

WHAT HAVE WE LEARNT FROM THE CONVERGENCE DEBATE?

Nazrul Islam

Emory University

Abstract. This paper surveys the convergence literature. It begins by laying out different definitions of convergence and by showing the link between the convergence issue and the growth theory debate. The paper then follows the convergence research conducted along four different approaches, namely the cross-section, panel, time-series, and distribution approaches. The paper shows the association of these methodological approaches with various definitions of convergence and highlights the connections among the convergence results. It shows that, despite some impressions to the contrary, there is considerable agreement among the results. Although the convergence research might not have solved the growth debate entirely, it has helped both the neoclassical and the new growth theories to adapt and evolve. The research on convergence has established new stylized facts regarding cross-country growth regularities. It has brought to fore the existence of large technological and institutional differences across countries and has given rise to new methodologies for quantifying and analyzing these differences. This is providing a new information base for analysis of technological and institutional diffusion and for further development of growth theory in general.

Keywords. Economic Growth; Convergence; Productivity

1. Introduction

A central issue around which the recent growth literature has evolved is that of convergence. Whether income levels of poorer countries of the world are converging to those of richer countries is by itself a question of paramount importance for human welfare. However, interest in this question has been fueled further by the fact that it became linked with the issue of validity of alternative growth theories. It has been generally thought that convergence was an implication of the neo-classical growth theory (NCGT), while the new growth theories (NGT) did not have this implication. Accordingly, it was believed that by testing for convergence, one could test for the validity of alternative growth theories. Given this connection, it is not surprising that the convergence issue has drawn the attention of many outstanding minds of the economics profession.

This attention has however led to many different interpretations of convergence and to a wide array of empirical results, so much so that a feeling of exasperation

is now not uncommon. Some have even expressed the view that the new growth literature in general and the convergence literature in particular have not produced anything new or substantive. In their surveys of the new growth literature, Durlauf and Quah (1999) and Temple (1999) try to refute this pessimistic appraisal. Both these surveys are comprehensive, and convergence is only one of the different growth topics covered. These papers therefore cannot give exclusive attention to the convergence debate. For example, Temple's survey does not include the discussion of the 'distribution approach' to convergence study. Durlauf and Quah pay considerable attention to the theoretical relationship among different growth models and relegate many details of the empirical results to a summary table. The otherwise excellent surveys by Sala-i-Martin (1996b) and De la Fuente (1997) leave out convergence studies that use the panel, time-series, and distribution approaches. Meanwhile, research on convergence continues, and new entrants to this research often do not display adequate awareness of the different ways in which convergence has already been investigated and of the connections that exist among the available results. Presentation of new convergence results in this backdrop often adds to the confusion.

The aim of this paper is to provide a comprehensive background in the context of which future growth and convergence research can be pursued and understood. In achieving this goal, the paper builds on the previous surveys, extends and updates them, and offers some different assessments of the contribution of the convergence literature to the understanding of economic growth. An important distinction of the present survey lies in its mode of exposition. The literature on convergence has unfolded in response to perceived logical inadequacies of previous works. Understanding of the analytical points therefore requires a perusal of how this literature has evolved over time. Accordingly, the present survey uses the *historical-logical* mode of exposition and often looks at the literature from the history of thought point of view.¹ Also, the survey is presented in a fairly non-technical fashion, so that those who are not engaged in research on growth and convergence can find it easy to follow.

The paper begins by identifying the definitions of convergence and the methodologies that have been used for their investigation. It shows that while there is some correspondence between the definitions and methodologies, this correspondence is not unique. The paper identifies the areas of overlaps and shows how the awareness about these overlaps helps in understanding the relationship among results obtained from different methodologies.

In evaluating the convergence results, the paper reexamines Barro's (1997, p. x) conclusion that the recent empirical research prompted by the advent of new growth theory has ironically helped vindicate the explanatory power of the neoclassical growth theory.² Temple (1999, p. 112) finds this conclusion to be only a partial characterization of the new growth evidence. Durlauf and Quah (1999) think that this conclusion is erroneous because it is based on research that uses only linear specifications of the growth-convergence equation and ignores the possibility of its non-linear specification.

Barro's conclusion derives from the general empirical finding supporting 'conditional convergence.' The present survey shows that some of the findings of the 'time series' and the 'distribution' approaches can also be interpreted as supportive of 'conditional convergence.' The paper notes the empirical difficulty in distinguishing 'club convergence' from 'conditional convergence.' However, in reflecting on the results, the survey emphasizes the metamorphosis that both the NCGT and the NGT have undergone under the impact of the convergence research. The paper shows that, as a consequence of the give and take between the NCGT and NGT, it is now possible, generally speaking, to explain both convergence and non-convergence behavior by appropriately chosen models of growth theory of both these varieties. This in some sense frustrates the original idea of using convergence as a criterion for validity of alternative growth theories. However, the fact that convergence research led both the NCGT and the NGT towards accommodation may not be a small achievement by itself.

The convergence research had other, wider ramifications. First, it has produced new stylized facts about cross-country growth regularities, such as 'persistence' and 'bi-modality.' The growth theory is now called upon to explain these facts. Second, the convergence research has brought to fore the importance of technological differences across countries and has led to the development of new methodologies for quantification of these differences. The results of these quantification efforts are providing a new information base for examining alternative models of technology generation and diffusion. This information base is also helpful for understanding the interaction among countries along other dimensions such as trade, migration, spread of institutions, etc. Finally, the convergence research has also provided the background for formulation of stochastic growth models. All in all, instead of proving futile, the convergence research has had impressive achievements and has opened up useful new lines of research.

The discussion of this paper is organized as follows. Section-2 discusses the link between the growth theory controversy and the issue of convergence, and it catalogues different definitions and methodological approaches used for the convergence research. Section-3 provides a brief description of different concepts of convergence. Section-4 reviews the initial evidence on convergence based on informal specifications of cross-section regressions. The formal, model-based growth-convergence equation that has become the mainstay of convergence research is presented in section-5. Section-6 reviews the cross-section results based on formal specifications. It also includes discussion of 'club convergence.' Section-7 reviews the panel approach to convergence study, including research on 'TFP-convergence.' Section-8 reviews the time series approach to convergence analysis. Section-9 discusses the distribution approach to convergence, including research on σ -convergence. Conclusions are drawn in section-10. The literature on convergence is too vast to make an all-inclusive survey possible. Despite efforts to be inclusive, some works are only briefly discussed here and others remain outside the purview. This however does not mean that these works are not important.

2. Growth Theory and the Issue of Convergence

The NCGT assumption of diminishing returns leads to the convergence implication.³ In the mid-eighties however a perception arose that convergence did not hold in large samples of countries. Romer (1994) identifies this perceived incongruity as one of the two origins of the NGT. The other origin, according to Romer, is NCGT's inability to generate long term growth from within the model. The NGT models try to solve these twin problems by avoiding diminishing returns in various ways. This lets these models avoid the convergence implication and to have endogenous long run growth. Thus we have the link between the convergence issue and the issue of validity of alternative growth theories. It is because of this link and broader significance that the convergence debate has been raging so forcefully for such a long time.

The debate has in turn led to many different interpretations of convergence. In applying NCGT to the study of cross-country growth regularities, researchers have, either explicitly or implicitly, added other assumptions (to the basic assumption of diminishing returns), and this has been the main reason for the emergence of different notions of convergence. The following, often encountered, dichotomies indicate some of the different ways in which convergence has been understood:

- (a) Convergence *within* an economy vs. convergence *across* economies;
- (b) Convergence in terms of *growth rate* vs. convergence in terms of income *level*;
- (c) β -convergence vs. σ -convergence;
- (d) *Unconditional* (absolute) convergence vs. *conditional* convergence;
- (e) *Global* convergence vs. *local* or *club*-convergence;
- (f) *Income*-convergence vs. *TFP* (total factor productivity)-convergence; and
- (g) *Deterministic* convergence vs. *stochastic* convergence.

It is not that all these different concepts of convergence were apparent from the very beginning. Research on convergence proceeded through several stages, and it is only with time that these different definitions emerged and gained currency. Convergence research has also witnessed the use of different methodologies, which may be classified broadly as follows:

- (a) Informal cross-section approach,
- (b) Formal cross-section approach,
- (c) Panel approach,
- (d) Time-series approach, and
- (e) Distribution approach.

There is some correspondence between the convergence definitions and the methodologies used. This correspondence is however not unique. For example, the informal and formal cross-section approaches, the panel approach, and the time-series approach (in part) have all studied β -convergence, either conditional or unconditional. These approaches have generally dealt with convergence *across*

economies and in terms of per capita income *level*. In addition, the formal cross-section approach and the panel approach have been used to study club-convergence and TFP-convergence. The cross-section approach has even been used to study σ -convergence. The time series approach has been used to investigate convergence both *within* an economy and *across*-economies. Finally, the distribution approach has gone beyond investigating just σ -convergence and has studied the entire shape of the distribution and intra-distribution dynamics. A useful way to start reviewing the convergence literature is therefore to provide a brief introduction to these different concepts of convergence.

3. Different Concepts of Convergence

3.1. *Convergence Within vs. Convergence Across*

Robert Solow (1970), in his exposition of growth theory, starts out by relating to six stylized facts about growth put forwarded by Kaldor (1971). Coming to the fifth and sixth of these,⁴ Solow pauses and makes the following comment:

“The remaining ‘stylized facts’ are of a different kind, and *will concern me less*, because they relate more to comparisons *between* different economies than to the course of events *within* any one economy.” (p. 3; my italics.)

It is somewhat ironic that one of the recent dissatisfactions with the Solow model concerns its alleged failure to explain *between-* or *across*-country variation in growth rate and income level. Historically, the main objective of the Solow model has been to show that once factor substitution is allowed, the economy could achieve *stable* dynamic equilibrium, instead of suffering from the inherent instability that characterized the previous Harrod-Domar growth model. In NCGT, no matter whether the economy starts off from a per capita capital stock that is lower or higher than the equilibrium capital level, the substitution possibility and diminishing returns force the economy to ‘converge’ to the equilibrium. Hence, this is a proposition of convergence, albeit *within* the economy. Paradoxically, the concept of convergence that arose and became associated with NCGT refers to an *across*-economy process.

3.2. *Convergence in Terms of Growth Rate vs. Convergence in Terms of Income Level*

The across-economy convergence may in turn be understood in two different ways, namely ‘convergence in terms of growth rate’ and ‘convergence in terms of income level.’ Both of these require extending the NCGT conception of technology to the world level. The specification of technological progress in NCGT is based on the following assumptions: (a) no resources are needed to generate technological innovation, (b) everybody benefits equally from it, and (c) nobody pays any compensation for benefiting from it. Extended to a global setting, these

assumptions imply that all countries share in the technological progress equally, and hence they all can grow at the same rate in the steady state. This yields the hypothesis of convergence in terms of *growth rate*. To this researchers often added the assumption that all countries have identical aggregate production function. This implies that steady state income *levels* of all countries are also identical. This yields convergence in terms of income level.

3.3. β -convergence vs. σ -convergence

Convergence in terms of both growth rate and income level requires what is called β -convergence. This follows from the assumption of diminishing returns, which imply higher marginal productivity of capital in a capital-poor country. With similar savings rates, poorer economies will therefore grow *faster*. If this scenario holds, there should be a *negative* correlation between the initial income level and the subsequent growth rate. This led to the popular methodology of investigating convergence, namely running what is now known as the *growth-initial level* regressions. The coefficient of the initial income variable in these regressions (say, β) is supposed to pick up the negative correlation. Convergence judged by the sign of β is known as the β -convergence.⁵

However, such researchers as Quah (1993a), Friedman (1994), and others have emphasized that convergence is a proposition regarding *dispersion* of the cross-sectional distribution of income (and growth rate), and a negative β from the growth-initial level regression does not necessarily imply a reduction in this dispersion.⁶ According to this view, instead of judging indirectly and perhaps erroneously through the sign of β , convergence should be judged directly by looking at the dynamics of dispersion of income level and/or growth rate across countries. This gave rise to the concept of σ -convergence, where σ is the notation for standard deviation of the cross-sectional distribution of either income level or growth rate.

Despite the limitations above, researchers have continued to be interested in β -convergence, in part because it is a necessary, though not sufficient, condition of σ -convergence. The other reason is that methodologies associated with investigation of β -convergence also provide information regarding structural parameters of growth models, while research along the distribution approach usually do not provide such information, as we shall see.

3.4. Unconditional Convergence vs. Conditional Convergence

From a conceptual point of view, the most important distinction is probably between conditional and unconditional convergence. Proceeding from the Solow model and assuming a Cobb-Douglas production function of the type

$$Y_t = K_t^\alpha (A_t L_t)^{1-\alpha} \quad (1)$$

(where Y , K , L , and A stand for output, capital, labor, and total factor productivity, respectively), the steady state level of per capita income, y^* , is given by

$$y^* = A_0 e^{gt} [s/(n + g + \delta)]^{\alpha/(1-\alpha)}, \quad (2)$$

where s is the investment rate, g and n are the assumed exponential growth rates of A_t and L_t , respectively.⁷ This shows clearly that the steady state income level of a country depends on the following six elements: A_0 , s , g , n , δ , and α , which may be combined in the vector θ .⁸ Unconditional convergence implies that all elements of θ are the same for the economies considered. In terms of the growth-initial level regression, this means that the sign of β should be negative even if no other variable is included on the right hand side. In contrast, the concept of conditional convergence emphasizes possible differences in the steady state and hence requires that appropriate variables be included on the right hand side of the growth-initial level regression in order to control for these differences. Which of the different elements of the vector θ should be allowed to vary and which not, continues to be an important issue, as we shall see.

3.5. Conditional Convergence vs. Club Convergence

The concept of conditional convergence is also related with the notion of 'club convergence.' The latter term can be traced back to Baumol (1986), but its more rigorous formulation owes to Durlauf and Johnson (1995) and Galor (1996). One property of the standard NCGT is *uniqueness* of its equilibrium, and the usual notion of convergence assumes this uniqueness. In the case of unconditional convergence, there is only one equilibrium-level to which *all* economies approach. In the case of conditional convergence, equilibrium differs by the economy, and each particular economy approaches its own but *unique* equilibrium. In contrast, the idea of club-convergence is based on models that yield *multiple* equilibrium.⁹ Which of these different equilibrium an economy will reach, depends on its initial position or some other attribute. A group of countries may approach a particular equilibrium if they share the initial location or attribute corresponding to that equilibrium. This produces club-convergence.

3.6. Income-convergence vs. TFP-convergence

Researchers have generally dealt with convergence in terms of per capita income, i.e., with *income convergence*. However, income convergence can be the joint outcome of the twin processes of capital deepening and technological catch-up. While most researchers have focused on parameters of the capital deepening process, other researchers, such as Dowrick and Nguyen (1989), Dougherty and Jorgenson (1996, 1997), Wolff (1991), and Dollar and Wolff (1994), have directed their attention to the process of technological catch-up. Since total factor productivity (TFP) is the closest measure of technology, these researchers have investigated whether countries have come closer in terms of TFP levels. This has given rise to the concept of *TFP-convergence*. Clearly, income convergence can get either accelerated or thwarted depending on whether initial TFP-differences narrow or widen over time.

3.7. *Deterministic Convergence vs. Stochastic Convergence*

Several researchers, such as Bernard and Durlauf (1996), Carlino and Mills (1993), Evans (1996), and Evans and Karras (1996a), Qi and Papell (1999), and others have investigated convergence using time series econometric methods. As we shall see in more detail below, ‘within convergence’ is actually a time series concept. However, researchers have used time series analysis to examine ‘across convergence’ too. From this point of view, Two economies, i and j , are said to converge if their per capita outputs, $y_{i,t}$ and $y_{j,t}$ satisfy the following condition:

$$\lim_{k \rightarrow \infty} E(y_{i,t+k} - a \cdot y_{j,t+k} | I_t) = 0, \quad (3)$$

where I_t denotes the information set at time t . This definition of convergence is relatively unambiguous for a two-economy situation. This is not so when convergence is considered in a sample of more than two economies. Researchers differ on defining convergence in such multi-country situations. Some have taken deviations from a *reference economy* as the measure of convergence. In this treatment, y_{it} in equation (3) is replaced by y_{1t} , where 1 is the index for the reference country. Others have based their analysis of convergence on deviations from the *sample average*. In this treatment, y_{it} is replaced by \bar{y}_t , the average for time t . This difference is not innocuous, as we shall see. The time series definitions of convergence can be related with the notions of conditional and unconditional convergence too. With $a = 1$, equation (3) represents a variant of *unconditional* convergence. On the other hand, if $a \neq 1$ then equation (3) may represent a variant of *conditional* convergence. Within this framework a distinction has also been made between ‘deterministic’ and ‘stochastic convergence.’ This distinction refers to whether ‘deterministic’ or ‘stochastic’ trend is allowed in testing for unit root in the deviation series.

From a chronological point of view, the study of convergence began with the notion of ‘absolute convergence’ and then moved to the concept of ‘conditional convergence.’ Both these concepts were initially studied using the notion of ‘ β -convergence.’ The notion of σ -convergence arose later. Alongside emerged the concepts of ‘club-convergence,’ ‘TFP-convergence,’ and the time series notions of convergence. There was also a chronological progression from the ‘informal cross-section’ to ‘formal cross-section,’ and then on to ‘panel’ approach to convergence study. The ‘time-series’ and the ‘distribution’ approaches developed alongside.

With this introduction to various convergence concepts, we can now proceed to follow the complex evolution of the convergence research.

4. Initial Cross-section Studies of Convergence

The specifications of ‘growth-initial level regressions’ used in initial studies of β -convergence were not *formally* derived from theoretical models of growth. This however does not mean that these studies did not have connection with growth

models. In fact, to the extent that this connection was less formal, some of these works could derive inspiration from *several* theoretical paradigms and therefore have multiple focus. Although the concepts of unconditional and conditional convergence were not rigorously distinguished yet, the evidence of these initial works can be attributed, with hindsight, to both these notions.

4.1. *Initial Evidence of Unconditional Convergence*

The most well known initial study of unconditional convergence is by Baumol (1986). The main part of his analysis is based on a sample of 16 OECD countries for which long term data were available from Maddison (1982). Baumol obtains a significant negative coefficient on the initial income variable in a growth-initial level regression for these countries, and takes this as strong evidence of (unconditional) convergence.¹⁰ However, prodded by Romer, Baumol also considers the relationship in an extended sample of 72 countries. In this larger sample, however, he does not find evidence of convergence.¹¹ Thus, Baumol's study produces evidence of both presence and absence of unconditional convergence, depending on the sample. Introspecting on the basis of the growth-initial level scatter-diagram, Baumol also coins the expression 'convergence-club.' He suggests that, while there is no convergence in the larger sample as a whole, there exist 'clubs' of countries within which evidence of convergence can be seen.¹² DeLong (1988) however shows that Baumol's finding of unconditional convergence in the 16-country OECD sample suffers from selection bias.¹³ Nevertheless, Baumol's finding of absence of unconditional convergence in the larger sample of countries became an important point of departure for further discussion of convergence.

4.2. *Initial Evidence of Conditional Convergence*

While Baumol's study focuses on unconditional convergence, other studies, such as by Kormendi and Meguire (1985) and Grier and Tullock (1989), provide evidence that can be interpreted as being of conditional convergence. Reflecting research interests of an earlier period, these studies also consider other issues such as of inflation-output trade off, Philips curve relationship, etc. Regressions in these studies therefore include additional variables representing these relationships. However, the basic neoclassical paradigm is preserved through the inclusion of labor, capital, and the *initial* income variable. In a sample of about fifty countries, Kormendi and Meguire's regressions yield a negative β , which can be taken as evidence of conditional convergence. Grier and Tullock (1989) extend Kormendi and Meguire's study to a larger sample size and longer sample period. This allows them to consider the issue of parameter stability across sub-samples and sub-periods. The evidence from their study is somewhat mixed.¹⁴

4.3. *Initial Evidence on Convergence and the Growth Theory controversy*

Some of the initial studies of convergence, such as Baumol (1986) and Kormendi and Meguire (1985), either pre-date or are contemporaneous with pioneering

NGT papers, and hence the growth theory controversy was not topical for them. The studies that are subsequent to the advent of NGT do make efforts to relate their results to the growth theory debate. However, to the extent that regression specifications of these studies are not formally linked with growth models, these efforts have to be limited and conjectural.¹⁵ On the other hand, the absence of a formal link with growth models saves these initial studies from the *within-across tension* of the convergence concept that will soon surface.

4.4. *Convergence and Human Capital*

The convergence studies discussed above do not include human capital as an explanatory variable. Yet, as we noticed, in striving to relax the constraint of diminishing returns, many NGT models rely on human capital related processes. Hence human capital was set to become an important variable in convergence research. This begins with Barro (1991),¹⁶ which is inspired directly by the NCGT-NGT controversy and is also one of the few studies that look at the convergence issue from the NGT perspective.¹⁷

Barro starts by abandoning the standard neoclassical format and, instead, emphasizes the simultaneity among growth, investment, and fertility.¹⁸ In view of this simultaneity, he runs separate sets of regressions with growth, investment, and fertility as dependent variables.¹⁹ His basic regressions of growth do not even include physical capital and labor as explanatory variables. Instead, the focus is now on *human capital*, which appears in all the regressions and proves important. Barro interprets the finding as vindication of the NGT-emphasis on human capital.²⁰

To study convergence, Barro includes the initial income variable in his regressions. He reports absence of unconditional convergence in a broad sample of 98 countries,²¹ and interprets this as supportive of the NGT.²² However, he finds that when the initial measures of human capital are included, β turns negative and significant. This leads Barro to conclude that the data support the convergence hypothesis in a "*modified sense*."²³ To the extent that Barro's growth regressions include other control variables, his 'convergence in a modified sense' can be viewed as the germination of the concept of 'conditional convergence.' However, since investment and labour force growth rates do not appear as control variables, which are the main conditioning variables from the NCGT-point of view, there may be some ambiguity regarding this interpretation. Towards the end of the paper however Barro presents growth-initial level regression of more conventional (meaning, neoclassical) format, i.e., inclusive of investment and population growth rates as controls,²⁴ and finds that $\hat{\beta}$ obtained from this regression is also negative.²⁵

While Barro emphasizes the role of human capital, De Long and Summers in a series of papers draw attention to the special role of equipment investment in growth. This discussion of the role of equipment investment, thought not focused on the income-convergence issue, has important bearing on the issue of TFP-convergence, as we shall see later.²⁶

Meanwhile, soon after Barro (1991), the convergence research enters its ‘formal specification’ stage. This begins with the formulation of the neoclassical growth convergence equation, which we present in the next section.

5. The Neoclassical Equation for Convergence Study

The evolution of the ‘conditional convergence’ concept from its initial, inchoate stage to the subsequent, precise stage and the associated transition of the convergence regression from informal specifications to formal, model-based specifications are accomplished in Barro and Sala-i-Martin (1992), (henceforth BS) and Mankiw, Romer, and Weil (1992) (henceforth MRW). In both these works, the regression specification is derived formally from the neoclassical growth model. MRW work with the original Solow-Swan model, while BS use the Cass-Koopmans’ optimal savings version of the NCGT. Since these works, the neoclassical growth-convergence equation has occupied the center stage of convergence research, and it is virtually impossible to review the literature without bringing this equation into the picture. Since its detailed derivation is available elsewhere,²⁷ we sketch in the following only the basic steps, primarily to introduce the notations. The exercise involves derivation of the law of motion around the steady state first and then translation of this motion into an estimable regression equation.

5.1. Deriving the Rate of Convergence

The dynamics of capital in the Solow model are given by

$$\dot{\hat{k}} = sf(\hat{k}) - (n + g + \delta)\hat{k}, \quad (4)$$

where, $\hat{k} = \frac{K}{AL}$ is capital per *effective* labor, $\dot{\hat{k}}$ is the time derivative of \hat{k} , and $f(\hat{k})$ is the production function normalized in terms of effective labour. Also, s , n , g , δ are the rates of investment, population growth, technological progress, and depreciation, respectively. First order Taylor expansion of the right hand side term around the steady state gives

$$\dot{\hat{k}} = [sf'(\hat{k}^*) - (n + g + \delta)](\hat{k} - \hat{k}^*). \quad (5)$$

Substituting for s using the steady state relationship, $sf(\hat{k}^*) = (n + g + \delta)\hat{k}^*$, gives

$$\dot{\hat{k}} = ([f'(\hat{k}^*)\hat{k}^*/f(\hat{k}^*)] - 1)(n + g + \delta)(\hat{k} - \hat{k}^*). \quad (6)$$

Under the assumption that capital earns its marginal product, $f'(\hat{k}^*)\hat{k}^*/f(\hat{k}^*)$ equals the steady state share of capital in income, α . In the Cobb-Douglas case, this will also be the exponent of capital in the production function. Using this relationship we get

$$\dot{\hat{k}} = \lambda(\hat{k}^* - \hat{k}), \text{ where} \quad (7)$$

$$\lambda = (1 - \alpha)(n + g + \delta). \quad (8)$$

Evidently, λ gives the speed at which the gap between the steady state level of capital and its current level is closed and has come to be known in the literature as the *rate of convergence*. The same rate holds for convergence in terms of income per effective labor. This is because $\hat{y} = f'(\hat{k})$, which upon expansion at \hat{k}^* and differentiation with respect to time gives

$$\dot{\hat{y}} = f'(\hat{k}^*)\dot{\hat{k}}. \quad (9)$$

As a first order approximation, we should therefore have,

$$\hat{y}^* - \hat{y} = f'(\hat{k}^*)(\hat{k}^* - \hat{k}). \quad (10)$$

By substitution we then get,

$$\dot{\hat{y}} = \lambda(\hat{y}^* - \hat{y}), \quad (11)$$

where λ is again the rate of convergence given by equation (8).

5.2. Deriving the Equation for Testing Convergence

Switching to logarithms, solving this first order non-homogeneous differential equation, and rearranging, we get from (11):

$$\ln \hat{y}(t_2) - \ln \hat{y}(t_1) = (1 - e^{-\lambda\tau})(\ln \hat{y}^*(t_1) - \ln \hat{y}(t_1)), \quad (12)$$

where t_1 denote the initial period, t_2 the subsequent period, and $\tau = (t_2 - t_1)$. If we now substitute for \hat{y}^* from equation (2) above, we get²⁸

$$\begin{aligned} \ln \hat{y}(t_2) - \ln \hat{y}(t_1) = & (1 - e^{-\lambda\tau}) \frac{\alpha}{1 - \alpha} \ln(s_{t_1}) \\ & - (1 - e^{-\lambda\tau}) \frac{\alpha}{1 - \alpha} \ln(n_{t_1} + g + \delta) - (1 - e^{-\lambda\tau}) \ln \hat{y}(t_1) \end{aligned} \quad (13)$$

Clearly, this is again the *growth-initial level* equation, but now the coefficients are formally linked with the structural parameter of the NCGT. For example, $\beta = -(1 - e^{-\lambda\tau})$, and hence it is now possible to recover the value of λ from the estimate of β . The value of λ , in conjunction with other estimated coefficients of the equation, yields values of other structural parameters of the model such as α . The question now is how to estimate this equation.

5.3. The Within-Across Tension in the Convergence Concept

It may be noted that the above derivation in λ and of equation (13) is entirely on the basis of the growth process *within* an economy, and there is no reference to

what is happening *across* economies. This shows that λ essentially refers to a within-economy process and is determined by the values of α , n , g , and δ of the economy concerned. It would, therefore, seem natural and proper to estimate equation (13) and λ on the basis of time series data of individual economies.

Interestingly however, researchers began estimating equation (13) using cross-section data. This is because, from the beginning, convergence arose as a concept pertaining to *across-economy* growth regularities. The question that resides in the mind of most people interested in convergence is not so much whether an individual country is closing the gap between its *own* current and steady state level of income as whether poorer countries are narrowing their gap with richer countries. From the latter point of view, cross section data is the natural place to look for evidence of convergence.

However, this introduces a tension in the interpretation of the convergence parameter λ . While according to equation (13), λ is the measure of speed at which an economy proceeds towards *its* own steady state level, the λ estimated from cross-section data is often interpreted as the speed at which poorer economies are closing their income gap with richer countries. In the case of unconditional convergence, this is not a problem, because then the steady state income levels of the currently rich and poor countries coincide, making both within and across interpretation of λ valid. However, under conditional convergence, where the steady states of the currently poor and rich countries are different, the across-interpretation of λ is no longer valid.²⁹ This tension was not so apparent as long as cross-country regression specifications were informal, because in those contexts it was possible to limit to only the reduced form, cross-sectional interpretation of β .³⁰ Thus, the formal derivation of the neoclassical convergence equation, while being a significant step forward, has also brought to fore the within-across tension in the interpretation of the convergence rate.

6. The Formal Cross-section Approach to Convergence Study

6.1. Cross-section Estimation of the Neoclassical Convergence Equation

One of the most successful implementations of the formal cross section approach is MRW itself. In this influential study, two of the six elements of θ , namely s and n , are allowed to differ across countries. The values of the rest three, namely α , g , and δ , are taken to be the same for all countries. Differences in A_0 are assumed to be part of the error term, and this allows estimation of the equation by the ordinary least squares (OLS) method. MRW find that with only s and n as explanatory variables, the regression runs well, but the implied values of the rate of convergence, λ , are too low.³¹ The accompanying result is that the implied values of α are too high.³² To overcome this problem, MRW augment the Solow model by including human capital as another argument of the production function in exactly the same way as is physical capital, only with a different exponent, say ϕ . Regression on the basis of this augmented model produces more desirable results. The value of α decreases to an empirically plausible level (0.48 for the

NONOIL sample), and the value of λ increases to around .02 (0.0142 for the NONOIL sample, implying a half-life of 49 years).

Similar results on conditional convergence across countries are presented in BS. These are drawn mainly from Barro (1991). However, the regressors are now clearly interpreted as determinants of the steady state. Also, the estimated regression coefficient is now traced back to yield the structural parameter λ . For a similar sample of 98 countries as MRW's NONOIL, BS report $\hat{\lambda}$ to be 0.0184.³³

6.2. *β -Convergence across 'regions'*

The formal growth-convergence equation has been used to study convergence across regions too. In particular, whether or not convergence holds for the US states has drawn considerable attention. When it comes to regions within the same country, the assumption of identical steady states, and hence of unconditional convergence, becomes more plausible. BS make this assumption to study convergence across US states using the Cass-Koopmans version of the NCGT. This helps them avoid the difficult problem of controlling for such 'deep' behavioural parameters as inter-temporal elasticity of substitution, time discount rate, etc., for which data are difficult to get.³⁴ The authors find significant evidence of convergence,³⁵ and the estimated rate of convergence proves to be in the neighborhood of two percent per year. Holtz-Eakin (1993), on the other hand, emphasizes the possible differences in steady state even among the US states and, thereby, considers a situation of conditional convergence. He uses a human capital augmented version of the neoclassical convergence equation and implements a variant of pooled regression. Upon inclusion of variables that either represent or proxy for the determinants of steady state, Holtz-Eakin obtains higher estimates of the rate of (conditional) convergence.³⁶

Regional convergence studies have been conducted in the context of other countries and continents too. Sala-i-Martin (1996b) presents a comprehensive study of convergence across regions of Japan, Germany, the UK, France, Italy, Spain, and Canada, in addition to that of the US. He generally assumes a situation of unconditional convergence, and obtains convergence rates that are close to two percent.³⁷ As we shall see, regional convergence has been studied using the panel and time series approaches too.

6.3. *Research on Club-Convergence*

The formal cross-section equation has been used to study 'club convergence' too, though not that frequently. In their contribution to this line of research, Durlauf and Johnson (1995), allude to theoretical models of multiple equilibrium and observe that convergence in large samples (*global* convergence) does not hold (or proves weak) because countries belonging to different equilibrium (or 'regimes') are lumped together. The proper thing, according to them, is to identify country groups, whose members share the same equilibrium, and then to check whether convergence holds within these groups (*local* convergence). The authors use initial levels of income and literacy levels to group the countries and find the rates of

convergence within the groups to be higher than that in the whole sample.³⁸ Also, estimated parameter values differ significantly across the groups, particularly when these groups are determined endogenously. The authors interpret the observed heterogeneity as indicative of the presence of different regimes.³⁹

Despite the conceptual distinction, it is not easy to distinguish 'club convergence' from 'conditional convergence' empirically.⁴⁰ This finds reflection in the problems associated with the choice of criteria to be used to group the countries in testing for club convergence. Clearly, steady state determinants cannot be used for this purpose, because differences in them cause equilibrium to differ even under conditional convergence. On the other hand, use of time-varying characteristics, such as initial levels of income or literacy, is also not free of problems. Suppose it is postulated that countries around a higher cut-off 1960 income (say \$10,000) have an equilibrium that is different from the equilibrium of the countries around a lower cut-off 1960 income (say \$5,000). However, the countries of the first group also had to cross the \$5,000 cut-off at some point in the past, and hence, other things being the same, the latter equilibrium should have applied to them too. This means that the initial income cut-off by itself cannot determine the equilibrium.⁴¹ There has to be other factors. However, these other factors cannot be associated with the determinants of the steady state, because then 'conditional convergence' will subsume 'club convergence.' On the other hand, leaving these factors unspecified (through endogenized grouping) is not satisfactory, because the exercise then does not yield any policy guidance. The underlying theoretical models also have to show that it is indeed these factors that determine which of the multiple equilibrium a particular economy will approach. Finally, it needs to be noted that if multiple equilibrium indeed depends on initial income cutoffs, the relationship between subsequent growth and initial income will not be linear. Appropriate non-linear specifications will have to be used to let hypothesized 'club convergence' to surface.⁴²

These theoretical and empirical difficulties make the research on club convergence challenging. However, some 'stylized facts' regarding cross-country growth regularities brought forward by the distribution approach to convergence study indicate that further research on multiple equilibrium and 'club convergence' is worthwhile. Desdoigts (1999) presents an important recent contribution in that direction.

6.4. *Evaluating the Cross-section Approach*

The formal convergence equation of MRW and BS has been used to explore a variety of other issues even within the cross-section set up. In particular, the MRW specification has been a popular point of departure, because it does not involve deep behavioral parameters and hence is easier to implement. Thus, Chua (1992) and Ades and Chua (1997) use it to study external economies arising from regional spillovers. Benhabib and Spiegel (1994) use it to study further the role of human capital. Rodriguez-Claire and Klenow (1997) use it to examine the impact of alternative measures of human capital on the convergence results presented in

MRW. Den Haan (1995) uses it to show that the distribution of the convergence parameter depends on the source of variation (whether it is technology or capital intensity) in the initial income. Nonneman and Vanhoudt (1996) augment the MRW model further by including accumulation of technological know-how. Tzanidakis and Kirizidis (1996) formulate an alternative way of testing the MRW model. Temple and Johnson (1998) use the MRW specification to examine the role of social capital in economic growth. Temple (1998b) discusses the influence of possible measurement errors and outliers on the results of the MRW. The research along this line continues.

However, the cross-section studies in general agree regarding the broad result of conditional convergence. There is no consensus regarding the precise value of either λ or α . However, the suggested values of λ are generally low and those of capital's share in income are high. The general conclusion is that capital has to be defined to include human capital so as to justify high values of the share of capital and slow diminishing returns.

The cross-section approach to convergence study however soon encountered some important limitations. These limitations often found expression in counter-intuitive implications of some of the cross-section results. For example, Cho and Graham (1996) draw attention to the counter-intuitive implication of the MRW results that on average the poorer countries are approaching their steady state from above, and the richer countries are doing so from below. Similarly, if MRW's estimated values of λ , α , and ϕ (together with their assumed values of g and δ) are used to compute the implied value of n , the results obtained are -0.001 , 0.0064 , and $.0028$ for the NONOIL, INTER, and OECD samples, respectively. These values are far from representative values of n in these samples, and the negative value of n for the NONOIL sample is quite problematic.

Thus while the switch from informal to formal specifications elevated the convergence discussion from one about broad presence or absence of convergence to one about *precise values* of structural parameters of the growth model, it also helped reveal some important problems. The basic limitation of the cross-section approach lies in the fact that having just one data point for a country provides a weak basis for estimation of the convergence parameter λ , which refers primarily to a within-country process. There is much too heterogeneity across countries to validate the assumption that cross-country data can be treated as multiple data of the same country. It is to overcome this basic limitation, which finds its manifestation in various ways, that the convergence research gradually moved from the cross-section to the panel approach.

7. The Panel Approach to Convergence Study

Income convergence, as noted earlier, can have two sources, namely capital deepening and technological diffusion.⁴³ This is reflected by the presence of the technology parameters, A_0 and g , alongside the parameters governing the capital deepening process, in the formula of the steady state income level given by equation (2). In studying convergence it is therefore necessary to take account

of both these processes. However, the cross-section approach finds it difficult to do so, because it generally relies on the assumption of identical technologies across countries. This homogeneity assumption rules out any systematic process of technological diffusion, contradicting both casual observation and findings of the empirical research.⁴⁴ When neglected, these technological differences act as a confounding factor in the data and hinder correct estimation of the parameters governing the capital deepening process. We begin the discussion of the panel approach by first noting this estimation problem.

7.1. Omitted Variable Bias Problem of the Cross-section Regression

The problem can be illustrated using equation (13). Although this equation is in terms of income per *effective* labor, in actual implementation, researchers invariably work with income per capita. Expressed in terms of per capita income and rearranging, we get from equation (13) the following

$$\ln y_{t_2} = (1 - e^{-\lambda\tau}) \frac{\alpha}{1 - \alpha} \ln s_{t_1} - (1 - e^{\lambda\tau}) \frac{\alpha}{1 - \alpha} \ln (n_{t_1} + g + \delta) + e^{-\lambda\tau} \ln y_{t_1} + (1 - e^{-\lambda\tau}) \ln A_0 + g(t_2 - e^{-\lambda\tau} t_1). \quad (14)$$

The A_0 term on the right hand side is the productivity shift term. MRW, for example, recognize the importance of this term and observe that, “the A_0 term reflects not just technology but resource endowments, climate, institutions, and so on; it may therefore differ across countries.” (p. 410–1)⁴⁵ However, in actual estimation, they regard A_0 as part of the error term and assume it to be uncorrelated with the included variables, s and n . This assumption however contradicts the expansive definition of A_0 that MRW themselves provide. Going by that definition, it is difficult to argue that A_0 is uncorrelated with the savings and fertility behavior of a nation.

Actually, in a cross-section regression, the necessity of econometric identification forces one to relegate A_0 to the error term. This is because there are no good measures of A_0 , and, even if some proxy variables are included, there still remains a part of A_0 that is unobservable or unmeasurable and yet correlated with the included variables.⁴⁶ However treating A_0 as part of the uncorrelated error term causes the well-known Omitted Variable Bias (OVB) problem.⁴⁷

7.2. Panel Estimation of the Convergence Equation

One of the advantages of the panel approach is that it can correct the OVB problem by allowing for technological differences across countries (at least the unobservable and unmeasurable part of it) in the form of *individual (country) effects*. Using notations of the panel data literature, equation (14) can be written as

$$y_{it} = (1 + \beta)y_{i,t-1} + \beta\Psi x_{i,t-1} + \eta_t + \varepsilon_{it}, \quad (15)$$

where $y_{it} = \ln y_{it}$, $y_{i,t-1} = \ln y_{i,t-1}$, $(1 + \beta) = e^{-\lambda\tau}$, $\Psi = (-\alpha/(1 - \alpha))$, $x_{i,t-1} = (\ln s_{i,t-1} - \ln(n_{i,t-1} + g + \delta))$, $\mu_i = (1 - e^{-\lambda\tau}) \ln A_0$, and $\eta_t = g(t_2 - e^{-\lambda\tau} t_1)$.⁴⁸

In these notations, μ_i , the individual, *country effect*, represents the term A_0 . There are many different ways to model and deal with μ_i . However, in view of the correlation of A_0 with s and n , it is clear that the *random effects*-specification of μ_i is not appropriate. An appropriate choice, it seems, is the *correlated effects*-model and the accompanying Minimum Distance (MD) estimator suggested by Chamberlain (1982, 1983).⁴⁹ Both Islam (1995) and Knight *et al.* (1993) use this estimator, and the results show that allowing for technological differences have a significant impact on the estimated values of the convergence parameters. The estimated values of λ now prove to be much higher,⁵⁰ and the implied values of α also now prove to be much lower and more in conformity with its commonly accepted empirical values.⁵¹ These show that in absence of technological shifts, diminishing returns set in rather quickly; and there is not much evidence of externality associated with physical capital. The panel results also indicate that the way human capital influences output is perhaps different from the way physical capital does.⁵²

Canova and Marcet (1995) also present an analysis of the European regional data and the OECD data showing the bias of the cross-section regressions. They use a Bayesian procedure to control for differences in steady state (that includes differences in A_0) and find significantly higher values of λ and lower values of α . They also show that the hypothesis of the same steady state across regions or countries is rejected.

An important by-product of the panel approach to convergence study is the estimated values of A_0 , which provide indices of relative productivity or technology levels across countries. These levels are found to differ enormously. According to results of Islam (1995), in a sample of 96 countries, the highest value of A_0 is about forty times larger than the lowest. This issue of technological differences has recently drawn attention of many other researchers. For example, Hall and Jones (1996, 1997, and 1999) use a cross-section growth accounting methodology to compute TFP indices in a sample of 121 countries and conclude that TFP differences play the most important role in explaining income differences across countries.⁵³ Based on his analysis, Prescott (1998) declares that "savings rate differences are of minor importance. What is all important is total factor productivity (TFP)." (p. 525)⁵⁴ An important contribution of the panel approach to convergence study has been bringing to fore and provide estimates of the large differences that exist in aggregate productivity across countries. This also facilitates the analysis of the important issue of TFP convergence.

7.3. Research on TFP-Convergence

Originally, TFP studies were based on time series data of individual countries and were focused on computation of TFP *growth rates*. These studies did not consider

the issue of convergence in TFP *levels*. International comparison of relative TFP *levels* was initiated by Jorgenson and Nishimizu (1978) and was carried forward by Christensen, Cummings, and Jorgenson (1981). Dougherty and Jorgenson (1996, 1997) have recently resumed this line of research. Wolff (1991) and Dollar and Wolff (1994) have also examined TFP-level convergence using a similar methodology, which begins with a growth-accounting exercise in a multi-country setting using time series data in order to get the TFP level indices across countries. These indices are next analyzed to check for TFP-level convergence. Employing this methodology to the G-7 countries, Dougherty and Jorgenson and Wolff all find evidence of TFP-convergence.⁵⁵ The use of this methodology to study TFP convergence in large samples of countries is however difficult, because it requires lengthy and detailed time series data, which are often not available for developing countries.

Not all researchers have adopted the time-series growth accounting methodology to study TFP-convergence. Dowrick and Nguyen (1989) for example examine TFP-convergence using a cross-section regression. The specification is similar to (13), but the authors proceed from the assumption of a common capital-output ratio for all countries of the sample. In their formulation, the initial income variable is relative to that of the USA, the most advanced country of the sample. This allows them to interpret the coefficient on the initial income variable of the equation as indicative of TFP-convergence. Their results support TFP-convergence in a sample of fifteen OECD countries. The problem with this approach is that the capital-output ratio may not be the same across countries, and labor productivity differentials may arise from differences in both technological level and capital intensity. This is particularly true for larger samples of countries.⁵⁶ This limits the applicability of the cross-section regression approach for studying TFP convergence.

By themselves, the panel studies of convergence have not yet answered the question of TFP convergence. There are inherent difficulties in separating technological change from capital deepening.⁵⁷ However, the fact that convergence-rates prove higher when panel estimation allows for technological differences in the form of country-effects indicates that these differences are either not narrowing or at least not narrowing fast enough to reinforce convergence that is resulting from diminishing returns. According to results in Islam (1995), the panel estimate of the convergence rate increases 7.2 times (relative to its OLS estimate that ignores technological differences) in the NONOIL sample, while the increase is 4.5 times in the OECD sample. This agrees with the general evidence that technological differences have narrowed much less in the larger NONOIL sample than in the OECD sample. However, further information is necessary in order to make more precise conclusions about the direction and magnitude of TFP movements in convergence.⁵⁸ In particular, this requires generation of TFP-level indices for several consecutive time periods so that the TFP dynamics can be seen more directly. The use of panel methodology can be of help in this regard.⁵⁹

7.4. *Panel Analysis and the Extended Cross-section Regressions*

Many researchers working with cross-section data recognize the importance of technological differences across countries. However they try to control for these differences by including additional regressors in the equation. This leads to extended specifications of cross-section growth regressions. One problem with these extended specifications is that they quickly lead to unwieldy sets of explanatory variables. This is one of the reasons why growth regressions of the initial stage fell into some disrepute. Using Leamer's extreme bound analysis, Levine and Renelt (1992) try to rescue the situation by conducting a sensitivity study of a huge number of the right hand side variables that had appeared in growth regressions.

One benefit of the switch from the informal to formal specification of growth regressions is that it provides a clear guidance about the variables to be included in these regressions.⁶⁰ Recently some researchers have on occasions abandoned the strictly model-based specification and reverted back to the previous style of extended specifications. Some do this on purely conjectural grounds and do not try to link the additional variables of the regression with the variables of the growth model. Others try to maintain this link and argue that the additional variables stand for the A_0 term of the equation. This has again resulted in a swelling of variables in growth and convergence regressions, and Sala-i-Martin (1997), following Levine and Renelt (1992)'s spirit, has responded with another effort to determine statistically the robustness of the included explanatory variables.

Durlauf and Quah (1999) provide a critique of the extended specifications of growth-convergence equation and of Levine and Renelt (1992) and Sala-i-Martin (1997)'s attempts to resolve the issue using mechanical statistical criterion. They rightly argue that the guidance about which variable to include in the equation has to come from theory. It is true that the theory does not provide specification of A_0 . But this is all the more a reason to make an essential distinction between variables that are strictly model-determined, such as capital and labor, and variables about which the model's guidance is less clear.⁶¹ It may be noted that even those researchers who justify the extended specification by linking the additional variables to the term A_0 often do not adhere to the restrictions that follow from such a link. For example, equation (15) suggests some clear relationships between the coefficient of the term A_0 and the coefficients of the terms representing s and n . Also, the coefficients of variables proxying for A_0 will have precise relationship among themselves. Second, many of these researchers frequently leave their results in the reduced form and do not trace out the implied values of the parameters of the growth model. Even if they do, they limit themselves to using the coefficient of the initial income variable to recover the rate of convergence only and ignore the task of recovering other parameters of the growth model. This shows that the claim in these studies about the link of the included additional variables with A_0 is tenuous. Finally, to the extent that inclusion of proxy variables still leaves a part of A_0 uncaptured, the cross-section regressions cannot be entirely free of the omitted variable bias.

It is in view of these problems of the cross-section growth regressions that the panel methodology for studying technological differences holds some appeal. This methodology suggests a two-stage procedure, the first of which consists of panel estimation based on only the strictly model based variables to produce estimates of A_0 . In the second stage, the estimated values of A_0 are analyzed to find out the determinants of technology differences and diffusion. Since theoretical guidance regarding specification of A_0 is lacking, the specifications at the second stage of the analysis may again be conjectural. However, this two-stage procedure at least preserves the primal status of the basic, strictly model-based variables such as labour and capital, and does not relegate these variables to the same status as of the conjectural variables. Canova and Marcet (1995) also propose a similar two step procedure, the first of which is devoted to estimation of the steady states. "Once steady state estimates are obtained for each unit," the authors suggest, "we can test, in a second step, what variables determine the cross-sectional distribution of steady states." (p. iv)

7.5. *Heterogeneity in Steady State Growth Rate*

In the panel studies discussed above, technological differences were allowed only in the form of variation in A_0 , implying variation in technology *level*. The growth rate of technology, g , was however assumed to be common. This homogeneity assumption regarding g may be claimed to conform better to the NCGT view of technology as a public good. However, as noted earlier, the NCGT was formulated as a description of growth process *within* an economy and therefore did not automatically suggest that the public good assumption regarding technology held equally across countries. It is therefore quite possible to argue that homogeneity in g needs to be relaxed, so that even under the NCGT paradigm, countries can differ in terms of not only level but also growth rate of technology.⁶²

Lee *et al.* (1997) do actually allow g to differ in their panel study of growth and convergence, and find that this leads to large increase in the estimated rate of convergence.⁶³ The authors do not report the corresponding values of α , the capital share. However, when worked out, these are likely to be very low, and this may indicate some problem. Binder and Pesaran (1999) also allow g to differ across countries. They compute the 'deterministic rate of technological progress' and find it to vary considerably across countries.⁶⁴ Empirical testing of heterogeneity of g is made difficult by the fact that data only give the *actual* growth rates, which are generally a combination of *steady state* and *transitional* growth rates. However, given the empirical evidence of very large differences in TFP level, the idea of g differing across countries deserves further attention.

However, it is important to note that the heterogeneity of g (in addition to A_0) leads to a virtual collapse of the convergence concept, so far as its across-dimension is concerned. This concept, which with hindsight can now be termed as the Usual (concept) of Conditional Convergence (UCC), arose to allow

differences in steady state *level* of per capita income. Convergence under heterogeneity of both A_0 and g implies that the economies are converging not only to different *levels* of per capita income but also to different *growth rates*. This may be termed as the Weak (notion) of Conditional Convergence (WCC), although some may wonder whether WCC is worth calling convergence at all. Also, heterogeneity in g makes NCGT almost indistinguishable from the NGT, so far as the cross-country steady state growth pattern is concerned. Unless data pertaining to the transitional phase are available separately, it is difficult to distinguish a situation of WCC from a non-convergence situation described by many NGT models. It may be difficult to capture these finer distinctions using only cross-section data. Panel data, with its both cross-section and time-series dimensions, can be helpful in this regard.

7.6. *The Issue of Endogeneity Bias*

The use of panel data has allowed researchers to examine several other econometric problems of convergence regressions. One of these is the problem of endogeneity bias. Researchers often use contemporaneous values of investment rate and labor force growth rate as explanatory variables in growth-convergence equations. Since output-growth, investment rate, etc. in a particular period are likely to be jointly determined, the possibility of endogeneity bias in such regressions is quite obvious.⁶⁵ Caseli *et al.* (1996) raise this issue and try to rectify the situation by estimating a variant of equation (15) using Arellano and Bond (1991)'s GMM procedure.⁶⁶ The results lead the authors to reject the Solow-Swan model, both in its original version⁶⁷ and in its MRW augmented version.⁶⁸ The authors then abandon the strictly model-based specifications and switch to extended specifications. Based on results from the latter regressions, they suggest that the estimated value of λ is around ten percent and view that such a value is compatible with the open economy version of the Cass-Koopmans variant of the neoclassical growth theory.

The worry about potential endogeneity-bias in the convergence regressions is quite justified. However it is not clear whether use of the Arellano-Bond GMM estimator is the best strategy to rectify this bias, because Monte Carlo studies have generally found this estimator to display large small sample bias.⁶⁹ It is possible that any reduction in the endogeneity bias in Caseli *et al.* (1996)'s results is outweighed by the introduction of the small sample bias. (The next sub-section discusses this issue in more detail.) Their switch to extended specifications invites the issues discussed in the subsection above. For example, the authors do not work out the value of α that corresponds to their suggested value of λ . Based on their results for the original Solow model, an λ equaling to ten percent would imply an α equal to 0.1258, which is too low an estimate of capital's share in output, even if capital is defined as physical capital only. Hopefully other, better ways to address the potential endogeneity bias problem of the growth-convergence equation will be found soon.

7.7. *Problems of the Panel Approach*

The panel approach is however not devoid of problems, a few of which are discussed below.

7.7.1 *Possibility of small sample bias*

The problem of small sample bias already surfaced in our discussion of the endogeneity bias above. From an econometric point of view, equation (15) represents a *dynamic* panel data model, and there exist many different estimators for such models. Indeed, a whole range of panel estimators has now been used to estimate equation (15) or its variants. These include the Least Squares with Dummy Variables (LSDV), the Minimum Distance (MD) estimator of Chamberlain, the GMM estimators of Arellano and Bond, conditional maximum likelihood estimator (MLE), etc. In addition, Nerlove (1999) uses pooled OLS, Generalized Least Squares (GLS), and unconditional maximum likelihood estimator. As noted earlier, panel estimators relying on the random-effects assumption are not appropriate for estimation of the growth-convergence equation. This implies that the GLS estimator and some variants of the maximum likelihood estimators (that treat country-effects as random) may not be suitable for estimation of equation (15). These considerations narrow down somewhat the list of possible candidate estimators.⁷⁰ However the problem remains, because theoretical properties of most of these estimators are asymptotic and similar.⁷¹ Their small sample performance cannot therefore be ascertained without Monte Carlo studies. As already mentioned, some Monte Carlo evidence is now available. However, Monte Carlo studies are more useful when these are tailored to the equations and data sets actually used for estimation. Monte Carlo studies using Summers-Heston data set and focusing on the convergence equation (15) indicate that, in general, estimators that do not use further lagged values of the dependent variable as instruments perform better than those which do.⁷² This suggests that in choosing which panel estimator to use, researchers need to pay careful attention to the possibility of small sample bias. In doing so, they may benefit from the already existing Monte Carlo evidence. In addition, they may conduct Monte Carlo studies customized to their specific data sets and specifications.

7.7.2 *The Issue of short frequency*

Some authors have contended that panel estimation of the growth-convergence equation is not appropriate, because it throws away the cross-section variation in data and relies on the within variation only. This is not desirable, they argue, because the cross-section dimension of data contains long-run features that are more pertinent to growth study. The within dimension of data, according to this argument, is flawed because of its short frequency.⁷³ These concerns are genuine. However several things need to be said in clarification. First, except for LSDV, most of other panel estimators use *both* within and between variation. This is particularly true of Chamberlain's MD estimator, which is one of the more

appropriate estimators for the growth-convergence equation. Second, whether or not an estimator uses *both* within and between variation cannot be the main criterion of an estimator's suitability in this case. For example, the random-effects GLS estimator uses both within and between variation, but this estimator is not suitable for estimating the convergence equation because it contradicts the correlation of the country effects with the included explanatory variables. Third, as noted earlier, the parameter λ refers to the speed at which an economy is moving to its *own* steady state. From this point of view, it may even be more appropriate to estimate the convergence equation using *within* variation. Finally, the frequency at which within variation is considered in most of the panel studies is five years. This is shorter than twenty-five years, which is often the range over which averages are computed in cross-section studies. However, it is not as short as a year, which is often the frequency at which data are considered in many convergence studies of the time-series and distribution approaches. Barro (1997) himself offers pooled regressions with data spanning over ten years. Hence, five-year spans may not be too short to study growth, particularly when several such five-year spans are combined to produce the estimates, so that the effects of cut-off years chosen to distinguish the panels are likely to get cancelled out.⁷⁴

Overall, the panel approach has several advantages in convergence research. Convergence studies using this approach helped reveal the fact that persistent technological differences are a major cause of income dispersion in large samples of countries. The approach has its problems.⁷⁵ However, by taking advantage of the cross-section and time-series dimensions of data, it is possible to try to deal with many of these problems. It is in view of this potential that Temple (1999) concludes in his survey that, "the use of panels is often the best way forward. ..." (p. 113)

8. The Time Series Approach to Convergence Study

A progression of convergence study from the cross-section to the panel and then to the time series approach can be viewed as a natural response to the across-within tension of the convergence concept. However, many researchers early on saw the scope of application of time series methods to the study of convergence. Most of these studies proceed from standard reduced form equations of the output process. However these equations may be linked with the model-based growth-convergence equation, as illustrated below.

8.1. The Time Series Equation for β -Convergence

The commonly used equation for the time series analysis of convergence can be derived directly from the equation for β -convergence given by (15). This generally involves the assumption that the $x_{i,t-1}$ remains unchanged over the sample period considered. In that case, $\beta\psi x_{i,t-1}$ becomes just another time invariant term, and it can be subsumed under the term μ_i . Also, note that substituting $t_2 = t$ and $t_1 = t - 1$ in the expression for η_t , we get

$$\eta_t = g(t_2 - (1 + \beta)t_1) = g[t - (1 + \beta)(t - 1)] = (1 + \beta)g - \beta gt. \quad (16)$$

For an individual economy, $(1 + \beta)g$ is a constant, and hence can also be subsumed under μ_i , so that η_t effectively reduces to $-\beta gt$. Introducing these changes and upon rearrangement and suppressing the country subscript i , and adding an *iid* error term in an arbitrary fashion, we get from (15) the following:

$$y_t = \mu - \beta gt + (1 + \beta)y_{t-1} + \varepsilon_t. \quad (17)$$

This is the Dickey-Fuller equation with a drift and linear trend. For convergence in the usual sense, β should be negative. In other words, $(1 + \beta)$ should be less than one. The question then reduces to whether or not y_t has a unit root.⁷⁶ To the extent that the trend in equation (17) is *deterministic*, a test of unit root based on this type of equation is thought to yield a test of ‘deterministic convergence.’ It is clear that the deterministic trend of equation (17) is a direct result of the deterministic specification of the technology term of the growth model, whereby $A_t = A_0 e^{gt}$.

Some researchers have however allowed the trend to be *stochastic*. Unit root tests based on equations with stochastic trend are thought to yield tests of ‘stochastic convergence.’ A stochastic trend in equation (17) will however require a stochastic specification of A_t . So far there have been few presentations of growth models with stochastic specification of A_t . Binder and Pesaran (1999) is an important step in that direction. They provide stochastic variants of the NCGT and the Ak models based on explicit stochastic formulation of A and L . These authors test the cross-section implications of these stochastic growth models for the capital-output ratio and find more support for the stochastic version of the NCGT as compared with the deterministic NCGT and the Ak (either stochastic or deterministic) growth models. Their paper therefore can provide the theoretical basis for the concept of ‘stochastic convergence.’⁷⁷

8.2. Time Series Analysis of Within-Convergence

While the traditional unit root analysis has been limited mainly to output series of developed countries, under the convergence paradigm, the analysis is extended to a larger sample of countries. Lee *et al.* (1997), for example, conduct an exercise along this line and find that, out of 102 countries for which the equation is fitted, only for a few the null of unit root can be rejected.⁷⁸ Since the Dickey-Fuller unit root tests take non-stationarity as the null,⁷⁹ Lee *et al.* also use the Kwiatkowski *et al.* (1992) test, which uses stationarity as the null.⁸⁰ Using this test, Lee *et al.* find that the number of countries for which stationarity can be rejected “fell steadily with the length of the truncation parameter.” When this parameter is set at eight, the number of rejections falls to only nine.

It is important to note here that the above tests are based on the assumption of $x_{i,t-1}$ being constant. However, elements of $x_{i,t-1}$ may change over time even within an economy.⁸¹ Lee *et al.* recognize the possibility of ‘once for all changes’

taking the form of 'shifts or take-offs,' and observe that their finding of non-stationarity for some of the countries may be the result of not taking account of these changes.⁸² There now exists a whole line of research focusing on the sensitivity of rejection of unit root with respect to allowance of 'breaks,' which is one of the forms that changes in $x_{i,t-1}$ may take. This research shows that introduction of simple trend breaks (either exogenous or endogenous) leads to large increase in the number of rejection of unit root.⁸³ In view of these considerations, it cannot be said that the time series analysis has produced evidence that 'within convergence' does not hold for many countries.

8.3. *Time Series Analysis of Across-Convergence*

Since convergence is primarily understood in the across-economy sense, time series studies also focus on across-economy convergence. Attention here is directed to testing of unit roots in deviations, as defined by equation (3). Notice that going by the definition of equation (3), there may be 'across-convergence' even if output processes of individual economies contain unit root and hence lack 'within-convergence.'

8.3.1 *Time series analysis of convergence across the US states*

From a chronological point of view, time series analyses of 'across-convergence' began with the analysis of convergence across regions or states of the US. For example, Carlino and Mills (1993) analyze per capita income of eight geographic regions of the US.⁸⁴ They define the deviation series as $Dy_{jt} = (\bar{y}_t - y_{jt})$, where y_{jt} is the log per capita output of the region j and \bar{y}_t is the average for the USA as a whole. Application of the augmented Dickey-Fuller test on a variant of the equation (17) with y_t replaced by Dy_t yields rejection of the unit root hypothesis for a majority of the regions, thus providing evidence of convergence.⁸⁵

In Carlino and Mills' deviation-setup above, the region specific intercept term of the equation stand for (log of) the term ' a ' in equation (3) and may be interpreted as standing for *time-invariant* differences in the determinants of the steady state across regions. This is equivalent to allowing some components of $x_{i,t-1}$ to vary, albeit only in the direction of i . In addition, they feel the necessity of allowing trend break, which can be interpreted as allowing $x_{i,t-1}$ to vary in the direction of t , albeit in a very restricted way. Thus Carlino and Mill's evidence may be characterized as of 'conditional convergence.' Lowey and Papell (1996) extend this analysis further by endogenizing the timing of the break.⁸⁶ These authors conduct the analysis at a further disaggregated level by dividing the US into 22 regions instead of 6. Again, the null of unit root is rejected for majority of the regions, confirming evidence of (conditional) convergence.

While Carlino and Mills and Lowey and Papell analyze the deviation data region by region, Evans and Karras' (1996b) conduct similar analysis by *pooling* the deviation data for individual states.⁸⁷ In view of the weakness of the standard Dickey-Fuller test, Evans and Karras use a modified version of the unit root test proposed by Levine and Lin (1993) that is designed specifically for pooled data.

The results show rejection of the unit root hypothesis even when trend breaks are not included. In this setup too, the *state specific* intercept term of the equation stands for (time-invariant) difference in steady state among the individual states. Hence, Evans and Karras' finding may also be interpreted as one of *conditional* convergence. In fact, the authors themselves contrast this aspect of their result with Barro and Sala-i-Martin (1992)'s result of *unconditional* convergence across US states.⁸⁸

8.3.2 *Time Series Analysis of Convergence across Countries*

Unit Root Analysis of Pooled Data for Countries: The time series analysis has been applied to investigate convergence across countries too. In fact, Evans and Karras (1996a) conduct a similar unit root analysis of pooled *deviation* (from average) data for a sample of 56 countries. The results favor rejection of unit root and by implication favor the conditional convergence hypothesis. Analogous results are also obtained in Evans (1996) from analysis of long historical data (1870–1989) for a sample of thirteen developed countries. Earlier Quah (1990) presented an analysis of per capita income (in the form of deviations from that of the US) for 114 countries ranging from 1970 to 1985. Noting both large N and T , Quah develops and applies the inference theory appropriate for 'random field data' and rejects the null of no unit root. However, unlike other researchers mentioned above, Quah does not allow for country specific intercepts in his analysis of the deviations. Accordingly his analysis is more of the unconditional convergence hypothesis, and it is not surprising that this hypothesis is rejected. As we saw, the hypothesis of unconditional convergence in large sample of countries was rejected by other methodologies too.⁸⁹

Vector Approach to Across-Convergence Study: In many of the studies above, the unit root analysis was conducted either on deviations of individual economies one-at-a-time. One problem with this analysis is that the source of rejection of the null of unit root is not always clear. When deviations are taken from a reference economy, such a rejection can occur because the output process of either the reference economy or the economy in question contain unit root while that of the other does not. On the other hand, the rejection may also occur because the output processes of both these economies contain unit root and they are not co-integrated. The problem remains when deviations are taken from the average of the sample. Presence of a unit root in the output process of just one of the economies of the sample will cause the average to contain unit root, and then deviations for all the economies (except for the one that has unit root) will also contain unit root. The problem persists when unit root analysis is conducted on pooled deviations for all countries. A second problem with one-at-a-time analysis of deviations from a reference economy arises when there is no agreement about which to choose as the reference economy. On the other hand, the dimension of the exercise becomes simply too large if deviations of all countries from the remaining all are to be studied one-at-a-time in a sizable sample.

An alternative approach, that albeit does not resolve all the problems above, is to conduct cointegration analysis in a vector setting.⁹⁰ Bernard and Durlauf (1995) take this route and analyze cointegration in a sample of fifteen developed countries.⁹¹ Their goal is to check whether the per capita output series of these economies are co-integrated or not, and if they were, whether the co-integration vector is of the form $(1, -1)$ or $(1, -a)$, where a is a constant. Under the vector approach, the hypotheses can be formulated in terms of conditions on the rank of the spectral density matrix at frequency zero of ΔDY_t and ΔY_t . Here, ΔY_t is the first difference of Y_t , which, in turn, is the vector of individual (per capita) output series, y_{it} . Similarly, DY_t is the vector of Dy_{it} , the deviation of (per capita) output of country i from that of the reference country having index 1.⁹²

Bernard and Durlauf use two sets of procedures to carry out the tests: one, based on Phillips and Ouliaris (1988), and the other, based on Johansen (1988). In either case, the conclusion is broadly similar: there is evidence of co-integration of the form $(1, -a)$ but not of the form $(1, -1)$. The authors interpret this result as showing that the countries 'shared common trends' but did not converge. However, as we have noted before, co-integration of the form $(1, -a)$ can also be interpreted as a manifestation of conditional convergence.⁹³

Broadly therefore time series analysis supports a variant of the conditional convergence hypothesis. Viewed in this way, the results produced by the time series studies of convergence have not been all that different from those produced by studies following either the cross-section approach or the panel approach. Although we tried in the above to link the time series equation with an underlying growth model, most of the time series studies of convergence do not do so. Their analysis is limited entirely to reduced form equations, and no attempt is made to link the estimation results with parameters of the growth model. This also limits the policy relevance of these studies.

9. The Distribution Approach and σ -Convergence

While the cross-section, panel, and (in part) time-series approaches have in one way or the other investigated β -convergence, the distribution approach focuses on σ -convergence and on changes in the cross-section income distribution as a whole. However, as noted earlier, this correspondence is not that simple. The distribution approach has actually proceeded along two lines. The first maintains a relationship with β -convergence and tries to work out the precise relationship between β and σ . The second emphasizes the limitations of β -convergence and focuses on the shape of the entire distribution. We begin the review with a discussion of the first.

9.1. Relationship between β -convergence and σ -convergence

In order to see the relationship between β - and σ -convergence, it is worthwhile to start from the decomposition of the cross-section variance into its constituent elements. In offering this decomposition, BS for example note that if all the terms

other than $y_{i,t-1}$ and ε_{it} in equation (15) are ignored, the evolution of σ_t^2 , variance of y_{it} , under suitable assumptions on ε_{it} , can be described by

$$\sigma_t^2 = (1 - \beta)^2 \sigma_{t-1}^2 + \sigma_\varepsilon^2 = \tilde{\beta}^2 \sigma_{t-1}^2 + \sigma_\varepsilon^2, \quad (18)$$

where σ_ε^2 is the variance of ε , and $\tilde{\beta} = (1 + \beta)$. Iterating backwards, this yields

$$\sigma_t^2 = \frac{\sigma_\varepsilon^2}{1 - \tilde{\beta}^2} + \left(\sigma_0^2 - \frac{\sigma_\varepsilon^2}{1 - \tilde{\beta}^2} \right) \tilde{\beta}^{2t}. \quad (19)$$

As $t \rightarrow \infty$, the above approaches the steady state value $\sigma_\infty^2 = \sigma_\varepsilon^2 / (1 - \tilde{\beta}^2)$. It is clear that σ_∞^2 increases with σ_ε^2 and decreases as β becomes more negative. What is more important is that σ_t^2 can monotonically *either increase or decrease* to σ_∞^2 depending on whether the initial variance σ_0^2 is smaller or greater than the steady state variance σ_∞^2 . This algebraic result again shows that a negative β cannot guarantee falling variance. In that sense, β -convergence is not sufficient for σ -convergence. However, it also shows that an empirical finding of *increasing* cross-sectional variance is not incompatible with β -convergence.

A more extended form of the relationship between β and σ obtains if the other right hand side variables of equation (15) are not ignored. Lee *et al.* (1997) work out this relationship as shown in equation (20) below:

$$\sigma_t^2 = \tilde{\beta}^{2t} \sigma_0^2 + [1 - \tilde{\beta}^{2t}] \sigma_{*0}^2 + \left[\frac{1 - \tilde{\beta}^{2t}}{1 - \tilde{\beta}^2} \right] \sigma_\varepsilon^2 + \left[T - \frac{1 - \tilde{\beta}^{2t}}{1 - \tilde{\beta}^2} \right] \sigma_g^2, \quad (20)$$

where σ_{*0}^2 is the cross-country variance in steady state per capita output in time 0, and σ_g^2 is variance of the steady state growth rate, g . Under the assumption of a common g , this last term drops out. Then the above expression reduces to the same as in equation (19), except that it now has the additional term involving σ_{*0}^2 . As expected, the latter term now also appears in the expression for σ_∞^2 :

$$\sigma_\infty^2 = \sigma_{*0}^2 + \frac{\sigma_\varepsilon^2}{1 - \tilde{\beta}^2}. \quad (21)$$

Substituting for σ_{*0}^2 in equation (20) we get

$$\sigma_t^2 = \sigma_0^2 + [1 - \tilde{\beta}^{2t}] (\sigma_\infty^2 - \sigma_0^2). \quad (22)$$

This is now the same as equation (19) above, because σ_{*0}^2 gets subsumed under σ_∞^2 . Equation (22) helps to see again that dispersion may either increase *or* decrease towards σ_∞^2 , depending on whether initial dispersion, σ_0^2 , is less or greater than σ_∞^2 .⁹⁴

Thus we see that β and σ^2 are algebraically related, and the value of one can be obtained from that of the other, provided some other conditions are satisfied. This implies that tests of these two concepts of convergence can also be related.

This is of particular importance for σ -convergence, because, unlike β -convergence, statistical tests for σ -convergence were not readily available.

9.2. Relationship between Tests for β - and σ -convergence

In trying to formulate a test for σ -convergence, Litchenberg (1994) for example observes that, ignoring other terms, from (15) we can also have

$$\frac{\sigma_t^2}{\sigma_{t-1}^2} = \tilde{\beta}^2 + \frac{\sigma_\varepsilon^2}{\sigma_{t-1}^2}. \quad (22)$$

The ratio $\sigma_t^2/\sigma_{t-1}^2$ provides evidence regarding σ -convergence, because it shows whether the dispersion is increasing or decreasing over time. One can try to get information about this ratio *directly*, as do for example Miller (1995) and Lee *et al.* (1997). Lichtenberg however shows that this ratio can be estimated *indirectly* from $\tilde{\beta}$ and the R^2 of the cross section regression estimating β . Since $1 - R^2 = \sigma_\varepsilon^2/\sigma_t^2$, it follows from equation (22) that

$$\sigma_t^2/\sigma_{t-1}^2 = R^2/\tilde{\beta}^2. \quad (23)$$

This shows that a test statistic for σ -convergence can be obtained from $\hat{\tilde{\beta}}$ by adjusting it using R^2 to account for the distribution of the shock term.⁹⁵

Litchenberg suggests that the test statistic obtained from equation (23) have an F -distribution with $[n - 2, n - 2]$ degrees of freedom.⁹⁶ However, Carree and Klomp (1997) point out that Litchenberg's conclusion regarding appropriate distribution for the test statistic implied by (23) is not entirely correct. They draw attention to the fact that F -distribution is valid if σ_t^2 and σ_{t-1}^2 are independent of each other, which will not be true provided $\tilde{\beta} \neq 0$.⁹⁷ These authors try to 'salvage Litchenberg's idea of using F -distribution to test for σ -convergence by showing that if the sample is large, so that $\hat{\tilde{\beta}}$ can be thought 'close' to $\tilde{\beta}$, then $T_2 = (\hat{\sigma}_t^2/\hat{\sigma}_{t-1}^2 - \hat{\tilde{\beta}}^2)/(1 - \hat{\tilde{\beta}}^2)$ will be distributed approximately as $F(n - 2, n - 1)$.⁹⁸ Carree and Klomp (1997) suggest yet another variant of the statistic that takes into account the variability in the estimate $\hat{\tilde{\beta}}$. This is given by $T_3 = (\hat{\sigma}_t^2/\hat{\sigma}_{t-1}^2 - (\hat{\tilde{\beta}}^2 - z_\alpha \hat{\sigma}_{\tilde{\beta}})^2)/(1 - (\hat{\tilde{\beta}}^2 - z_\alpha \hat{\sigma}_{\tilde{\beta}})^2)$, where $\hat{\sigma}_{\tilde{\beta}}$ is the standard error of estimated $\tilde{\beta}$ and z_α is the adopted critical value from the standard normal distribution.⁹⁹

Evidence regarding σ -convergence can therefore be obtained both directly and indirectly, and the evidence may or may not be formalized in the form of statistical tests. We now look at some evidence regarding σ -convergence.

9.3. Evidence regarding σ -convergence

Not unexpectedly, evidence regarding σ -convergence depends very much on the sample. For the OECD countries, data have generally favored σ -convergence. Lee *et al.* (1997), for example, compute variance of cross-section distribution of log of

per capita income for different samples of countries for 1961 to 1989 and plot them against time. Their results show that the variance for the OECD sample has decreased over time. Miller (1995) and other researchers have produced similar results regarding the OECD sample. Going by the classification above, these are *direct* evidence of σ -convergence.

On the other hand, Litchenberg (1994) uses his procedure to formally test the hypothesis of σ -convergence. He runs a simple regression of $\ln GDP_{85}$ on $\ln GDP_{60}$ for the OECD countries and uses the result to compute the test statistic as per equation (23) above.¹⁰⁰ Application of the critical values from F -distribution results in a non-rejection of the null of non-convergence.¹⁰¹ However, Carree and Klomp (1997) redo the exercise using their statistics and reverse the conclusion. For the period 1960–85, all of Carree and Klomp's three statistics report convergence.¹⁰² In other words, the conclusion of σ -convergence for the OECD countries is upheld. Similar evidence of σ -convergence has been found for other smaller samples of countries. Evidence of σ -convergence has been reported for the US states too.¹⁰³

However, for large, global samples of countries, the evidence generally indicates a rise in variance. For example, according to Lee *et al.* (1997)'s computation, output-variance in the sample of 102 countries increased from 0.77 to 1.24 between 1961 and 1989. Other researchers have furnished similar evidence. There are different ways in which these results can be interpreted. For example, a rising σ in the global sample may indicate that the steady state dispersion, σ_{∞}^2 , itself has increased, which, in turn, may be the result of increased dispersion of the determinants of steady state. Alternatively, it is possible that σ_{∞}^2 has remained unchanged, but the initial variance, σ_0^2 , was less than σ_{∞}^2 , so that the variance increased from below towards the steady state variance.¹⁰⁴ In either case, the outcome is not incompatible with conditional β -convergence.

Similarly, it is not known how much of the decrease in variance in small sample of developed economies is due to negative β and how much due to reduction in other items of (20), including reduction in the dispersion of the steady state determinants. All this indicates that more knowledge is needed about the dynamics of the steady state determinants in order to understand the changes in the cross-section variance of income.

The research on σ -convergence focuses on only one feature of the cross-section distribution, namely the variance. It has been argued that attention should rather be given to evolution of the entire shape of the distribution. This line of research also goes beyond the anonymity of distribution and identifies the position of individual or groups of countries within the distribution, and it notices how these positions change over time. We now turn to a review of this line of convergence research.

9.4. *Study of the Evolution of the Cross-sectional Distribution*

The research focussing on the shape of the cross-section distribution of income has been carried forward almost single handedly by Danny Quah.¹⁰⁵ To capture more details about the distribution, Quah focuses on the probability mass at

different quantiles. A simple plotting of the cross-section distribution of the global sample for successive years already displays two features: first, the cross-section distribution is not collapsing, and, second, this distribution is becoming more *bi-modal*. However, since it is not known whether the plotted distributions are of steady state or not, and because the plots of distribution cannot tell the position of individual countries, Quah performs a more formal analysis using the following framework involving Markov transition matrix:

$$F_{t+1} = MF_t, \quad (26)$$

where F_t is the cross-section distribution at time t , and F_{t+1} is the same at time $t + 1$, and M is the transition matrix that maps F_t onto F_{t+1} . The goal is to know M , which determines the evolution of the distribution. Assuming M to be unchanged over time, we have,

$$F_{t+s} = M^s F_t, \quad (27)$$

where s is any particular length of time (number of years, say). Letting $s \rightarrow \infty$ this also allows to obtain the steady state distribution.

Quah calibrates M using actual data. Both the calibrated transition matrices and ergodic distributions obtained on their basis lead to similar conclusions. The first is ‘persistence.’ The values of the diagonal elements of the one-year M -matrix are in the neighborhood of 0.9, implying that most of the countries continue to remain in the same position (or range) of the distribution. Second, whatever mobility (within the distribution) exists, it works to ‘thin out the middle,’ and ‘pile up of probability mass at the two tails.’ This is Quah’s result of growing ‘twin-peakedness’ or bi-modality of the distribution. The results do not change if higher order specifications are used. In fact, these make the bi-modal property and ‘poverty piling up’ even more pronounced.¹⁰⁶ This exercise is extended in Quah (1993a) where he lets the quantiles to evolve (instead of being fixed).¹⁰⁷ However, the results remain more or less intact. In fact, now the dynamics of the quantiles further confirm these results. Thus, the formal analysis using transition Markov transition matrix confirms what informal plotting of distribution of successive years already suggested.¹⁰⁸

The findings of ‘persistence’ and ‘bi-modality’ are valuable additions to the known stylized facts about cross-country growth regularities. Growth theories now need to explain these facts. The Markov analysis itself does not help in the explanation. It is another type of reduced form analysis. The transition matrix M is memory-less, and no growth theory is required for its estimation; no structure is imposed on the data.¹⁰⁹ In fact, Quah makes it explicit that it is his intention not to be restricted by assumptions of long term growth. However, Quah has made suggestions about the directions in which explanations for the stylized facts have to be sought. One is along models of multiple equilibrium that yield ‘club convergence.’ At a more fundamental level, Quah (1997) suggests that the dynamics toward bi-modality can be explained by ‘spatial spillovers,’ which in turn are determined not so much by ‘openness’ of an economy as by ‘who trades with

whom,' i.e., the trading partners. More research will be necessary to resolve these issues. But these are useful directions for further research on growth.

10. Conclusions

Research on convergence has indeed proceeded in many directions using many different definitions and methodologies. It is therefore not unreasonable to feel somewhat dazed by the variety of results and conclusions encountered in the literature. A close review however reveals that at a broad level there is considerable agreement among the results. For example, despite differences in approach and methodology, the finding of *conditional* β -convergence has remained relatively robust. This has been true both for small samples of developed economies and for large, global samples. For developed economies, researchers have in fact often reported unconditional convergence. Similarly, once it is remembered that σ -convergence research generally focuses on *unconditional* convergence, it becomes clear that results regarding σ -convergence largely agree with those regarding β -convergence. Evidence of σ -convergence is found precisely in those small samples of developed economies for which there is also evidence of *unconditional* β -convergence. On the other hand, in large global samples, neither unconditional β -convergence nor σ -convergence holds. Finally, time series analysis of both within and across convergence has produced evidence that can be interpreted as of *conditional* convergence.

However, the convergence research has not produced consensus at more concrete levels. For example, agreement about estimated values of structural parameters of growth models, such as the rate of convergence and the elasticity of output with respect to capital, has proved elusive. In fact, not all approaches to convergence research have been equally concerned with values of structural parameters.¹¹⁰ Given the differences in approach, sample, data, model, estimation technique, etc., absence of consensus regarding parameter values is not surprising. However, some generalities have emerged even in this regard. It has been observed that the more differences in the steady state of economies are controlled for (either by sample selection or by inclusion of relevant variables in the regression), the higher are the resulting convergence rates. In particular, convergence rates prove to be much higher when technological differences across countries are taken into account.

The implications of the convergence results have several dimensions. First of all, the welfare implication of the conditional convergence finding for global samples is rather limited, because it only means that poor countries are moving toward their own steady states, and knowing this may be of little solace if those steady state income levels are themselves very low. This shows that it is more important to focus on the determinants of the steady state levels. Also, improvements in steady state levels cause transitional growth, and a sequence of transitional growth brought about by intermittent improvements in steady state levels may not be that different from the 'growth effect' postulated by many NGT models.

The implications of the convergence results for the growth theory debate are more controversial. As Barro noted, the general finding of conditional convergence conforms to the basic NCGT assumption of diminishing returns to capital. The rate of conditional convergence gets extra attention because of its relation with the elasticity of output with respect to capital. A lower convergence rate is associated with a higher elasticity value, which in turn can be indicative of externality if this value proves to be much larger than that obtained from national accounts. This has some bearing on the first generation new growth models that rely on externality as the source for long run growth. However, the new growth models have themselves moved on from externality to profit motivated research as the source of growth. Hence the debate about externality via the rate of conditional convergence may not be as significant now as it was at the beginning of the growth debate.¹¹¹

More important in this regard is to notice the metamorphosis that both NCGT and NGT have undergone as a consequence of the convergence research. On the one hand, absence of absolute convergence in large samples of countries has forced NCGT to recognize the differences in steady state income levels across countries. Many researchers now also allow the steady state growth rates to differ across countries under NCGT, making it consistent even with 'divergent' behavior. On the other hand, the empirical finding of conditional convergence has led to the emergence of many NGT models that yield convergence implication. These NGT models have followed different routes to reach this implication.¹¹² However, this shows that the convergence implication can no longer be associated solely with the NCGT. As a result of this give and take between NCGT and NGT, it is now possible to explain both convergence and non-convergence behavior by appropriately chosen models of growth theory of both the varieties. This may prove disappointing from the original idea of using convergence as a test for validity of alternative growth theories. However, the fact that convergence debate forced both types of growth theories to move toward accommodation may not be a mean achievement by itself.

However, the convergence research has had other important payoffs. First, it has furnished new stylized facts regarding cross-country growth regularities, such as 'persistence' and 'bi-modality.' The growth theory faces the task of explaining these facts, and this is drawing attention anew to models of multiple equilibrium. Second, convergence research has highlighted the existence of very large productivity or technological differences across countries and has given rise to new methodologies for quantification of these differences. As a result of these quantification efforts, a new information base is emerging for the study of determinants of technology differences and diffusion. The empirical finding of large technological and institutional differences is leading to theoretical developments too. New models of technology diffusion are being proposed, along both NGT and NCGT traditions.¹¹³ The information base furnished by the convergence research will prove helpful in discriminating among alternative models of technology diffusion.

There are several other directions in which the convergence research is leading the growth literature to develop. Most of the convergence studies so far have been conducted on the basis of closed economy growth models. In reality, however, the economies are not closed, and actual data embody the effects of interaction that occur among economies along the lines of trade, technology transfer, capital flow, labor migration, diffusion of institutions, etc. Barro, Mankiw, and Sala-i-Martin (1995), Quah (1996c), and others attempt to include capital mobility in the explanation of cross-country growth regularities. There have been recent advances in formal modeling of the impact of trade on growth too.¹¹⁴ However, empirical studies embodying the insights of these theoretical developments are yet very few.

The convergence studies so far have also been based on deterministic growth models. This has limited both theoretical and policy implications of the empirical time series studies of convergence. However, researchers are now formulating stochastic growth models and deriving their implications for cross-section and time series data. These stochastic formulations shift the discussion from the usual notions of convergence to the ergodicity of variables. Future growth and convergence research has to pay special attention to the time series properties of the technology or productivity term, A .

Finally, empirical research on growth and convergence of the recent period has heavily depended on the Summer-Heston data set. While this data set has earned appreciation,¹¹⁵ and is ubiquitous in terms of use, it has also been the target of considerable criticism.¹¹⁶ Part of the future research effort may also be usefully directed toward improving and generating data that are necessary for better understanding of the growth and convergence issues.

Acknowledgments

I would like to thank Jesus Felipe, Joy Mazumdar, John Pencavel, Jonathan Temple, and several anonymous referees for their comments on earlier drafts of this paper. Special thanks are due to Dale Jorgenson, whose encouragement was very crucial for writing and publishing this paper. For correspondence, send e-mail to nislam@emory.edu.

Notes

1. This does not mean that what follows is a chronological narration *per se*. It is chronological to the extent that this is required to follow the logical progression.
2. As Barro (1997) puts it, "It is surely an irony that one of the lasting contributions of endogenous growth theory is that it stimulated empirical work that demonstrated the explanatory power of the neoclassical growth model." (p. x)
3. We include here the Inada (1963) conditions as part of the NCGT assumption of diminishing returns. This is important because, as we shall see, a particular variant of the NGT retains diminishing returns but relaxes the Inada conditions to obtain asymptotic growth.

4. The fifth of these stylized facts was that the growth rate of per capita output varied widely across countries, and the sixth was that economies with high share of profits in income had higher investment to output ratios.
5. In this paper we shall use β as the generic notation for the coefficient on the initial level variable in the growth-initial level regressions. Note that negative β can be interpreted as evidence of convergence in terms of *both* income level and growth rate.
6. They point out that a negative β can just be an example of the more general phenomenon of reversion to the mean, and, by reading convergence in it, growth researchers are falling into Galton fallacy.
7. See for example Mankiw, Romer, and Weil (1992) or Barro and Sala-i-Martin (1995) for the derivation.
8. In the case of the Cass- Koopmans model, θ also has similar set of elements with s replaced by parameters for the rate of time preference and the elasticity of inter-temporal substitution in consumption.
9. For models with multiple equilibria, see, for example, Azariadis and Drazen (1990).
10. The numerical magnitudes of β from different studies are not directly comparable because of the differences in regression specification.
11. The numerical results of this regression were not presented, but Baumol reported that it yielded 'slightly positive slope,' indicating a process of rather divergence.
12. One example was the OECD group, already considered. Another example, according to him, was the group of formerly centrally planned countries. According to Baumol, such clubs consisted of countries which had certain degree of homogeneity in 'product mix and education' enabling them to share in the 'public good properties of the innovations and investments of other nations.' (p. 1080) The idea of technological diffusion is also reflected here.
13. The proper criteria for sample selection for convergence study, DeLong (1988) points out, is *ex-ante* income level, and not *ex-post*. In particular, he shows that, if, guided by the *ex-ante* criterion, Baumol's OECD sample is modified slightly, the result of unconditional convergence no longer holds. Baumol largely accepts this criticism. See Baumol and Wolff (1988).
14. For the OECD sub-sample, they find the coefficient on the initial income variable to be negative. For the larger ROW sub-sample, this coefficient turns out to be positive. However, upon splitting ROW into (three) smaller samples, they find the sign of β to vary.
15. Grier and Tullock (1989) were themselves quite keen about this limitation. They make only the broad observation that their results are generally supportive of NCGT, and that these results may show directions for further development of NGT. For their further comment on this point, see (p. 260).
16. There is an earlier tradition of cross-country growth regressions conducted by development economists that do include human capital variables. See for example Krueger (1968). These regressions are often intended to determine patterns of economic development or to ascertain determinants of growth in general. Chenery, Robinson, and Syrquin (1985) provide a useful summary of this body of work. The current wave of empirical growth research however is not that aware of and does not recognize much connection with this earlier tradition of work.
17. Also, Barro was not an outsider to this debate. He embarks on this empirical work after already making his own contribution to the development of NGT. In Barro (1990) he examines the role of government spending in the setting of an *Ak*-style model of endogenous growth. Becker and Barro (1988) and Barro and Becker (1989)

address the issue of fertility choice. These works and his close association with the genesis of NGT, gave Barro quite a few propositions to test out, and he is quite explicit that he is doing so 'using recent theories of economic growth as a guide.' (p. 437)

18. Barro argues for this simultaneity on the basis of new theories of growth. In particular, he cites extensively the conclusions of Rebelo (1991), Barro (1990), Romer (1990), Barro and Becker (1989) and Becker, Murphy and Tamura (1990).
19. Barro's work is, thus, not limited to the convergence issue; instead it addresses a host of other issues related to growth. It therefore, to some extent, shares the characteristic of multiple focus mentioned earlier.
20. Other right hand side variables of Barro's regressions include government consumption (reflecting his earlier interest in it), an index of relative inflation, indicators of political stability and some regional and continent dummies. Although the inflation variable is interpreted as a measure of 'market distortion,' its inclusion can as well be linked to the output-inflation literature, in which Barro himself played a no small role. See for example, Barro (1976) and (1978). In fact, Barro has returned to this issue, as can be seen in Barro (1997).
21. The correlation between the average growth rate of per capita real gross domestic product between 1960 and 1985 (GR6085) and the 1960 value of real per capita GDP (GDP60) is reported to be positive 0.09.
22. As Barro (1991) puts it, "This finding accords with recent models, such as Lucas (1988) and Rebelo (1991), that assume constant returns to a broad concept of reproducible capital, which includes human capital. In these models the growth rate of per capita product is independent of the starting level of per capita product." (p. 408)
23. The inclusion of human capital causes β to be negative in the regressions with investment as dependent variable as well, and Barro interprets this too as "consistent with the convergence implication of the neoclassical growth model." (Barro 1991, p. 427). With the ratio of private investment to GDP as the dependent variable, the coefficient on the initial income is significant and ranges between $-.0093$ and $-.0098$.
24. The argumentation for the specification is, however, not neo-classical. Instead, Barro observes that NGT relationships among growth, investment, and fertility imply that residuals from the growth regression will be positively related with those from investment regression, and negatively related with residuals from fertility regression. He justifies the conventional format of the growth regression as an alternative way of checking whether the stipulated relationships among the residuals are true.
25. The coefficient is a little different in magnitude from that obtained from earlier set of growth regressions that do not include investment and fertility rates as control variables. Barro interprets this as showing that the negative effect of the initial income on growth does not work through its effects on investment or fertility. Instead, it works mainly through lower rate of return on investment. (p. 430)
26. De Long and Summers (1991) show that 'equipment investment' has strong positive effect on growth, and this effect does not depend on 'education infrastructure.' Inclusion of human capital variables, which appears in Barro's regression, does not affect the coefficient of the equipment investment variable in De Long and Summers' study. These authors interpret their regression results as showing the presence of positive externalities associated with equipment investment. De Long and Summers (1993) extend these results for developing countries. Some authors have however expressed skepticism about De Long and Summers' claim. For example, Auerbach *et al.* (1994) argue that countries of the De Long Summers' sample are too heterogeneous to have

the same technology, so that equipment investment was also capturing impact of technological diffusion. Blomstrom, Lipsey, and Zejan (1996) argue that direction of causality runs from growth to equipment investment, rather than the other way round. Temple (1998a) and Temple and Voth (1998) return to the issue and provide some support for De Long and Summers' view. With hindsight from subsequent research, it may be said that the roles of human capital and equipment investment need not be counter-posed. While the research has confirmed human capital's important role in growth, it is not conclusive about the precise way in which this role is played. It is also not clear what the best way is to develop human capital. Equipment investment may have role both in facilitating human capital's role and in the development of human capital (through, for example, learning by doing).

27. See for example Barro and Sala-i-Martin (1995), Mankiw (1995), and Durlauf and Quah (1999).
28. An important point here concerns whether to make $\ln \hat{y}^*$ contingent on t_1 or not. If it is assumed that the determinants of the steady state income remain constant between t_1 and t_2 , it does not matter whether $\ln \hat{y}^*$ is made contingent on or not. Generally, in implementation, it is assumed that such determinants of steady state income as s , n , δ , and α , remain the same between t_1 and t_2 . This suggests that it is not necessary to make steady state income contingent on the initial period in considering the transitional dynamics. This is certainly the case in the continuous time setting, where the difference between t_1 and t_2 is instantaneous. However, when dealing with steady state income per capita (as is usually the case), it is necessary to be careful. As we can see from equation (2), the formula for steady state income per capita has the term $A_0 e^{gt}$, which will differ if evaluated at t_2 instead of at t_1 . In the logarithmic version of the equation, this will generally imply differences in terms involving g and may not affect the basic conclusions. However, it is worth to be aware about the issue.
29. Unfortunately, researchers do not always pay adequate attention to this within-across tension in the interpretation of the convergence rate parameter λ . Durlauf and Quah (1999) takes appropriate note of this important issue.
30. It is taken as evidence showing that, other things held constant, countries starting with lower levels of income grow faster. See for example Kormendi and Meguire (1985, p. 147). This however also meant that no information regarding the structural parameter values could be obtained.
31. For the NONOIL sample, for example, λ equals 0.00606, implying a half-life of 114 years, which is indeed very long. It needs to be mentioned that NONOIL, INTER, and OECD are three samples used in the MRW study. The NONOIL is the sample of 98 countries which include almost all the sizable countries of the Summer-Heston data set (Summers and Heston; 1988, 1991) except those for which extraction of oil was the dominating source of income. The INTER sample is a sub-sample of NONOIL consisting of 76 countries for which the quality of data is better. Finally, OECD is the sample of 22 OECD member countries. Much of the subsequent convergence literature has used these three samples.
32. MRW do not report the results of the restricted version of this regression, hence we do not have an unique estimate of α . Based on the coefficient of the s variable, its value would be 0.82, while based on the coefficient of $(n + g + \delta)$, it would be 0.68. Both of these are for the NONOIL sample and are far greater than the share of capital in the national income in the countries of this sample, computed on the basis of the national accounts data. These values agree with the estimates produced earlier by Romer (1989a) from a growth accounting exercise under similar assumptions.

33. However, it is to be noted that although BS use the neoclassical convergence equation to recover λ , they do not apply it strictly to determine the right hand side variables of the regression.
34. However, in some of their regressions BS include regional dummies and a variable proxying for output composition. Inclusion of these variables makes it a little ambiguous whether it is unconditional or conditional convergence that is investigated.
35. This was true in terms of both per capita income and product and for different time periods considered. BS left it as an unresolved puzzle why this rate proved to be similar in terms of income and product.
36. Recently, Caselli and Coleman (2001) have reexamined convergence across US states and have emphasized the role of structural transformation in the observed convergence.
37. Shioji (1995) investigates convergence across prefectures of Japan. Lusigi, Piesse, Thirtle (1998) look at converge among countries of the African continent.
38. The authors conduct two sets of exercises. In the first, the countries are grouped on the basis of arbitrarily chosen cut off levels of initial income and literacy. Apprehending selection bias in such grouping, the researchers present a second exercise in which the grouping is endogenized using the 'regression-tree' method. The results obtained from these two methods of grouping however prove to be qualitatively similar.
39. Since this instability pertains to groups classified according to both initial income and human capital levels, the authors conclude that both of these variables are important in identifying the 'regimes.'
40. Durlauf and Quah (1999) also discuss this problem in their survey.
41. The *ratio* of (initial income and literacy) levels may have better promise in this regard, because although the countries cross all the level values sooner or latter, they may not do so observing the same ratio of these levels. Durlauf and Johnson found that output dominated literacy as a criterion for group/regime identification.
42. In fact, one of Durlauf and Quah (1999)'s main arguments for rejection of Barro's conclusion regarding convergence research is that this research has so far limited itself to use of linear specifications only.
43. Some authors refer to other sources of convergence, such as changes in sectoral composition (alternatively expressed as structural transformation). Changes in sectoral composition usually find reflection in changes in capital intensity, so that the latter may subsume the former. Also, one-sector models, that have been the basis of most of the convergence studies, do not allow changes in sectoral composition to surface directly.
44. For recent empirical evidence on R&D spillovers see for example Coe and Helpman (1995), Coe, Helpman, and Hoffmaister (1997), and Bayomi, Coe, and Helpman (1999). For earlier historical discussion of technological diffusion see in particular Gerschenkron (1953), who coined the expression 'advantages of backwardness.'
45. Durlauf and Johnson (1995) also note the issue of potential differences across countries in aggregate production function. Observing wide variation in the estimated parameter values across groups, they observe that, "aggregate production function differs substantially across countries." In particular, large differences in the intercept term – which is related with A_0 of equation (14) – lead them to conclude that, "different economies have access to different aggregate technologies." (p. 375) Hence, they express the view that "...the Solow growth model should be supplemented with a theory of aggregate production function differences in order to fully explain international growth patterns." See also Bernard and Jones (1996b)

46. The exercise by De la Fuente (1996) shows that even after accounting for A_0 by carefully constructed variables, a part remains that prove significant in the regressions.
47. There are other possible sources of bias for the cross-section regression as well. See Lee *et al.* (1997) and Evans and Karras (1996a) on this point. Lee *et al.* (1997), for example, draw attention to another possible bias of the cross section growth regression. They start with an equation similar to (13) and assume t and $t-1$ to be one year apart. Since the cross-section regression generally considers growth over a long (say, 25 year) period, they try to have the equation correspond to that one by assuming initial $t = 0$, and iterating it forward to T . This gives rise to, among others, a composite error term $\xi_{iT} = \sum_{j=0}^{T-1} \beta^j \varepsilon_{i,T-j}$. Thus, possible serial correlation in ε now acts as a source of bias. So does possible across-country variation in A_0 and g . See Lee *et al.* (1997, Appendix A) for details.
48. Recall the earlier discussion about whether the variables representing determinants of steady state should be made contingent on the initial period or not. The formulation in equation (14) and (15) is on the basis of such contingent \hat{y}^* . However under the assumption that s and n remain constant between t and $t-1$, we can replace $x_{i,t-1}$ in the equation above by x_{it} .
49. Instead of hiding or sidetracking, the correlated effects model allows the correlation between μ_i and x_{it} 's to come to fore and play out its role in the estimation process.
50. The estimated values of λ for the NONOIL, INTER and OECD samples in Islam (1995) prove to be 0.0434, 0.0417, and 0.0670, respectively. Knight *et al.* report an even higher value, 0.0652, for the NONOIL sample.
51. The estimated values of α in Islam (1995) for the three samples above are 0.4397, 0.4245 and 0.2972, respectively. The value of α reported for NONOIL sample in Knight *et al.* (1993) is even lower, 0.335.
52. Incorporation of human capital in the panel analysis leads to 'anomalous' results, with the coefficient of the human capital variable turning out to be negative and generally insignificant. This agrees with earlier results regarding human capital obtained from pooled regressions (see for example Gregorio (1992)) and also the results obtained by Benhabib and Spiegel (1994). To the extent that the human capital data are weak, there may be some data issues related with this result. The schooling data used for construction of human capital variable are yet to be adjusted for quality differences. Also, many processes of human capital formation that occur outside of formal schooling are not included in this variable. Nevertheless, panel results indicate that the channel of influence of human capital on output may be more complicated than suggested by MRW though their proposal of multiplicative inclusion of human capital in the aggregate production function, alongside physical capital. Benhabib and Spiegel (1994) also find support for multiple and more complex channels of influence of human capital on output. Panel results in Islam (1995) show very strong positive correlation between measures of human capital and estimated values of A_0 . This provides a basis for the suggestion that the route along which human capital influences output may run through A_0 . In particular, it seems that human capital impacts output largely through its influence on the overall technological level.
53. Islam (1999) provides a recent review of international comparison of TFP levels across countries.
54. The finding of large total factor productivity differences across countries also helps explain the issue of capital flows that Lucas (1990) and others discuss. Bernard and Jones (1996a) also provide evidence on productivity differences. Maddison (1987) and

- Fagerberg (1994) provide excellent discussions of productivity differences across countries.
55. Dougherty and Jorgenson's second step analysis is limited to graphical treatment. Wolff, on the other hand, runs regression of subsequent TFP growth on initial TFP level.
 56. Dowrick and Nguyen try to distinguish between these two sources by including an interaction term in the regression. The interaction is between initial income variable and the average investment rate over the period. They conclude in favor of technological diffusion. The procedure, however, involves several simplifying assumptions.
 57. There is a large body of literature emphasizing this issue. See for example Abramovitz (1956) and Abramovitz and David (1973). Recently, Wolff (1991) has also drawn attention to this issue. Also related to this issue is the literature on the 'embodiment' hypothesis.
 58. De la Fuente (1996) presents an interesting decomposition of income-convergence among Spanish regions into its two sources, namely technological diffusion and capital deepening.
 59. The indices produced by the panel method contain ordinal as well as cardinal information, which can both be helpful in answering questions regarding TFP-convergence.
 60. For example, the specification based on the original version of the NCGT allows for only two variables, namely the investment rate and the labor force growth rate, assuming that g and δ are the same across countries. The specification based on the MRW augmented version of the NCGT allows for three variables.
 61. The discussion here is in the context of the NCGT model and the growth-convergence equation that is derived from that model. That does not mean that all growth-convergence equations have to proceed from this model. The general point is that no matter what model is used, the inclusion of variables needs to be decided by the model, and not on mechanical statistical grounds.
 62. Regarding heterogeneity of g see for example Romer's comment on Mankiw's paper in Mankiw (1995).
 63. For NONOIL, INTER, and OECD samples respectively the estimated value of this rate turns out to be 0.1845, 0.1521, and 0.1495, respectively. The correspondence between Lee *et al.*'s samples and MRW's NONOIL, INTER, and OECD samples is approximate.
 64. According to their Table-1, the coefficient of variation in this rate is 0.85, 0.66, and 0.36 for the 'Full,' 'Intermediate,' and 'OECD' samples respectively. These samples consist of 72, 58, and 20 countries, respectively.
 65. It may be noted that by itself, equation (15) does not pose a problem of endogeneity, because in this specification the right hand side variables are pre-determined. However, Chamberlain's MD estimation procedure uses both past and *future* values of x to substitute out μ_i (and y_{i0}) and hence requires strict exogeneity of x_{it} 's for the validity of estimation.
 66. This procedure eliminates μ_i by first differencing and then uses lagged values of y_{it} 's and x_{it} 's as instruments.
 67. This is because the results do not support the restriction that the coefficients of the s and n variable should be equal in magnitude but opposite in sign.
 68. This is because the estimated coefficient of the human capital variable turns out to be negative.
 69. See for example Alonso-Borrego and Arellano (1999), Kiviet (1995), Harris and Matyas (1996), Islam (2000), Jusdon and Owen (1997), and Ziliak (1997). However,

it needs to be mentioned that the instrument set used for the Arellano and Bond (1991) estimator in the Monte Carlo study is not exactly the same that Caseli *et al.* seems to have used.

70. Also the pooled OLS is not strictly a panel estimator, because it ignores the individual effect. Only those estimators that explicitly allow for individual effects and account for these effects in the estimation process can be properly called panel estimators. Sometimes, researchers have called their regressions as *panel* when, in fact, these are *pooled* regressions. Barro (1997) for example divides up the 1960–90 period into decades. He then pools the data and uses OLS. This is a pooled OLS, and not really an exercise in panel estimation. A pooled regression may also be conducted in a SURE multi-equation framework with the equations distinguished by the time period. This involves some specification of the error covariance matrix and therefore has some semblance to the (GLS) estimation under random-effects assumption. However, the error covariance structures in these two cases are not the same. Hence the SURE method that applies to data stacked by time periods and ignores the individual effect does not actually represent panel estimation.
71. Note that the asymptotic properties of panel estimators can be considered in the direction of $N \rightarrow \infty$ or $T \rightarrow \infty$ or both N and T going to infinity. Further, N and T can go to infinity either at the same rate or different rates. In fact, properties of panel estimators may differ depending on the direction in which asymptotics are considered. For example, Amemiya (1967, 1971) shows that, although LSDV is biased in the direction of $N \rightarrow \infty$, it is consistent in the direction of $T \rightarrow \infty$.
72. See for example Islam (2000). The dynamic panel data estimators considered in this study include: LSDV, two instrumental variable estimators by Anderson and Hsiao (1981, 1982), two GMM estimators by Arellano and Bond (1991), 2SLS, 3SLS, generalized 3SLS and Minimum Distance estimators by Chamberlain (1982, 1983). The Arellano-Bond GMM estimator, which depends heavily on lagged y 's as instruments, displayed large bias.
73. Barro (1997) and Durlauf and Quah (1998) have made these arguments.
74. This was shown in Islam (1995) by comparing the single cross-section results from a 25-year span and pooled OLS results obtained from five 5-year spans constituting the same 25-year period. The difference in results was negligible.
75. Some researchers have felt other problems than the ones discussed above. For example, Barro (1997) expresses concern about presence of auto-correlation induced by measurement errors. This problem can however be dealt with through use of such estimators as Chamberlain's MD estimator, which uses a heteroskedasticity and auto-correlation consistent variance-covariance matrix of the error term.
76. If we cannot reject the $H_0 : (1 + \beta) = 1$, then by implication we cannot reject the null $\beta = 0$, i.e., we cannot reject the hypothesis that there is no convergence.
77. Li and Papell (1999) offer a different definition of 'deterministic' and 'stochastic' convergence.
78. The exact number of rejections depends on the particular specification and variant of the Dickey-Fuller test. The simple DF test yields three rejections at 5 percent significance level. Noting that standard Dickey-Fuller tests have low power, Lee *et al.* also use a test proposed by Im, Pesaran and Shin (1995), referred to as the t-bar test, which is based on the average value of the DF statistics obtained across countries. The results remain basically unchanged. To account for potential serial correlation, they also adopt the augmented Dickey-Fuller specification with the number of lags determined by use of the SBC information criterion. This, however, does not change the results

that much. Based on the ADF t -statistics, the number of rejections of the unit root null ranges between 3 and 14, depending on the number of lags and whether data were demeaned or not. Lee *et al.* refer to allowing for 'common time specific effects.' (p. 21) It is unclear what is meant because apparently the procedure works on individual time series separately. The use of an analogous t -bar statistic for the ADF set-up, as proposed in Im *et al.*, does not affect these results by that much.

79. A non-rejection of the null therefore leaves a wide range of other non-unit root alternatives still compatible with the evidence.
80. With this test however the outcome depends on the degree of truncation.
81. The presence of $x_{i,t-1}$ in equation (15) is linked to the notion of *conditional* convergence when considered in an across-country setting. However, the notions of 'unconditional' and 'conditional' convergence have some relevance in the within economy context too. Equation (2) shows that the steady state of a particular economy can change over time if its determinants change over time. The term $x_{i,t-1}$ in equation (15) basically stands for these determinants, and the time subscript indicates the possibility of such changes. An analysis of 'within convergence' that allows for changes in the steady state over time can be thought to be an analysis of 'conditional convergence.' On the other hand if the analysis constrains $x_{i,t-1}$ to be time-invariant, then it may be termed as analysis of 'unconditional convergence.' Whether or not taking into account of changes in $x_{i,t-1}$ will affect the outcome of the unit root tests depends on the nature of changes and the time series properties of $x_{i,t-1}$ itself.
82. Lee *et al.* (1997) observe that Solow model does not have *internal explanation* for such changes. However, that does not mean that the model does not allow for the *possibility* of such changes.
83. For details, see Ben-David and Papell (1995, 1997), Ben-David, Lumsdaine, and Papell (1997), Lumsdaine and Papell (1997), and Zivot and Andrews (1992).
84. The regions are: New England, Mideast, Great Lakes, Plains, Southeast, Southwest, Rocky Mountains, and Far West.
85. They also impose a time series structure on the error term ε_t .
86. Carlino and Mills set the trend break exogeneously for the year 1946.
87. This makes their results directly comparable with Barro and Sala-i-Martin (1992)'s analogous results obtained from the cross-section approach.
88. However, as we noted earlier, Barro and Sala-i-Martin also include regional dummies and an index of composition of output in part of their analysis. Hence, it is a moot point whether their result is entirely of unconditional convergence.
89. For further discussion of issues involved with tests of unit root in panel data, see for example, Hall, Robertson, and Wickens (1993) and Hall and Urga (1995).
90. It is clear that unit root analysis of deviations is a type of cointegration analysis. This is obvious when deviations are taken from a reference country. When the deviations are taken sample average, the interpretation is not clear, as we noticed earlier.
91. They also consider two sub-samples: the first consisting of 11 European countries, and, the second, consisting of 6 European countries that show a high degree of 'pairwise cointegration.'
92. The basic idea, from Engle and Granger (1987), is that if the number of distinct stochastic trends in Y_t is less than n (which would imply co-integration), then the spectral density matrix at frequency zero of ΔY_t , i.e., of $f_{\Delta Y}(0)$, is not of full rank. If all n countries are converging in per capita output, then $f_{\Delta DY}(0)_i = 0, \forall i$, or equivalently, the rank of $f_{\Delta DY}(0)$ is zero. So, in operational terms, the task is to look at the spectral density matrix at frequency zero of ΔDY and ΔY , and check for the rank

conditions (more concretely, number of co-integrating vectors). For the 15-country sample, the US is the reference country. For the European sub-samples, the reference country is France.

93. Bernard and Durlauf were inclined to interpret the shared common trends as indicative of club convergence. However, this is not clear. We have noted earlier the difficulty in distinguishing evidence of club convergence from that of conditional convergence in general.
94. Drawing attention to this result, Lee *et al.* (1997) note that σ -convergence is not an implication of the Solow model.
95. Short of this adjustment, hypothesis of absence of σ -convergence will be rejected too often.
96. Note that the above relationships are based on a host of simplifying assumptions and ignoring, in particular, differences in the steady state. Litchenberg think that similar relationships, hold even when differences in steady state are allowed. However, there are a few issues here. First, the relationship involves, apart from conditioning on the left-hand side, both σ_u^2 and σ_ε^2 on the right hand. Second, it is derived on the assumption that the steady state is not time-contingent, so that \hat{y}_{it}^* is the same as $\hat{y}_{i,t-1}^*$. Third, the relationship above is in terms of income per effective worker; it will be more complicated when transformed in terms of income per capita. However, the basic ideas are evident from this algebra.
97. They also point out that the correct degrees of freedom is [(n-1), (n-1)] instead of [(n-2), (n-2)]
98. They denote original Litchenberg statistic by T_1 .
99. The adjustment will cause a type-I error that is lower than the significance level. One feature of these tests is that the comparison of variance is limited to that of first and last period only. In order to make use of the information in between, Carree and Klomp, proceeding from simplified version of equation (15), iterating backwards, and making use of the independence assumption, derive the following expression for the first difference of cross-section variance,

$$\Delta\sigma_t^2 = \tilde{\beta}^{2(t-1)}[(\tilde{\beta}^2 - 1)\sigma_0^2 + \sigma_\varepsilon^2]. \quad (24)$$

Under the null of no convergence, $(\tilde{\beta}^2 - 1) = 0$. Hence the equation above can be used to test this null by regressing $\Delta\sigma_t^2$ on $\tilde{\beta}^{2(t-1)}$ and using the $t(T-2)$ distribution, if $\tilde{\beta}$ is known. In practice, one has only an estimate of $\tilde{\beta}$ and the distribution will hold only approximately. This is their statistic, T_4 .

100. He obtains a slope coefficient of 0.715 and $R^2 = 0.802$. This gave $R^2/\hat{\beta}^2 = 1.57$ which is equal to $\text{var}(\ln GDP85)/\text{var}(\ln GDP60)$, the test statistic.
101. The probability value is 0.31.
102. When a later sample period of 1972–94 is considered, they however find that convergence does not hold any more. This, in their view, shows that while during initial years there has been significant reduction in variance, in the more recent years σ_t^2 has become close to σ_ε^2 and hence no further pronounced tendency for it to decrease is found. In other words, as the countries get closer to the steady states, the transition-component of the dynamics recede, and idiosyncratic shocks take over, which then displays no systematic tendency to decrease.
103. See for example, Sala-i-Martin (1996), Miller (1995).
104. Lee *et al.* (1997) point out that increased variance may also be the result of increasing dispersion in g .

105. Quah has produced a series of articles based on this line of research. These include Quah (1993b, 1996a, 1996b)
106. Quah also allows the one year transition matrix to be iterated 23 times and compares the resulting matrix (which he calls 'stationary estimate') with the 23-year transition matrix that is obtained from actual calibration of the data. He finds that the long run matrix shows stronger persistence than found in the 'stationary estimate.'
107. He fits a VAR model to forecast the quantiles, $Q(t)$, and then takes the convolution with M raised to the appropriate power to get the dynamic evolution of the sequence of distributions.
107. For example, the results are contingent on the arbitrary grid that is used to discretize the point in time empirical distributions. However, as Quah notes, inappropriate discretization may destroy the Markov property of an otherwise well behaved first order Markov process. Also, conclusions regarding piling up of probability mass are contingent on the choice of discretizing grid and may not be robust. (See Quah 1993a, p. 437.) Finally, Quah observes that the VAR models are estimated on the basis of only about 20 data points and hence are not that precise
108. Quah notes certain technical shortcomings of this analysis. For example, the results are contingent on the arbitrary grid that is used to discretize the empirical distributions. Quah notes that inappropriate discretization may destroy the Markov property of an otherwise well-behaved first order Markov process. Also, conclusions regarding piling up of probability mass is contingent on the choice of discretizing grid and may not be robust. (Quah 1993a, p. 437) Similarly, Quah observes that the VAR models are estimated on the basis of only twenty data points and hence are not that precise.
109. This contrasts with research on conditional β -convergence, which imposes structure and uses growth theory to decide on specification, choice of right hand side variables, etc. De la Fuente (1996) also draws attention to this feature of Quah's analysis.
110. From this point of view, convergence studies can be classified into two broad groups. One group, comprised mainly of the cross-section and the panel approach, imposes theoretical structure (derived from formal growth models) on the data and produces estimates of structural parameters of growth models. The other group, comprised mainly of the time series and the distribution approach, tends to avoid structure and thus resembles to reduced-form analysis of the output data.
111. Also as Solow (1994) notes, the externality-based NGT models suffer from the 'generalized Domar problem' requiring the social or aggregate elasticity of output with respect to capital to be exactly equal to one. While the estimated rates of conditional convergence have differed, the implied values of this elasticity have generally remained well short of being equal to one. As Chad Jones (1995a, 1995b) has shown, there are other ways to test for the validity of the new variants of the NGT.
112. For example, Jones and Manuelli (1990) relax the Inada conditions while preserving diminishing returns so that there can be both convergence and asymptotic growth. Ventura (1997), Duffy and Papageorgiou (2000), and others use the CES production function and show that for a certain range of the value of the elasticity of substitution between capital and labor, there can be long run growth. Tamura (1991) allows the strength of externality to depend on the level of human capital accumulation, and this leads to the convergence implication. Barro and Sala-i-Martin (1995) and Basu and Weil (1998) show how convergence can occur via technological diffusion. Ortigueira and Santos (1997) compare rates of convergence under alternative NGT models. Kocherlakota and Yi (1995) also discuss the problem of distinguishing between NCGT and NGT on the basis of convergence regressions.

113. Recent models of technological diffusion include Grossman and Helpman (1991), Parente and Prescott (1994a, 1994b), Barro and Sala-i-Martin (1997), Basu and Weil (1998), and Howitt (2000). As noted earlier, development economists and economic historians were aware all along about technological differences across countries, and many of them provided descriptive account of the technological diffusion process. Nelson (1968) provides an early formal model of technology diffusion. He notices that "the result that there are significant unexplained productivity differences across countries has been largely unnoticed," (p. 1223) and that "there are no reasons why (the assumption of) a common linear homogeneous production function... should be held sacrosanct." (p. 1229) In an earlier paper, Nelson and Phelps (1966) emphasized the role of human capital in technological diffusion. Many of the ideas currently used by the NGT models of technological diffusion can be found in these earlier works.
114. NGT models that allow role of trade in growth include Krugman (1990) and Grossman and Helpman (1991). For discussion of the role of trade in growth under the NCGT paradigm see Richard Baldwin (1993) and Mazumdar (1996). For an early discussion of the subject, see Robert Baldwin (1966). Ventura (1997) provides a recent elaborate integration of trade and growth. The issue of institutions has also gained importance. McGuire and Olson (1996) offer a model of interaction between institutions and growth. There has been significant development in collection of data on institutional quality across countries too, and these data have been employed to establish the importance of institutions in explaining cross-country growth regularities. See for example Knack and Keefer (1995).
115. The *American Economic Association (AEA)* has awarded prize to Alan Heston and Robert Summers for their work in putting together this data set.
116. See for example, Bardhan (1995).

References

- Abramovitz, Moses (1956) Resource and Output Trends in the US since 1870. *American Economic Review* 46, 5–23.
- Abramovitz, Moses and Paul David (1973) Reinterpreting American Economic Growth: Parables and Realities. *American Economic Review* 63, 428–437.
- Ades, Alberto and Hak B. Chua (1997) Thy Neighbor's Curse: Regional Instability and Economic Growth. *Journal of Economic Growth* 2, 279–304.
- Alonso-Borrego, Cesar and Manuel Arellano (1999) Symmetrically Normalized Instrumental-Variable Estimation Using Panel Data. *Journal of Business & Economic Statistics* 17, 36–49.
- Amemiya, T. (1967) A Note on the Estimation of Balestra-Nerlove Models, Technical Report No. 4, Institute for Mathematical Studies in Social Sciences, Stanford University.
- Amemiya, T. (1971) The Estimation of the Variance in a Variance-Component Model. *International Economic Review* 12, 1–13.
- Anderson, T. W. and C. Hsiao (1981) Estimation of Dynamic Models with Error Components. *Journal of American Statistical Association* 76, 598–606.
- Anderson, T.W. and C. Hsiao (1982) Formulation and Estimation of Dynamic Models Using Panel Data. *Journal of Econometrics* 18, 47–82.

- Arellano, Manuel and Stephen Bond (1991) Some Tests of Specification for Panel Data: Monte Carlo Evidence and an Application to Employment Equations. *Review of Economic Studies* 58, 277–297.
- Azariadis Costas and Allan Drazen (1990) Threshold Externalities in Economic Development. *Quarterly Journal of Economics* 105, 501–526.
- Auerbach, Alan J., Kevin A. Hassett, and Stephen D. Oliner (1994) Reassessing the Social Returns to Equipment Investment. *Quarterly Journal of Economics* 109, 789–802.
- Baldwin, Robert (1966) The Role of Capital-Goods in the Theory of International Trade. *American Economic Review* 56, 841–848.
- Baldwin, Richard E. (1992) Measurable Dynamic Gains from Trade. *Journal of Political Economy* 100, 162–74.
- Bardhan, Pranab (1995) The Contribution of Endogenous Growth Theory to the Analysis of Development Problems: An Assessment, in *Handbook of Development Economics*, Vol. III, edited by J. Behrman and T. N. Srinivasan, New York, Elsevier Science, pp. 2984–2998.
- Barro, Robert J. (1976) Rational Expectations and the Role of Monetary Policy. *Journal of Monetary Economics* 2, 1–32.
- Barro, Robert J. (1978) Unanticipated Money, Output, and the Price Level in the United States. *Journal of Political Economy* 86, 549–580.
- Barro, Robert J. (1990) Government Spending in a Simple Model of Endogenous Growth. *Journal of Political Economy* 98, 5, part II, S103–S125.
- Barro, Robert J. (1991) Economic Growth in a Cross Section of Countries. *Quarterly Journal of Economics* 106, 407–443.
- Barro, Robert J. (1997) *Determinants of Economic Growth*, Cambridge, MIT Press.
- Barro, Robert J. and Gary Becker (1989) Fertility Choice in a Model of Endogenous Growth. *Econometrica* 57, 481–501.
- Barro, Robert J. and Jong-Wha Lee (1993) International Comparisons of Educational Attainment. *Journal of Monetary Economics* 32, 363–94.
- Barro, Robert J., N. Gregory Mankiw and Xavier Sala-i-Martin (1995) Capital Mobility in Neoclassical Models of Growth. *American Economic Review* 85, 103–115.
- Barro, Robert J. and Xavier Sala-i-Martin (1992) Convergence. *Journal of Political Economy* 100, 223–51.
- Barro, Robert J. and Xavier Sala-i-Martin (1995) *Economic Growth*, McGraw Hill, New York, 1995.
- Barro, Robert J. and Xavier Sala-i-Martin (1997) Technological Diffusion, Convergence, and Growth. *Journal of Economic Growth* 2, 1–27.
- Basu, Susanto and David N. Weil (1998) Appropriate Technology and Growth. *Quarterly Journal of Economics* 113, 1024–1054.
- Baumol, William J. (1986) Productivity Growth, Convergence and Welfare: What the Long Run Data Show? *American Economic Review* 76, 1072–85.
- Baumol, William J. and Edward N. Wolff (1988) Productivity Growth, Convergence and Welfare: Reply. *American Economic Review* 78, 1155–1159.
- Bayomi, Tamim, David Coe, and Elhanan Helpman (1999) R&D Spillovers and Global Growth. *Journal of International Economics* 42, 399–428.
- Becker, Gary and Robert J. Barro (1988) A Reformulation of the Economic Theory of Fertility. *Quarterly Journal of Economics* 103, 1–25.
- Becker, Gary, Kevin M. Murphy and Robert Tamura (1990) Human Capital, Fertility and Economic Growth. *Journal of Political Economy* 98, S12–S37.
- Ben-David, Dan and David H. Papell (1995) The Great Wars, the Great Crash, and Steady State Growth: Some New Evidence about an Old Stylized Fact. *Journal of Monetary Economics* 36, 453–475.
- Ben-David, Dan and David H. Papell (1997) Slowdowns and Meltdowns: Post-War Growth Evidence from 74 Countries. *Review of Economics and Statistics* 80, 561–71.

- Ben-David, Dan, Robin Lumsdaine, and David Papell (1997) Unit Roots, Postwar Slowdowns, and Long-Run Growth: Evidence from Two Structural Breaks, Department of Economics, University of Houston, mimeo.
- Benhabib, Jess and Mark M. Spiegel (1994) The Role of Human Capital in Economic Development: Evidence from Aggregate Cross-Country Data. *Journal of Monetary Economics* 34, 143–173.
- Bernard, Andrew and Steven N. Durlauf (1996) Interpreting Tests of the Convergence Hypothesis. *Journal of Econometrics* 71, 161–173.
- Bernard, Andrew and Steven N. Durlauf (1995) Convergence in International Output. *Journal of Applied Econometrics* 10, 97–108.
- Bernard, Andrew and Charles I. Jones (1996) Technology and Convergence. *Economic Journal* 106, 1037–1044.
- Binder, Michael and M. Hashem Pesaran (1999) Stochastic Growth Models and Their Econometric Implications. *Journal of Economic Growth* 4, 139–183.
- Blomstrom, Magnus, Robert E. Lipsey, and Mario Zejan (1996) Is Fixed Investment the Key to Economic Growth? *Quarterly Journal of Economics* 111, 269–276.
- Canova, Fabio and Albert Marcet (1995) The Poor Stay Poor: Non-convergence across Countries and Regions, Discussion Paper No. 1265, Center for Economic Policy Research (CEPR), London.
- Carlino, G. A. and L. O. Mills (1993) Are the US Regional Incomes Converging? A Time Series Analysis. *Journal of Monetary Economics* 32, 335–346.
- Carree Martin and Luuk Klomp (1995) Testing the Convergence Hypothesis: A Comment. *Review of Economics and Statistics* 79, 683–86.
- Caselli, Francesco, Gerardo Esquivel and Fernando Lefort (1996) Reopening the Convergence Debate: A New Look at Cross Country Growth Empirics. *Journal of Economic Growth* 1, 363–89.
- Caselli, Francesco and Wilbur John Coleman II (2001) The US Structural Transformation and Regional Convergence: A Reinterpretation. *Journal of Political Economy* 109, pp. 584–616.
- Cass, David (1965) Optimum Growth in an Aggregative Model of Capital Accumulation. *Review of Economic Studies* 32, 233–40.
- Chamberlain, Gary (1982) Multivariate Regression Models for Panel Data. *Journal of Econometrics* 18, 5–46.
- Chamberlain, Gary (1983) Panel Data, in *Handbook of Econometrics*, Zvi Griliches, and Michael Intriligator, eds. (Amsterdam: North Holland), pp. 1247–1318.
- Chenery, Hollis, Sherman Robinson, and Moshe Sirquin (1985) *Industrialization and Growth: A Comparative Study*, New York, Oxford University Press.
- Cho, Dongchul and Stephen Graham (1996) The Other Side of Conditional Convergence. *Economics Letters* 50, 285–290.
- Christensen, Lauritis R., Dianne Cummings, and Dale W. Jorgenson (1981) Relative Productivity Levels, 1947–1973: An International Comparison. *European Economic Review* 16, 61–74.
- Chua, Hak B. (1992) Regional Spillovers and Economic Growth, Department of Economics, Harvard University, 1992.
- Coe, David T. and Elhanan Helpman (1995) International R&D Spillovers. *European Economic Review* 39, 859–887.
- Coe, David T., Elhanan Helpman, and Alexander W. Hoffmaister (1997) North-South and R&D Spillovers. *Economic Journal* 107, 134–149.
- De Gregorio, Jose (1992) Economic Growth in Latin America. *Journal of Development Economics* 39, 59–84.
- De Long, Bradford J. (1988) Productivity Growth, Convergence, and Welfare: A Comment. *American Economic Review* 78, 1138–54.
- De Long, Bradford J. and Lawrence Summers (1991) Equipment Investment and Economic Growth. *Quarterly Journal of Economics* 106, 445–502.

- De Long, Bradford J. and Lawrence Summers (1993) How Strongly Do Developing Economies Benefit from Equipment Investment? *Journal of Monetary Economics* 32, 395–415.
- De Long, Bradford J., and Lawrence H. Summers (1994) Equipment Investment and Economic Growth: Reply. *Quarterly Journal of Economics* 109, 803–807.
- De la Fuente, Angel (1996) On the Sources of Convergence: A Close Look at the Spanish Regions, Discussion Paper No. 1543, Center for Economic Policy Research (CEPR), London.
- De la Fuente, Angel (1997) The Empirics of Growth and Convergence: A Selective Review. *Journal of Economic Dynamics and Control* 21, 23–73.
- Den Haan, Wouter J. (1995) Convergence in Stochastic Growth Models: The Importance of Understanding Why Income Levels Differ. *Journal of Monetary Economics* 35, 65–82.
- Desdoigts, Alain (1999) Patterns of Economic Development and the Formation of Clubs. *Journal of Economic Growth* 4, 305–330.
- Dickey, D. and W. Fuller (1979) Distribution of the Estimators for Autoregressive Series with a Unit Root. *Journal of the American Statistical Association* 74, 427–431.
- Dickey, D. and W. Fuller (1981) Likelihood Ratio Tests for Autoregressive Time Series with a Unit Root. *Econometrica* 49, 1057–1072.
- Dollar, David and Edward Wolff (1994) Capital Intensity and TFP Convergence in Manufacturing, 1963–1985, in William J. Baumol, Richard R. Nelson, and Edward N. Wolff, eds., *Convergence of Productivity: Cross National Studies and Historical Evidence*. New York, Oxford University Press.
- Domar, Evsey (1946) Capital Expansion, Rate of Growth and Employment. *Econometrica* 14, 137–147.
- Dougherty, Chrys and Dale W. Jorgenson (1996) International Comparison of Sources of Growth. *American Economic Review* 86, 25–29.
- Dougherty, Chrys and Dale W. Jorgenson (1997) There is No Silver Bullet: Investment and Growth in the G7. *National Institute Economic Review* 162, 57–74.
- Dowrick, Steve and Duc-Tho Nguyen (1989) OECD Comparative Economic Growth 1950–85: Catch-Up and Convergence. *American Economic Review* 79, 1010–30.
- Duffy, John and Chris Papageorgiou (2000) A Cross-country Empirical Investigation of the Aggregate Production Function Specification. *Journal of Economic Growth* 5, 87–120.
- Durlauf, Steven N. and Danny T. Quah (1999) The New Empirics of Economic Growth, in John Taylor and Michael Woodford (eds.), *Handbook of Macroeconomics*, Vol. 1A, Amsterdam, North-Holland.
- Durlauf, Steven. and Paul A. Johnson (1995) Multiple Regimes and Cross-Country Growth Behavior. *Journal of Applied Econometrics* 10, 365–384.
- Engle, Robert and Clive Granger (1987) Cointegration and Error Correction Model. *Econometrica* 55, 251–76.
- Evans, Paul (1996) Using Cross-country Variances to Evaluate Growth Theories. *Journal of Economic Dynamics and Control* 20, 1027–1049.
- Evans, Paul and Georgios Karras (1996a) Convergence Revisited. *Journal of Monetary Economics* 37, 249–265.
- Evans, Paul and Georgios Karras (1996b) Do Economies Converge? Evidence from a Panel of US States. *Review of Economics and Statistics* 78, 384–388.
- Fagerberg, Jan (1994) Technology and International Differences in Growth Rates. *Journal of Economic Literature* 32, 1147–1175.
- Friedman Milton (1994) Do Old Fallacies Ever Die? *Journal of Economic Literature* 30, 2129–2132.
- Galor Oded (1996) Convergence? Inference from Theoretical Models. *Economic Journal* 106, 1056–1069.
- Gerschenkron, Alexander (1953) Economic Backwardness in Historical Perspective, in Bert F. Hoselitz (ed.) *The Progress in Underdeveloped Areas*, Chicago, Chicago University Press.

- Gregorio Jose De (1992) Economic Growth in Latin America. *Journal of Development Economics* 39, 59–84.
- Grier, Kevin B. and Gordon Tullock (1989) An Empirical Analysis of Cross-National Economic Growth, 1951–1980. *Journal of Monetary Economics* 24, 259–276.
- Grossman, Gene M. and Elhanan Helpman (1991) *Innovation and Growth in the Global Economy*. Cambridge, MIT Press.
- Hall, Robert E. and Charles I. Jones (1996) The Productivity of Nations, NBER Working Paper No. 5812, Cambridge, MA.
- Hall, Robert E. and Charles I. Jones (1997) Levels of Economic Activity across Countries. *American Economic Review* 87, 173–177.
- Hall, Robert E., and Charles I. Jones (1999) Why Do Some Countries Produce So Much More Output Than Others? *Quarterly Journal of Economics* 114, 83–116.
- Hall, Stephen and Giovanna Urga (1995) Stochastic Common Trends and Long-Run Relationships in Heterogeneous Panels, Discussion Paper No. 27–95, Center for Economic Forecasting, London Business School.
- Hall, Stephen, Donald Robertson, and Michael Wickens (1993) How to Measure Convergence with an Application to the EC Economies, Discussion Paper No. 19–93, Center for Economic Forecasting, London Business School.
- Harris, Mark N. and Laszlo Matyas (1996) A Comparative Analysis of Different Estimators for Dynamic Panel Data Models, Department of Economics, Monash University.
- Harrod, Roy (1993) An Essay in Dynamic Theory. *The Economic Journal* 49, 14–33.
- Holtz-Eakin, Douglas (1993) Solow and the States: Capital Accumulation, Productivity, and Economic Growth. *National Tax Journal* 46, 425–439.
- Howitt, Peter (2000) Endogenous Growth and Cross-country Income Differences. *American Economic Review* 90, 829–846.
- Im, Pesaran, and Shin (1995) Testing for Unit Roots in Heterogeneous Panels, Working Paper, Amalgamated Series 9526, Department of Applied Economics, University of Cambridge.
- Inada, Ken-Ichi (1963) On a Two-sector Model of Economic Growth: Comments and Generalization. *Review of Economic Studies* 30, 119–127.
- Islam, Nazrul (1995) Growth Empirics: A Panel Data Approach. *Quarterly Journal of Economics* 110, 1127–1170.
- Islam, Nazrul (1999) International Comparison of Total Factor Productivity: A Review. *Review of Income and Wealth* 45, 493–518.
- Islam, Nazrul (2000) Small Sample Performance of Dynamic Panel Data Estimators in Estimating the Growth-Convergence Equation: A Monte Carlo Study. *Advances in Econometrics* 15, 317–339.
- Johansen S. (1988) Statistical Analysis of Co-integration Vectors. *Journal of Economic Dynamics and Control* 12, 231–54.
- Jones, Larry E. and Rodolfo Manuelli (1990) A Convex Model of Equilibrium Growth: Theory and Policy Implications. *Journal of Political Economy* 98, 1008–1038.
- Jones, Charles I. (1995a) Time Series Tests of Endogenous Growth Models. *Quarterly Journal of Economics* 110, 495–525.
- Jones, Charles I. (1995b) R&D-Based Models of Economic Growth. *Journal of Political Economy* 103, 759–784.
- Jorgenson, Dale W. and M. Nishimizu (1978) US and Japanese Economic Growth, 1952–1974. *Economic Journal* 88, 707–726.
- Judson, Ruth A. and Ann L. Owen (1997) Estimating Dynamic Panel Data Models: A Practical Guide for Macroeconomists, Finance and Economic Series No. 3, Federal Reserve Bank, Washington, D.C.
- Kaldor, Nicholas (1971) Capital Accumulation and Economic Growth, in F. A. Lutz and D. C. Hague (eds.) *Theory of Capital*, New York: St. Martin's Press.

- Kelly, Morgan (1992) On Endogenous Growth with Productivity Shocks. *Journal of Monetary Economics* 30, 47–56.
- Kiviet, Jan F. (1995) On Bias, Inconsistency, and Efficiency of Various Estimators in Dynamic Panel Data Models. *Journal of Econometrics* 68, 53–78.
- Klenow, Peter J. and Andres Rodriguez-Clare (1997) *NBER Macroeconomics Annual 1997*, pp. 73–103, Cambridge and London: MIT Press.
- Knight Malcolm, Norman Loyaza and Delano Villanueva (1993) Testing for Neoclassical Theory of Economic Growth. *IMF Staff Papers* 40, 512–541.
- Knack, Stephen and Philip Keefer (1995) Institutions and Economic Performance: Cross-country Tests Using Alternative Institutional Measures. *Economics and Politics* 7, 207–227.
- Kocherlakota, Narayana R. and Kei-Mu Yi (1995) Can Convergence Regressions Distinguish between Exogenous and Endogenous Growth Models? *Economics Letters* 49, 211–215.
- Koopmans T.C. (1965) On the Concept of Optimal Economic Growth, in *The Economic Approach to Development Planning*, Pontifical Academy of Sciences, Amsterdam: North-Holland.
- Kormendi, Roger C. and Philip G. Meguire (1985) Macroeconomic Determinants of Growth: Cross-country Evidence. *Journal of Monetary Economics* 16, 141–163.
- Krueger, Anne O. (1968) Factor Endowments and Per Capita Income Differences among Countries. *Economic Journal* 78, 641–659.
- Krugman, Paul (1990) *Rethinking International Trade*, Cambridge: MIT Press.
- Kwiatkowski, Denis, Peter P. C. Phillips, Peter Schmidt, Yongcheol Shin (1992) Testing the Null Hypothesis of Stationarity against the Alternative of Unit Root. *Journal of Econometrics* 54, 159–178.
- Lee Kevin, M. Hashem Pesaran, and Ron Smith (1997) Growth and Convergence: A Multicountry Empirical Analysis of the Solow Growth Model. *Journal of Applied Econometrics* 12, 357–392.
- Levin, Andrew and Chien-Fu Lin (1993) Unit Root Tests in Panel Data, Dept. of Economics, University of California-San Diego.
- Levine, Ross and David Renelt (1992) A Sensitivity Analysis of Cross Country Growth Regressions. *American Economic Review* 82, 4, 942–963.
- Lichtenberg, Frank R. (1994) Testing the Convergence Hypothesis. *Review of Economics and Statistics* 76, 576–579.
- Lowey, Michael B. and David H. Papell (1996) Are US Regional Incomes Converging? Some Further Evidence. *Journal of Monetary Economics* 38, 587–598.
- Lucas Robert E. Jr. (1988) On the Mechanics of Economic Development. *Journal of Monetary Economics* 22, 3–42.
- Lucas Robert E. Jr. (1990) Why Doesn't Capital Flow from Rich to Poor Countries? *American Economic Review* 80, 92–6.
- Lucas Robert E. Jr. (1993) Making a Miracle. *Econometrica* 61, 251–272.
- Lusigi, Angela, Jenifer Piesse, Colin Thirtle (1998) Convergence of Per Capita Incomes and Agricultural Productivity in Africa. *Journal of International Development* 10, 105–115.
- Lumsdine, Robin and David Papell (1997) Multiple Trend Breaks and the Unit Root Hypothesis. *Review of Economics and Statistics* 79, 212–218.
- Maddison, Angus (1982) *Phases of Capitalist Development*, Oxford University Press, Oxford.
- Maddison, Angus (1987) Growth and Slowdown in Advanced Capitalist Economies: Techniques of Quantitative Assessment. *Journal of Economic Literature* 25, 649–698.
- Miller Ronald I. (1995) Time Series Estimation of Convergence Rates, Department of Economics, University of Columbia, 1995 (typescript).

- Mankiw, N. Gregory, Romer, David, and David Weil (1992) A contribution to the Empirics of Economic Growth. *Quarterly Journal of Economics* 107, 407–37.
- Mankiw, N. Gregory (1995) The Growth of Nations. 4 *Brookings Papers on Economic Activity* No. 1, 275–325.
- Mazumdar, Joy (1996) Do Static Gains from Trade Lead to Medium-Run Growth? *Journal of Political Economy* 104, 1328–1337.
- McGuire, Martin and Mancur Olson (1996) The Economics of Autocracy and Majority Rule: The Invisible Hand and the Use of Force. *Journal of Economic Literature* 34, 72–96.
- Nelson, Richard R. (1968) A Diffusion Model of International Productivity Differences in Manufacturing Industry. *American Economic Review* 58, 1219–1248.
- Nelson, Richard R. and Edmund S. Phelps (1966) Investment in Humans, Technological Diffusion, and Economic Growth. *American Economic Review* 56, 69–75.
- Nerlove Marc (1999) Properties of Alternative Estimators of Dynamic Models: An Empirical Analysis of Cross-country Data for the Study of Economic Growth, in Cheng Hsiao, Kajal Lahiri, Lung-Fei Lee, and Hashem Pesaran (eds.), *Analysis of Panels and Limited Dependent Variable Models*, Cambridge: Cambridge University Press.
- Nonneman, Walter and Patrick Vanhoudt (1996) A Further Augmentation of the Solow Model and the Empirics of Economic Growth for OECD Countries. *Quarterly Journal of Economics* 111, 943–53.
- Ortigueira, Salvador and Manuel S. Santos (1997) On the Speed of Convergence in Endogenous Growth Models. *American Economic Review* 87, 383–399.
- Parente, Stephen L. and Edward C. Prescott (1994a) Barriers to Technology Adoption and Development. *Journal of Political Economy* 102, 298–321.
- Parente, Stephen L. and Edward C. Prescott (1994b) Changes in Wealth of Nations, *Quarterly Review*, Federal Reserve Bank of Minneapolis, Spring, 3–16.
- Phillips, P. C. B. and S. Ouliaris (1988) Testing for Co-integration Using Principal Components Method. *Journal of Economic Dynamics and Control* 12, 205–230.
- Prescott, Edward C. (1998) Needed: A Theory of Total Factor Productivity. *International Economic Review* 39, 525–551.
- Li, Qing and David Papell (1999) Convergence of International Output: Time Series Evidence for 16 OECD Countries. *International Review of Economics and Finance* 8, 267–80.
- Quah, Danny (1990) International Patterns of Economic Growth: I. Persistence in Cross-country Disparities, Department of Economics, MIT, Cambridge.
- Quah, Danny (1993a) Galton's Fallacy and Tests of the Convergence Hypothesis. *Scandinavian Journal of Economics* 95, 427–443.
- Quah, Danny (1993b) Empirical Cross-Section Dynamics in Economic Growth. *European Economic Review* 37, 426–434.
- Quah, Danny (1996a) Empirics for Economic Growth and Convergence. *European Economic Review* 40, 1353–75.
- Quah, Danny (1996b) Twin Peaks: Growth and Convergence in Models of Distribution Dynamics. *Economic Journal* 106, 1045–1055.
- Quah, Danny (1996c) Convergence Empirics Across Economies with (Some) Capital Mobility. *Journal of Economic Growth* 1, 95–124.
- Rebelo, Sergio (1991) Long Run Policy Analysis and Long Run Growth. *Journal of Political Economy* 99, 500–21.
- Romer, Paul (1986) Increasing Returns and Long Run Growth. *Journal of Political Economy* 94, 1002–1036.
- Romer, Paul (1989a) Crazy Explanations for Productivity Slowdown, *Brookings Papers on Economic Activity*, Washington D. C., 163–202.

- Romer, Paul (1989b) Capital Accumulation in the Theory of Long Run Growth, in *Modern Business Cycle Theory*, Robert J. Barro, ed., Harvard University Press, Cambridge MA.
- Romer, Paul (1990) Endogenous Technological Change. *Journal of Political Economy* 98, S71-S102.
- Romer, Paul (1994) Origins of Endogeneous Growth. *Journal of Economic Perspectives* 8, 3-22.
- Sala-i-Martin, Xavier (1996a) The Classical Approach to Convergence Analysis. *Economic Journal* 106, 1019-1036.
- Sala-i-Martin, Xavier (1996b) Regional Cohesion: Evidence and Theories of Regional Growth and Convergence. *European Economic Review* 40, 1325-1352.
- Sala-i-Martin, Xavier (1997) I Just Ran Two Million Regressions. *American Economic Review* 87, 178-183.
- Shioji, Etsuro (1996) Regional Growth in Japan, Discussion Paper No. 1425, Center for Economic Policy Research.
- Solow, Robert M. (1956) A Contribution to the Theory of Economic Growth. *Quarterly Journal of Economics* 70, 65-94.
- Solow, Robert M. (1970) *Growth Theory: An Exposition*, Cambridge University Press, London.
- Solow, Robert M. (1994) Perspectives on Growth Theory. *Journal of Economic Perspectives* 8, 45-54.
- Summers, Robert, and Alan Heston (1988) A New Set of International Comparisons of Real Product and Price Levels Estimates for 130 Countries. 1950-85. *Review of Income and Wealth* 34, 1-26.
- Summers, Robert, and Alan Heston (1991) The Penn World Table (Mark 5): An Expanded Set of International Comparisons, 1950-1988. *Quarterly Journal of Economics* 106, 327-368.
- Swan, Trevor W. (1956) Economic Growth and Capital Accumulation. *Economic Record* 32, 334-361.
- Tamura, Robert (1991) Income Convergence in an Endogenous Growth Model. *Journal of Political Economy* 99, 522-540.
- Temple, Jonathan (1998) Equipment Investment and the Solow Model. *Oxford Economic Papers* 50, 39-60.
- Temple, Jonathan (1998) Robustness Tests of the Augmented Solow Model. *Journal of Applied Econometrics* 13, 361-75.
- Temple, Jonathan (1999) The New Growth Evidence. *Journal of Economic Literature* 37, 112-156.
- Temple, Jonathan and Paul A. Johnson (1998) Social Capability and Economic Growth. *Quarterly Journal of Economics* 113, 965-90.
- Temple, Jonathan and Hans-Joachim Voth (1998) Human Capital, Equipment Investment, and Industrialization. *European Economic Review* 42, 1343-1362.
- Tzanidakis, George and Theo Kirizidis (1995) A Test of a Modern Version of the Solow Model. *Economics Letters* 3, 587-590.
- Verspagen, Bart (1992) Endogenous Innovation in Neo-classical Growth Models: A Survey. *Journal of Macroeconomics* 14, 631-662.
- Ventura, Jaume (1997) Growth and Interdependence. *Quarterly Journal of Economics* 112, 57-84.
- Wolff, E. N. (1991) Capital Formation and Productivity Convergence. *American Economic Review* 81, 565-579.
- Young, Alwyn (1995) The Tyranny of Numbers: Confronting the Statistical Realities of the East Asian Growth Experience. *Quarterly Journal of Economics* 110, 641-680.

Zivot, Eric and Donald Andrews (1992) Further Evidence on the Great Crash, the Oil Price Shock, and the Unit Root Hypothesis. *Journal of Business and Economic Statistics* 10, 251–270.