research efforts, the welfare effects of changes in R&D expenditure can be large. Population growth does not seem to have the large negative effects that are frequently conjectured. High inequality lowers growth, perhaps by raising social and political instability. The depth of financial intermediation seems important to subsequent development. Democracies do not do noticeably better than autocratic regimes, but countries that extend economic freedoms and protect property rights grow faster. Big government and high taxation may have a negative effect, but the evidence is still somewhat ambiguous. Government spending on infrastructure is beneficial. Openness to trade also appears to be a good thing, although we do not yet know enough about the conditions under which this is true.

In my view, those are the main findings to emerge. Some, like those on finance and on inequality, tend to contradict the earlier conventional wisdom. Others, like those on the importance of infrastructure, reinforce it. In some areas, such as the size of government and the effect of democracy, the main point to note is the lack of support for positions that are often too fiercely held. At least some knowledge of the average pattern is the beginning of wisdom, and although we have not learnt as much as might be hoped, it is always worth remembering how little we knew when we started.

References

Abramovitz, Moses. 1986. "Catching Up, Forging Ahead, and Falling Behind," *J. Econ. Hist.*, 46:2, pp. 385–406.

Adelman, Irma and Cynthia Taft Morris. 1968. "An Econometric Model of Socioeconomic and Political Change in Underdeveloped Countries," *Amer. Econ. Rev.*, 58, pp. 1184–218.

Ades, Alberto and Hak B. Chua. 1997. "Thy Neighbor's Curse: Regional Instability and Economic Growth," *J. Econ. Growth*, 2:3, pp. 279–304.

Aghion, Philippe and Peter Howitt. 1998. *Endogenous Growth Theory*. Cambridge MA: MIT Press.

Alderman, Harold; Jere R. Behrman, David R. Ross, and Richard Sabot. 1996. "The Returns to Endogenous Human Capital in Pakistan's Rural Wage Labour Market," *Oxford Bul. Econ. Statist.*, 58:1, pp. 29–55.

Alesina, Alberto and Roberto Perotti. 1994. "The Political Economy of Growth: A Critical Survey of the Recent Literature," *World Bank Econ. Rev.*, 8:3, pp. 351–71.

Atkinson, Anthony B. 1995. "The Welfare State and Economic Performance," *Nat. Tax J.*, 48:2, pp. 171–98.

Barro, Robert J. 1991. "Economic Growth in a Cross-section of Countries," *Quart. J. Econ.* 106:2, pp. 407–43.

———. 1997. *Determinants of Economic Growth: A Cross-Country Empirical Study*. Cambridge, MA: MIT Press.

Barro, Robert J. and Xavier Sala-I-Martin. 1992. "Convergence," *J. Polit. Econ.*, 100:2, pp. 223–51.

———. 1995. *Economic Growth*. Boston: McGraw-Hill.

Baumol, William J. 1986. "Productivity Growth, Convergence, and Welfare: What the Long-Run Data Show," *Amer. Econ. Rev.*, 76:5, pp. 1072–85.

Benabou, Roland. 1996. "Inequality and Growth," in *NBER Macroeconomics Annual*. Ben Bernanke and Julio Rotemberg, eds. London and Cambridge: MIT Press, pp. 11–74.

Ben-David, Dan. 1993. "Equalizing Exchange: Trade Liberalization and Income Convergence," *Quart. J. Econ.*, 108:3, pp. 653–79.

Ben-David, Dan and A. K. M. Atiqur Rahman. 1996. "Technological Convergence and International Trade," CEPR Discussion Paper No. 1359.

Benhabib, Jess and Jordi Gali. 1995. "On Growth and Indeterminacy: Some Theory and Evidence," *Carnegie-Rochester Conference Series on Public Policy*, 43, pp. 163–211.

Benhabib, Jess and Boyan Jovanovic. 1991. "Externalities and Growth Accounting," *Amer. Econ. Rev.*, 81:1, pp. 82–113.

Benhabib, Jess and Mark M. Spiegel. 1994. "The Role of Human Capital in Economic Development: Evidence from Aggregate Cross-Country Data," *J. Monet. Econ.*, 34:2, pp. 143–73.

Bernard, Andrew B. and Charles I. Jones. 1996a. "Productivity Across Industries and Countries: Time Series Theory and Evidence," *Rev. Econ. Statist.*, 78:1, pp. 135–46.

———. 1996b. "Technology and Convergence," *Econ. J.*, 106:437, pp. 1037–44.

Bhagwati, Jagdish and Bent Hansen. 1972. "Should Growth Rates Be Evaluated at International Prices?" in *Development and Planning*. Jagdish Bhagwati and Richard S. Eckaus, eds. London: George Allen and Unwin, p. 53–68.

Binder, Michael and M. Hashem Pesaran. 1996. "Stochastic Growth," U. Cambridge, Dept. Applied Econ. working paper no. 9615.

Bollen, Kenneth A. 1990. "Political Democracy: Conceptual and Measurement Traps," *Stud. Comp. Int. Devel.*, 25:1, pp. 7–24.

Brander, James and Steve Dowrick. 1994. "The Role of Fertility and Population in Economic Growth: Empirical Results from Aggregate Cross-National Data," *J. Population Econ.*, 7:1, pp. 1–25.

Bruno, Michael and William Easterly. 1998. "Inflation Crises and Long-Run Growth," *J. Monet. Econ.*, 41:1, pp. 3–26.

Caballero, Ricardo J. and Adam B. Jaffe. 1993. "How High Are the Giants' Shoulders?" in *NBER Macroeconomics Annual*. Olivier Jean Blanchard and Stanley Fischer, eds. London and Cambridge: MIT press, pp. 15–74.

Canjels, Eugene and Mark W. Watson. 1997. "Estimating Deterministic Trends in the Presence of Serially Correlated Errors," *Rev. Econ. Statist.*, 79:2, pp. 184–200.

Carroll, Christopher D. and David N. Weil. 1994. "Saving and Growth: A Reinterpretation," *Carnegie-Rochester Conf. Ser. Public Policy*, 40, pp. 133–92 .

Caselli, Francesco; Gerard Esquivel, and Fernando Lefort. 1996. "Reopening the Convergence Debate: A New Look at Cross-Country Growth Empirics," *J. Econ. Growth*, 1:3, pp. 363–90.

Chenery, Hollis B. 1988. "Structural Transformation: A Program of Research," in *The State of Development Economics*. Gustav Ranis and T. Paul Schultz, eds. Oxford: Basil Blackwell.

Cho, Dongchul and Stephen Graham. "The Other Side of Conditional Convergence," *Econ. Letters*, 50, pp. 285–90.

Coe, David T. and Elhanan Helpman. 1995. "International R&D Spillovers," *Europ. Econ. Rev.*, 39:5, pp. 859–87.

Crafts, Nicholas F. R. and Gianni Toniolo, eds. 1996. *Economic Growth in Europe since 1945*. Cambridge: Cambridge U. Press.

Deaton, Angus and Ron Miller. 1996. "International Commodity Prices, Macroeconomic Performance and Politics in Sub-Saharan Africa," *J. African Econ.*, pp. 99–191.

Deininger, Klaus and Lyn Squire. 1996. "A New Data Set Measuring Income Inequality," *World Bank Econ. Rev.*, 10:3, pp. 565–91.

De Long, J. Bradford. 1997. "Cross-Country Variations in National Economic Growth Rates: the Role of 'Technology'," in *Technology and Growth*. Jeffery Fuhrer and Jane Sneddon Little, eds. Federal Reserve Bank of Boston.

De Long, J. Bradford and Lawrence H. Summers. 1991. "Equipment Investment and Economic Growth," *Quart. J. Econ.*, 106:2, pp. 445–502.

Dowrick, Steve and Duc-Tho Nguyen. 1989. "OECD Comparative Economic Growth 1950–85: Catch-up and Convergence," *Amer. Econ. Rev.*, 79:5, pp. 1010–30.

Durlauf, Steven N. and Paul A. Johnson. 1995. "Multiple Regimes and Cross-Country Growth Behaviour," *J. Appl. Econometrics*, 10:4, pp. 365–84.

Durlauf, Steven N. and Danny T. Quah. Forthcoming. "The New Empirics of Economic Growth," in *Handbook of Macroeconomics*. John Taylor and Michael Woodford, eds.

Easterly, William. 1994. "Economic Stagnation, Fixed Factors, and Policy Thresholds," *J. Monet. Econ.*, 33:3, pp. 525–57.

———. 1997. "Life During Growth," World Bank manuscript.

Easterly, William and Ross Levine. 1997. "Africa's Growth Tragedy: Policies and Ethnic Divisions," *Quart. J. Econ.*, 112, pp. 1203–50.

———. 1998. "Troubles with the Neighbours: Africa's Problem, Africa's Opportunity," *J. African Econ.*, 7:1, pp. 120–42.

Easterly, William and Sergio Rebelo. 1993. "Fiscal Policy and Economic Growth," *J. Monet. Econ.*, 32:3, pp. 417–58.

Eaton, Jonathan and Samuel Kortum. 1994. "International Patenting and Technology Diffusion," NBER working paper no. 4931.

———. 1995. "Engines of Growth: Domestic and Foreign Sources of Innovation," NBER working paper no. 5207.

Edwards, Sebastian. 1993. "Openness, Trade Liberalization, and Growth in Developing Countries," *J. Econ. Lit.*, 31:3, pp. 1358–93.

Evans, Paul. 1996. "Using Cross-Country Variances to Evaluate Growth Theories," *J. Econ. Dynam. Control*, 20, pp. 1027–49.

———. 1997. "How Fast Do Economies Converge?" *Review Econ. Statist.*, 79:2, pp. 219–25.

Fagerberg, Jan. 1994. "Technology and International Differences in Growth Rates," *J. Econ. Lit.*, 32, pp. 1147–75.

Frankel, Jeffrey A.; David Romer, and Teresa Cyrus. 1996. "Trade and Growth in East Asian Countries: Cause and Effect?" NBER working paper no. 5732.

Gemmell, Norman. "Evaluating the Impacts of Human Capital Stocks and Accumulation on Economic Growth: Some New Evidence," *Oxford Bul. Econ. Statist.*, 58:1, pp. 9–28.

Gerschenkron, Alexander. 1952. "Economic Backwardness in Historical Perspective," in *The Progress of Underdeveloped Areas*. Bert F. Hoselitz, ed. Chicago: U. Chicago Press. Reprinted in A. Gerschenkron, *Economic Backwardness in Historical Perspective*, New York: Frederick A. Praeger, 1962.

Gramlich, Edward M. 1994. "Infrastructure Investment: A Review Essay," *J. Econ. Lit.*, 32:3, pp. 1176–96.

Griliches, Zvi. 1979. "Issues in Assessing the Contribution of R&D to Productivity Growth," *Bell J. Econ.*, 10:1, pp. 92–116.

———. 1992. "The Search for R&D Spillovers," *Scand. J. Econ.*, Supplement, 94, S29–47.

Griliches, Zvi and Jacques Mairesse. 1995. "Production Functions: The Search for Identification," NBER manuscript.

Hall, Robert E. and C. I. Jones. 1997. "Fundamental Determinants of Output per Worker across Countries," Stanford U., manuscript.

Hanushek, Eric A. 1992. "The Trade-off between Child Quantity and Quality," *J. Polit. Econ.*, 100:1, pp. 84–117.

Hanushek, Eric A. and Dongwook Kim. 1995. "Schooling, Labor Force Quality, and Economic Growth," NBER working paper no. 5399.

Harberger, Arnold C. 1987. "Comment," *NBER Macroeconomics Annual*. Stanley Fischer, ed. London and Cambridge: MIT Press, pp. 255–58.

Heston, Alan. 1994. "A Brief Review of Some Problems in Using National Accounts Data in Level of Output Comparisons and Growth Studies," *J. Devel. Econ.*, 44:1,29–52.

Heston, Alan and Robert Summers. 1988. "What We Have Learned about Prices and Quantities from International Comparisons," *Amer. Econ. Rev.*, 78:2, pp. 467–73.

———. 1996. "International Price and Quantity Comparisons: Potentials and Pitfalls," *Amer. Econ. Rev.*, 86:2, pp. 20–24.

Islam, Nazrul. 1995. "Growth Empirics: A Panel Data Approach," *Quart. J. Econ.*, 110:4, pp. 1127–70.

Jones, Charles I. 1995a. "Time Series Tests of Endogenous Growth Models," *Quart. J. Econ.*, 110:2, pp. 495–525.

———. 1995b. "R&D-Based Models of Economic Growth," *J. Polit. Econ.*, 103:4, pp. 759–84.

———. 1997a. "Convergence revisited," *J. Econ. Growth*, 2:2, pp. 131–53.

———. 1997b. "On the Evolution of the World Income Distribution," *J. Econ. Perspectives*, 11:3, pp. 19–36.

Jones, Charles I. and John C. Williams. 1997. "Measuring the Social Return to R&D," Stanford U., manuscript.

Kakwani, Nanak. 1997. "Growth Rates of Per-Capita Income and Aggregate Welfare: An International Comparison," *Rev. Econ. Statist.*, 79:2, pp. 201–11.

Kelley, Allen C. 1988. "Economic Consequences of Population Change in the Third World," *J. Econ. Lit.*, 26:4, pp. 1685–728.

King, Robert G. and Levine, Ross. 1993. "Finance and Growth: Schumpeter Might Be Right," *Quart. J. Econ.*, 108:3, pp. 717–37.

———. 1994. "Capital Fundamentalism, Economic Development and Economic Growth," *Carnegie-Rochester Conf. Ser. Public Policy*, 40, pp. 259–92.

King, Robert G. and Sergio T. Rebelo. 1993. "Transitional Dynamics and Endogenous Growth in the Neoclassical Model," *Amer. Econ. Rev.*, 83:4, pp. 908–31.

Klenow, Peter J. and Andres Rodriguez-Clare. Forthcoming. "The Neoclassical Revival in Growth Economics: Has It Gone Too Far?" *NBER Macroeconomics Annual*.

Knack, Stephen and Keefer, Philip. 1995. "Institutions and Economic Performance: Cross-Country Tests Using Alternative Institutional Measures", *Econ. Polit.*, 7:3, pp. 207–27.

Kocherlakota, Narayana R. and Kei-Mu Yi. 1995. "Can Convergence Regressions Distinguish between Exogenous and Endogenous Growth Models?" *Econ. Letters*, 49:2, pp. 211–15.

Koop, Gary; Jacek Osiewalski, and Mark F. J. Steel. 1995. "Measuring the Sources of Output Growth in a Panel of Countries," CORE discussion paper no. 9542.

Kormendi, Roger C. and Philip G. Meguire. 1985. "Macroeconomic Determinants of Growth: Cross-Country Evidence," *J. Monet. Econ.*, 16:2, pp. 141–63.

Kravis, Irving B. 1984. "Comparative Studies of National Incomes and Prices," *J. Econ. Lit.*, 22:1, pp. 1–39.

Kuznets, Simon. 1966. *Modern Economic Growth*. New Haven: Yale U. Press.

Leamer, Edward E. 1983. "Let's Take the Con Out of Econometrics," *Amer. Econ. Rev.*, 73:1, pp. 31–43.

———. 1985. "Sensitivity Analyses Would Help," *Amer. Econ. Rev.*, 75:3, pp. 308–13.

Lee, Jong-Wha and Robert J. Barro. 1997. "Schooling Quality in a Cross Section of Countries." NBER working paper no. 6198.

Lee, Kevin; M. Hashem Pesaran, and Ron Smith. "Growth and Convergence in a Multi-Country Empirical Stochastic Solow Model," *J. Appl. Econometrics*, 12, pp. 357–92.

Leung, Charles and Quah, Danny T. "Convergence, endogenous growth, and productivity disturbances," *J. Monetary Econ.*, 1996, 38:3, 535–547.

Levine, Ross. "Financial Development and Economic Growth: Views and Agenda," *J. Econ. Lit.*, 35, pp. 688–726.

Levine, Ross and David Renelt. 1991. "Cross-Country Studies of Growth and Policy," World Bank working paper WPS no. 608.

———. 1992. "A Sensitivity Analysis of Cross-Country Growth Regressions," *Amer. Econ. Rev.*, 82:4, pp. 942–63.

Levine, Ross and Sara J. Zervos. Forthcoming. "Stock Markets, Banks, and Economic Growth," *Amer. Econ. Rev.*

Lucas, Robert E. "On the Mechanics of Economic Development," *J. Monet. Econ.*, 22, pp. 3–42.

Mankiw, N. Gregory. 1995. "The Growth of Nations," *Brookings Pap. Econ. Act.*, 1, pp. 275–310.

Mankiw, N. Gregory; David Romer, and David N. Weil. 1992. "A Contribution to the Empirics of Economic Growth," *Quart.J. Econ.*, 107:2, 407–37.

Manksi, Charles F. "Identification of Endogenous Social Effects: The Reflection Problem," *Rev. Econ. Stud.*, 60, pp. 531–42.

Marris, Robin. 1984. "Comparing the Incomes of Nations: A Critique of the International Comparison Project," *J. Econ. Lit.*, 22:1, pp. 40–57.

Mauro, Paolo. 1995. "Corruption and Growth," *Quart. J. Econ.*, 110:3, pp. 681–712.

Moreno, Ramon and Bharat Trehan. 1997. "Location and the Growth of Nations," *J. Econ. Growth*, 2:4, pp. 399–418.

Nuxoll, Daniel A. 1994. "Differences in Relative Prices and International Differences in Growth Rates," *Amer. Econ. Rev.*, 84:5, pp. 1423–36.

Parente, Stephen L. and Edward C. Prescott. 1994. "Barriers to Technology Adoption and Development," *J. Polit. Econ.*, 102:2, pp. 298–321.

Perotti, Roberto. 1996. "Growth, Income Distribution, and Democracy: What the Data Say," *J. Econ. Growth*, 1, pp. 149–87.

Persson, Torsten and Guido Tabellini. 1994. "Is Inequality Harmful for Growth?" *Amer. Econ. Rev.*, 84:3, pp. 600–21.

Pesaran, M. Hashem and Ron Smith. 1995. "Estimating Long-Run Relationships from Dynamic Heterogeneous Panels," *J.Econometrics*, 68:1, pp. 79–113.

Pritchett, Lant. 1996a. "Population, Factor Accumulation, and Productivity," World Bank working paper no. 1567.

———. 1996b. "Where Has All the Education Gone?" World Bank working paper no. 1581.

———. 1997. "Divergence, Big Time," *J. Econ. Perspectives*, 11:3, pp. 3–17.

———. 1998. "Patterns of Economic Growth: Hills, Plateaus, Mountains, and Plains," manuscript, World Bank.

Quah, Danny T. 1993. "Empirical Cross-Section Dynamics in Economic Growth," *Europ. Econ. Rev.*, 37:2–3, pp. 426–34.

———. 1997. "Empirics for Growth and Distribution: Stratification, Polarization, and Convergence Clubs," *J. Econ. Growth*, 2:1, pp. 27–59.

Ramey, Garey and Valerie A. Ramey. 1995. "Cross-Country Evidence on the Link between Volatility and Growth," *Amer. Econ. Rev.*, 85:5, pp. 1138–51.

Ranis, Gustav. 1984. "Typology in Development Theory: Retrospective and Prospects," in *Economic Structure and Performance*. Moshe Syrquin, Lance Taylor and Larry E. Westphal, eds. London: Academic Press.

Ranis, Gustav and John C. H. Fei. 1988. "Development Economics: What Next?" in *The State of Development Economics*. Gustav Ranis and T. Paul Schultz, eds. Oxford: Basil Blackwell.

Robinson, Sherman. 1971. "Sources of Growth in Less Developed Countries: A Cross-Section Study," *Quart. J. Econ.*, 85:3, pp. 391–408.

Rodrik, Dani. 1998. "Where Did All the Growth Go? External Shocks, Social Conflict and Growth Collapses," CEPR working paper no.1789.

Rogoff, Kenneth. 1996. "The Purchasing Power Parity Puzzle," *J. Econ. Lit.*, 34, pp. 647–68.

Romer, David. 1996. *Advanced Macroeconomics*. Boston: McGraw-Hill.

Romer, Paul. 1986. "Increasing Returns and Long-Run Growth," *J. Polit. Econ.*, 94:5, pp. 1002–37.

Rousseeuw, Peter J. and Annick M. Leroy. 1987. *Robust Regression and Outlier Detection*. New York: Wiley.

Sachs, Jeffrey D. and Andrew M. Warner. 1995a. "Economic Reform and the Process of Global Integration," *Brookings Pap. Econ. Act.*, pp. 1–118.

———. 1995b. "Natural Resource Abundance and Economic Growth," NBER working paper no. 5398.

———. 1997. "Fundamental Sources of Long-Run Growth," *Amer. Econ. Rev.*, 87:2, pp. 184–88.

Sala-I-Martin, Xavier. 1996. "The Classical Approach to Convergence Analysis," *Econ. J.*, 106, pp. 1019–36.

Sarel, Michael. "Demographic Dynamics and the Empirics of Economic Growth," *IMF Staff Papers*, 1995, 42:2, 398–410.

Scholing, Eberhard and Vincenz Timmermann. "Why Growth Rates Differ: Measuring

'Unmeasurable' Influences," *World Devel.*, 16:11, pp. 1271–94.

Sims, Christopher A. "Commentary," *Fed. Res. Bank St. Louis Rev.*, May 78:3, pp. 173–78.

Slemrod, Joel. 1995. "What Do Cross-Country Studies Teach about Government Involvement, Prosperity, and Economic Growth?" *Brookings Pap. Econ. Act.*, pp. 373–431.

Summers, Robert and Alan Heston. 1988. "A New Set of International Comparisons of Real Product and Price Levels Estimates for 130 Countries, 1950–1985," *Rev. Income Wealth*, 34:1, pp. 1–25.

———. 1991. "The Penn World Table: Mark 5: An Expanded Set of International Comparisons, 1950–1988," *Quart. J. Econ.*, 106:2, pp. 327–68.

Temple, Jonathan R. W. 1998a. "Robustness Tests of the Augmented Solow Model," *J. Applied Econometrics*, 13:4, pp. 361–75.

———. 1998b. "Equipment Investment and the Solow Model," *Oxford Econ. Papers*, 50:1, pp. 39–62.

Temple, Jonathan R. W. and Paul A. Johnson. 1998. "Social Capability and Economic Growth," *Quart. J. Econ.*, 113:3, pp. 965–90.

Temple, Jonathan R. W. and Hans-Joachim Voth. 1998. "Human Capital, Equipment Investment, and Industrialization," *Europ. Econ. Rev.*, 42:7, pp. 1343–62.

Ventura, Jaume. 1997. "Growth and Interdependence," *Quart. J. Econ.*, 112, pp. 57–84.

Wolf, Holger C. 1994. "Growth Convergence Reconsidered," *Weltwirtschaftliches Archiv*, 130:4, pp. 747–59.

Young, Alwyn. 1992. "A Tale of Two Cities: Factor Accumulation and Technical Change in Hong Kong and Singapore," in *NBER Macroeconomics Annual*. Olivier Jean Blanchard and Stanley Fischer, eds. London and Cambridge: MIT Press, pp. 13–54.

———. 1995. "The Tyranny of Numbers: Confronting the Statistical Realities of the East Asian Growth Experience," *Quart. J. Econ.*, 110:3, pp. 641–80.