Galton's Fallacy and Tests of the Convergence Hypothesis

by

Danny Quah*
LSE Economics Department
December 1993

* This paper is substantially revised from an earlier version with the same title. I thank many colleagues, especially Milton Friedman, Marc Nerlove, and the LSE Macro Breakfast Group, for helpful comments. I also gratefully acknowledge the hospitality of the Institute for Empirical Macroeconomics at FRB Minneapolis. All data-derived graphs and calculations contained here were executed using the author's time-series, random-fields econometrics shell tSrF.

Galton's Fallacy and Tests of the Convergence Hypothesis
by
Danny Quah
LSE Economics Department
December 1993

ABSTRACT

Recent tests for the convergence hypothesis derive from regressing average growth rates on initial levels: a negative initial level coefficient is interpreted as convergence. These tests turn out to be plagued by Galton's classical fallacy of regression towards the mean. Using a dynamic version of Galton's fallacy, I establish that coefficients of arbitrary signs in such regressions are consistent with an unchanging cross-section distribution of incomes. Alternative, more direct empirics used here show a tendency for divergence, rather than convergence, of cross-country incomes.

Communications to: D. T. Quah, LSE, Houghton Street, London WC2A 2AE.
[Tel: +44-171-955-7535, Email: `dquah@lse.ac.uk`]
[(URL) `http://econ.lse.ac.uk/~dquah/`]

Nontechnical Summary

What causes some countries to grow rapidly, others only slowly? Do some regions stagnate while others boom? Is there a tendency for poorer countries to catch up with richer ones? (That the last might be true is also known as the "convergence hypothesis".) Questions such as these have motivated both the new growth theory and recent empirical studies of cross-country growth. This paper shows that the usual interpretations given in those empirical studies suffer from a classical fallacy in regression theory. Finding that the countries initially poorer also happen to grow faster—using standard regression techniques—turns out to be uninformative on the convergence hypothesis. The paper then also presents new empirical evidence showing that in the post War era rich countries tend to become richer, poor countries poorer, while the middle-income group of countries appears to be an (eventually) vanishing class.

Keywords: Cross-country growth, convergence, Galton's fallacy, regression towards the mean, transition matrix, stochastic kernel

JEL Classification: C10, C22, C23, E17, O40

Communications to: D. Quah, S479 LSE Economics Department, Houghton Street, London WC2A 2AE [Tel: (+44-71) 955-7535; Fax: (+44-71) 831-1840; Email: dquah@lse.ac.uk]

Feathers hit the ground before their weight can leave the air.

R.E.M.

I. Introduction

Do the incomes or productivity levels of different economies have a tendency to converge? Numerous researchers have recently examined this question by calculating cross-section regressions of measured growth rates on initial levels. See for instance Barro (1991), Barro and Sala-i-Martin (1991, 1992), Baumol (1986), DeLong (1988), Dowrick and Nguyen (1989), Mankiw, Romer, and Weil (1992), and many others. Such regressions have, in fact, become known as "Barro regressions," and are widely used to analyze empirical growth dynamics. Evidently, in a Barro regression, a negative coefficient on initial levels is taken to indicate convergence.

This paper clarifies what such initial level regressions are able to uncover. As used in this literature, the term *convergence* can mean a number of different things:

(a) Countries originally richer than average are more likely to turn below average eventually, and vice versa—the cycle repeats;
(b) Whether a country income is eventually above or below average is independent of that economy's original position;
(c) Income disparities between countries have neither unit roots nor deterministic time trends; and
(d) Each country eventually becomes as rich as all the others—the cross-section dispersion diminishes over time.

Cases (a) and (b) vaguely correspond to the notions of *mixing* and *ergodicity* in econometrics (see e.g. White (1984)). Case (c) is one formulation of persistence in income disparities: from a time-series perspective, it is the natural way to examine dependence on initial conditions. This particular probability model raises interesting econometric issues in the context of unit root random fields (see Quah (1992a)); it is, however, quite different in spirit and in substance from initial level regressions. Case (d) is closest to the notion of poorer countries eventually catching up with richer countries.

If (a) and (b) are the cases of interest, then models for studying transitional characteristics—for example, that used in the income distribution and earnings mobility literature—would seem appropriate. Thus, Quah (1992b) attempts to uncover such effects in the context of heterogeneous Markov chains. Overall, however, the work using initial levels regressions strongly suggests case (d) as being of interest.

This paper first shows, in Section 2, that the widely-used initial level regressions described above, in fact, shed no light on convergence in the sense of (d). It does this by developing an analogy between such regressions and Francis Galton's classical fallacy of regression towards the mean.[1] Recall that Galton, in aristocratic manner, was concerned about the sons of tall fathers regressing into a pool of mediocrity along with the sons of everyone else—he inferred this from observing that taller-than-average fathers had sons who turned out to be not as much above average as the fathers themselves. He could not, however, reconcile this with the population of male heights continuing to display significant cross-section dispersion. I show—using exactly the same reasoning which resolves that paradox—that a negative cross-section regression coefficient on initial levels is, in fact, perfectly consistent with the absence of convergence in the sense of (d).

Barro (1991), Barro and Sala-i-Martin (1991), and Mankiw, Romer, and Weil (1992) have recently shown that what they call "conditional convergence" occurs. Some of these authors further claim that this conditional convergence happens at approximately the same rate over different time periods and across different cross-sectional samples. The results described in the previous paragraph extend—in a straightforward way—to cover these cases of conditional convergence: simply apply the argument to the *residuals* of output growth, after conditioning on exogenous

[1] After completing the first draft of this paper, I came across Friedman (1992), which makes much the same point as this. The discussion in Section 3 below, however, appears nowhere else that I have seen.

variables of interest.

While Galton's formulation is convenient for analyzing observations at two points in time, it offers little by way of interesting dynamics. Extending the analysis to permit such dynamics, I show, again in Section 2, that a given cross-section distribution—replicating itself over time—is consistent with *arbitrary* signs on the cross-section initial levels regression coefficient. In other words, the sign of the initial levels regression coefficient says nothing about whether there is convergence or divergence.

Section 3 describes an alternative, more direct means for examining the convergence hypothesis, using a model of a cross-country distribution of incomes, evolving dynamically. The results here show that while one can uncover pockets of convergence, the data overall show divergence instead. The picture that emerges is one of a world where countries tend—in the long run—towards either the very rich or very poor, with the middle income classes disappearing. The disparity between the rich and poor, further, appears to be widening. Finally, Section 4 briefly concludes.

II. Galton's Fallacy Dynamics

I will argue here that calculating a cross-section regression to explain time-averaged growth rates gives an example of why it is inappropriate to attempt to draw dynamic implications from cross-section evidence. The earliest example of this that I know is the famous Galton's fallacy of regression towards the mean. In this section I make that connection explicit.

Consider Figure 2.1: such a graph is certainly not original with the current paper. It is arguably one of the original motivations for the entire endogenous growth literature, and has appeared in publications ranging from Paul Romer's academic papers to magazines like The Economist.

Figure 2.1 shows a wide dispersion of average growth rates for economies having low initial income levels. (The data source and a listing of the different economies studied here are given in the Data Appendix.) There is little sign of a negative correlation—in the cross-section—between initial conditions and time-averaged growth rates. But why might one expect such a negative correlation?

Suppose that after controlling for appropriate exogenous differences—human capital, government policy, natural resources—different economies are hypothesized to have the same long-run income level, or, alternatively, the same growth path. Then in a picture like Figure 2.1—again, after controlling for exogenous variables—should not there be a negative correlation in the cross section, so that countries initially starting lower catch up with or converge to those countries already ahead? Controlling for exogenous differences might be as simple as just looking at Figure 2.1, but using the residuals from some first-stage regression.

An alternative representation of this convergence idea is given in Figure 2.2: Such a graph shows a collapsing of the cross-section distribution—possibly after controlling for exogenous variables—about an underlying growth path that happens to be unique across countries. It seems intuitive that this dynamic characteristic should manifest in a negative correlation of time-averaged growth rates with initial conditions (such as Figure 2.1 ought to show).[2]

That such a negative correlation does not, in fact, imply a collapsing of the cross-section distribution is easily seen from the reasoning in Galton's fallacy. (I repeat: the argument here applies whether one takes just the raw data or the residuals from a first stage regression, i.e., controls for exogenous conditioning. Thus, the distinction between conditional and unconditional convergence is irrelevant here.) Suppose, contrary to Figure 2.2, that the cross-section distribution is not collapsing, but replicates itself because it happens to be the stationary distribution for many independent and identically-distributed country incomes $Y_j(t)$. For t_1 and t_2 arbitrary points in time, the cross-section regression of $Y(t_2)$ on a

[2] This simple idea resurfaces in many current policy debates—among them, whether East Germany is catching up to West Germany, whether Europe is becoming a two-track system (one for the rich, another for the poor), and whether the US poor are becoming poorer. See for instance Zimmerman (1992).

constant and $Y(t_1)$ is just:

$$P\left[\,Y(t_2) \mid 1, Y(t_1)\right] = E_C Y(t_2) + \lambda\left(\,Y(t_1) - E_C Y(t_1)\,\right),$$

where

$$\lambda = \mathrm{Var}_C^{-1}\left(Y(t_1)\right) \cdot \mathrm{Cov}_C\left(Y(t_2), Y(t_1)\right),$$

the C subscript denoting *cross-section*. The assumption that the cross-sections are in stationary state gives:

$$\mathrm{Var}_C(Y(t_1)) = \mathrm{Var}_C(Y(t_2)).$$

The Cauchy-Schwarz inequality,

$$\mathrm{Cov}_C(Y(t_2), Y(t_1)) \leq \mathrm{Var}_C^{1/2}(Y(t_2))\mathrm{Var}_C^{1/2}(Y(t_1)),$$

then implies that $\lambda \leq 1$, so that:

$$P\left[\,Y(t_2) - Y(t_1) \mid 1, Y(t_1)\right] = \mu - (1-\lambda)Y(t_1), \tag{$\dagger$}$$

with the coefficient on $Y(t_1)$ necessarily nonpositive.

If $t_1 < t_2$ then ($\dagger$) is a regression of (scaled) growth rates on an initial condition: the conclusion above then implies that the coefficient on the initial condition is always no greater than zero, even when the cross-section distribution remains invariant over time. A negative sign on the initial-condition coefficient, therefore, does not indicate a collapsing cross-section distribution. This, of course, is simply Galton's fallacy of regression towards the mean.

As formulated here, the argument takes the extreme assumption that the cross-section distribution is invariant. This makes the point particularly clear, but there is no real need for that time invariance. The coefficient in this initial-conditions regression depends only on λ, or equivalently, the relation between $\mathrm{Cov}_C(Y(t_2), Y(t_1))$ and $\mathrm{Var}_C(Y(t_2))$ and $\mathrm{Var}_C(Y(t_1))$. Therefore, the "initial-condition coefficient" can be negative—λ strictly less than 1—even if $\mathrm{Var}_C(Y(t_2))$

exceeds $\mathrm{Var}_C(Y(t_1))$—the cross-section distribution can diverge even when the initial conditions regression shows a negative correlation between time-averaged growth rates and initial levels! Further, nothing in the argument relies on t_2 exceeding t_1: all the same (negative) conclusions follow, even when the researcher is estimating a "final-condition" (or orrab) regression or a "middle-condition" regression. Figures 2.3 and 2.4 show that the message in Figure 2.1 is unchanged by replacing the initial conditions with income levels at other arbitrary dates. This kind of cross-section regression is thus completely uninformative for the dynamics of the distribution.

Galton's Fallacy reasoning, in discussing the empirics of growth, is sometimes phrased as follows: a negative initial condition provides a force for the cross-section distribution to collapse; on the other hand, ongoing disturbances provide a force in the opposite direction.[3] This phrasing is, however, specious. To see why, consider starting out the cross-section distribution with all countries at the same point. Then, over time, the distribution must tend towards its stationary version. Provided that the countries are not perfectly correlated, there must be a *divergence* from the initial distribution. All this happens with neither a change in the underlying time-series representation for each country, nor a change in the covariance properties across countries. Thus, any regression calculated from data on these countries must remain unchanged, independent of whether the cross-section is collapsing or diverging.

If regression (†) is perturbed so that the initial condition appearing on the right is not that used to calculate the average growth rate on the left-hand side, then we have:

$$P\left[Y(t_2) - Y(t_0) \mid 1, Y(t_1)\right] = \mu' - (\nu - \lambda)Y(t_1),$$

where:

$$\nu = \mathrm{Var}_C^{-1}\left(Y(t_1)\right) \cdot \mathrm{Cov}_C\left(Y(t_0), Y(t_1)\right).$$

[3] These different forces underly Barro and Sala-i-Martin (1991, 1992)'s distinction between β- and σ-convergence.

All the same observations made earlier about λ now apply equally to ν: the sign of $\lambda - \nu$ depends on $\mathrm{Cov}_C\,(Y(t_0), Y(t_1))$ and $\mathrm{Cov}_C\,(Y(t_2), Y(t_1))$ in addition to the variances of $Y(t_2)$, $Y(t_1)$, and $Y(t_0)$. No sensible inference can thus be drawn on the dynamics of the cross-section distribution by the cross-section regressions that are typically estimated. In particular, perturbing the initial condition gives no more information on the convergence properties of the distributions over time.

The strong independence and identical distribution assumption I used above plays a role only in simplifying the calculations. With heterogeneity, any time-invariant cross-section distribution is a probability mixture of the different individual time series ergodic distributions. Weak forms of dependence across countries will not affect their cross-section distribution being approached by the empirical (really just a version of the Glivenko-Cantelli law). With strong dependence or small numbers of countries, the empirical cross-section distribution is a non-degenerate random element in the space of distributions. While the calculations then become much more difficult, the flavor of the results is unaffected.

It is useful, therefore, to eschew this kind of analysis, and seek a better way to model the dynamics of this large cross-section. Figure 2.5 shows a three-dimensional plot of the natural logarithm of per capita country incomes, arrayed in both time and cross-sectional unit: this graphs the raw data behind the calculations previously undertaken in Figures 2.1, 2.3, and 2.4. While the ordering of these data along the time axis is obvious—time proceeds naturally sequentially—that along the country axis is not. Here, I have arbitrarily taken the ordering given in World Bank country codes—reported in, among other places, Summers and Heston (1991).

This graph shows the complications for dynamic inference—are countries converging over time? diverging? growing in a way depending on size?—off of this data set. The variation in the data is as rich in the cross-section dimension as it is in the time-series. Growth rate regressions, averaging out the time dimension, are simply one—and it turns out, misleading—way to analyze such a data structure. Standard econometric methods are not well-suited to working with such data structures.

III. Direct Tests of Convergence

An alternative, more transparent way to see whether convergence occurs is to examine directly the cross-section distributions of output per worker over time. Because the entire world might be growing, and because the analysis is most naturally performed abstracting from this world-wide comovement, I calculated output per worker for each country, divided by the same figure for the world. (Thus, a number like 2 below indicates twice the world average, and so on.) This normalization is an easy—but imperfect—way to abstract from world-wide growth and fluctuations; a more accurate, but also more difficult, normalization can be obtained using methods in Quah and Sargent (1993). Figures 3.1 (a)–(d) show point-in-time (smoothed) estimates of the densities of the cross-country distribution of this normalized series: (a)–(c) are for 1962, 1974, and 1985 respectively, while (d) is taken over the entire sample 1962–1985.[4]

The key message from this sequence of graphs is that the cross-section distributions do not appear to be collapsing. Instead, they seem to be fluctuating over time. Filling in the gaps from 1963–1973 and 1975–1984 with additional graphs of the densities for each of those years would simply reinforce that idea.

Two natural questions arise here: First, recall expositions of the neoclassical growth model, as e.g. in Mankiw, Romer, and Weil (1992). Those make clear that a nondegenerate distribution as in these graphs could well be consistent with that growth model in steady state—see in particular equation (6) of Mankiw, Romer, and Weil (1992) and the surrounding discussion. Those models, however, cannot produce a diverging distribution over time.[5] An informative examination of this

[4] The estimates were obtained using a gaussian kernel with bandwidth selected automatically, as suggested in Silverman (1986) 3.4.2; the fast fourier transform was used to calculate the resulting kernel estimator. A reflection method (Silverman (1986), 2.10) took into account nonnegativity in the original data.

[5] In this discussion, I have absorbed the cross-country conditioning variables

prediction therefore should examine the behavior of the cross-section distribution in steady state. But which of Figures 3.1 (a)–(d) is steady state?

Second, confronted with the representations of the data in Figures 3.1 (a)–(d), a researcher is tempted to ask if the "bump" in the density between 2 and 4 in 1985 contains the same countries as that in 1974. Graphs such as these cannot tell us that.

Both questions can be answered, however, by developing a probability model of transitions—law of motion—for the distributions in Figures 3.1 (a)–(d) that can, at the same time, be used to generate a characterization of the steady state. Quah (1993b) provides exactly such a model: he uses discrete Markov chains to approximate and estimate a law of motion for the evolving distributions. While that work has not yet been able to give any measures of precision, the results suggest two interesting empirical characteristics: (1) the ergodic distribution of these cross-country incomes is one where many countries are rich and many are poor, with the middle-income countries a vanishing class; and (2) intra-distribution mobility is nontrivial—over sufficiently long periods of time, output per worker wanders outside of any fixed neighborhood. These results appear at odds with the predictions of the standard neoclassical growth model, as described in e.g. Mankiw, Romer, and Weil (1992).

But those conclusions on the ergodic distribution bear an important qualification: they are conditional on the arbitrary grid that was used to discretize the point-in-time empirical distributions. From Markov chain theory, we know that a continuous first-order Markov process (otherwise well-behaved) need not even retain the Markov property when inappropriately discretized. More directly, in a discrete setting, statements to the effect that probability mass piles up in a certain part of the distribution *have* to be related to the choice of the discretizing grid. It

---

of Mankiw, Romer, and Weil into unobservable country-specific processes. Quah (1993a) is a study in the same spirit as the current one, but explicitly considers conditioning information. A brief discussion of the results there will appear below.

is important, therefore, to check the robustness of those conclusions.

In the current work, I explore a different representation for the dynamics in the evolving distributions of Figure 3.1. Fix a set of increasing, nonredundant probabilities $P = \{p_1, p_2, \ldots, p_n\}$. For income distribution $\mathcal{F}_t$ in each period t, the set P determines a corresponding set of quantiles $Q(t) = \{q_1(t), q_2(t), \ldots, q_n(t)\}$, where, by definition, the quantiles and probabilities are related by $p_k = \mathcal{F}_t\,(q_k(t))$, for $k = 1, \ldots, n$. When P is chosen to be equally spaced on the open unit interval $(0, 1)$, each quantile-set pair $\{\, Q(t), Q(t+1) \,\}$ defines an $(n+1) \times (n+1)$ *fractile transition probability matrix* $M(t)$ of transitions from $\mathcal{F}(t)$ to $\mathcal{F}(t+1)$. (The terminology *fractile Markov chain* is from Geweke, Marshall, and Zarkin (1986); it is natural also to call fractile the transition probability matrix of such a Markov chain.) Notice that given probabilities P, the fractile matrix $M(t)$ is uniquely determined even if the quantiles Q are not. If the fractile matrix M is assumed to be time-invariant, then it can be estimated by averaging across both time and countries. Further, if the quantile sets $Q(t)$ happen to be time-invariant then the representation here collapses to a version of that used by Quah (1993b).

In the subsequent analysis, I impose time-invariance in M, but leave $\{\, Q(t) :$ integer $t \,\}$ as a serially correlated vector process. Therefore, the work here generalizes that in Quah (1993b), and serves as a check on the robustness of conclusions there. The dynamic evolution of the sequence of income distributions can be obtained by first forecasting Q, and then taking the convolution with M raised to the appropriate power. Because M is a fractile matrix, we immediately know that there is an ergodic distribution, and further that that ergodic distribution is uniform relative to the quantiles Q (implied by P). The corresponding ergodic distribution on the original state space is then found by combining the preceding fact with knowledge of the sequence of quantile sets $\{\, Q(t) : \text{integer } t \,\}$. In words, the transition matrix M parametrizes intradistribution mobility, while the vector sequence Q parametrizes movements in the entire distribution.

I have estimated four- and five-state fractile Markov chain models for the cross-country income distributions, again taken relative to world average.[6] Tables 3.1 and 3.2 give the corresponding transition probabilities M. The first feature that strikes one is the extreme immobility over time apparent in these transition probabilities. The very richest remain so with probability at least 98%; the very poorest, with probability at least 95%. The interior diagonal entries in both tables are lower than the extreme diagonal endpoints: middle-income economies are less likely to remain where they are in the distribution. From the tables, we see that those middle-income economies face about equal probabilities of rising or falling. Notice that as estimated, both M's imply that the corresponding fractile Markov chains are irreducible and have all states ergodic. Thus, the long run, steady state distributions are unique, and by the previous discussion, are also uniform (relative to Q).

To complete the description of these transition dynamics, we now consider the behavior of the quantile sets $\{\, Q(t) : \text{integer } t\,\}$. Over the estimation period 1962–1984, these varied in quite suggestive ways. With $n = 4$, the median q_2 wandered between 0.61 in 1962 (half the world's economies had per capita incomes no greater than 61% of the world average) to 0.70 in 1975, and then 0.65 in 1984. During this period, q_2 never rose above 0.76 (1978) nor fell below 0.60 (1964). By contrast, the 0.25 quantile q_1 drifted downwards from an average of 0.36 over 1962–4 to 0.27 over 1982–4; the 0.75 quantile q_3 upwards from 1.29 to 1.50. Similar behavior (naturally) manifests with $n = 5$: the 0.2 quantile q_1 drifted downwards from an average of 0.31 over 1962–4 to 0.23 over 1982–4; the 0.8 quantile q_4 upwards from 1.57 to 2.04. In words, the median roughly stays put while the upper quantiles wander upwards, and the lower quantiles downwards. These movements at the extremes are, for the most part, monotone from year to year over 1962–84: the preceding description of time averages at the beginning and end points of the sample are thus representative of general tendencies.[7]

---

[6] The four-state specification, of course, has $P = \{0.25, 0.50, 0.75\}$, and similarly the five-state specification, $P = \{0.2, 0.4, 0.6, 0.8\}$.

[7] I have also experimented with VAR models for Q, and then forecasting from

Putting together these observations on the dynamics of Q and the immobility patterns in M, we come to much the same conclusions as Quah (1993b). The cross-country income distributions evolve over time in a way that shows a tendency for the rich remaining rich and the poor, poor, and for the gap between rich and poor to widen. I do not present the calculations here, but it should be clear from the structure of the transition probabilities in Tables 3.1 and 3.2 that iterating those transition probabilities shows the same effects as uncovered in the ergodic distributions in Quah (1993b): in the long run, there is a piling up of probability mass in the tails, and a thinning out in the middle.

From the discussion of Section 2 above, the usual cross-country growth regressions shed little light on the dynamics just described in the previous paragraph. Thus, the evidence on "convergence" that different researchers have uncovered do not really bear on our empirical conclusions. Instead, the empirical results here simply highlight the error in interpreting previous findings as having demonstrated convergence.

One important aspect that other researchers have been able to investigate with cross-section regressions, and that I have not here, is the influence of conditioning information: savings, schooling, political stability, and so on. Quah (1993a), in a different but related setting as that used above, has studied the effects of conditioning information. That work concludes that much the same divergence and immobility results as obtained here remain, even after conditioning.

---

them to get an idea of future tendencies. The time sample for estimation comprises only about 20 points, however, so that not very much of interest was precisely estimated. The exercise nevertheless did confirm the drift tendencies in the extreme quantiles.

IV. Conclusion

The main goal of this paper was to criticize standard cross-section regression tests of the convergence hypothesis. Section 2 did that, in some detail. By drawing an analogy of these regression tests with those of the classical Galton's fallacy, and then extending the reasoning for dynamics, I have shown why these cross-section tests are misleading for the hypothesis of interest.

The alternative, more direct tests described in this paper provide evidence against the convergence hypothesis. They show, instead, a world with economies tending—in the long run—towards either the very rich or very poor, with the middle-income classes vanishing. The rich-poor income disparity, in addition, appears to be widening.

What are the next steps in this alternative, empirical research program, one that escapes the strictures of the classical Galton's fallacy? First, the lessons here call for theoretical growth models that generate predictions for the dynamics of the entire cross-economy distribution, not just for those of a single economy. Second, there is scope for studying time-heterogeneity in these evolving distributions, in ways that are richer than simply breaks in trend or unit-root nonstationarity. (Quah (1992b) is a first step.) Third, many of the issues raised here extend naturally to empirical studies of regional and geographical dynamics and of the dynamics of large cross-sections of commodity and asset prices. Such investigations are currently under way.

References

Barro, Robert J. (1991) "Economic growth in a cross-section of countries." *Quarterly Journal of Economics* 106(2), 407–443, May

Barro, Robert J., and Xavier Sala-i-Martin (1991) "Convergence across states and regions." *Brookings Papers on Economic Activity* (1), 107–182, April

——— (1992) "Convergence." *Journal of Political Economy* 100(2), 223–251, April

Baumol, William J. (1986) "Productivity growth, convergence, and welfare." *American Economic Review* 76(5), 1072–85, December

DeLong, J. Bradford (1988) "Productivity growth, convergence, and welfare: A comment." *American Economic Review* 78(5), 1138–55, December

Den Haan, Wouter J. (1992) "Convergence in stochastic models." Working Paper, Economics Dept., UCSD, July

Dowrick, S., and D. Nguyen (1989) "OECD comparative economic growth 1950–85: Catch-up and convergence." *American Economic Review* 79(5), 1010–1030, December

Friedman, Milton (1992) "Do old fallacies ever die?" *Journal of Economic Literature* 30(4), 2129–2132, December

Geweke, John, Robert C. Marshall, and Gary A. Zarkin (1986) "Mobility indices in continuous time Markov chains." *Econometrica* 54(6), 1407–1423, November

Mankiw, N. Gregory, David Romer, and David N. Weil (1992) "A contribution to the empirics of economic growth." *Quarterly Journal of Economics* (429), 407–437, May

Quah, Danny (1992a) "International patterns of growth: I. Persistence in cross-country disparities." Working Paper, LSE, London, October

——— (1992b) "International patterns of growth: II. Persistence, path dependence, and sustained take-off in growth transition." Working Paper, LSE, October. (first draft July 1990)

———— (1993a) “Dependence in growth and fluctuations across economies with mobile capital.” Working Paper, LSE, London, June

———— (1993b) “Empirical cross-section dynamics in economic growth.” *European Economic Review* 37, 426–434, April

Quah, Danny, and Thomas J. Sargent (1993) “A dynamic index model for large cross sections.” In *New Research on Business Cycles, Indicators, and Forecasting,* ed. James Stock and Mark Watson (Chicago IL: University of Chicago Press). forthcoming

Silverman, Bernard W. (1986) *Density Estimation for Statistics and Data Analysis* (New York NY 10001: Chapman and Hall)

Summers, Robert, and Alan Heston (1991) “The PennWorld Table (Mark 5): An expanded set of international comparisons, 1950–1988.” *Quarterly Journal of Economics* 106(2), 327–368, May

White, Halbert (1984) *Asymptotic Theory for Econometricians* (New York NY: Academic Press)

Zimmerman, David J. (1992) “Regression toward mediocrity in economic stature.” *American Economic Review* 82(3), 409–429, June

Table 3.1: Cross-country output per worker, relative to world average*
4-state fractile Markov Chain, 1962–84

	Quantile			
	[0.25]	[0.50]	[0.75]	[1.00]
(667)	0.96	0.04		
(690)	0.04	0.93	0.03	
(667)		0.03	0.95	0.02
(690)			0.02	0.98

* The states are arrayed in increasing order, thus the lower right-hand portion of the table shows transitions from the rich to the rich. The numbers in parentheses on the right are the number of country/year pairs beginning in a particular state. These numbers should be equal: they differ from each other by at most the number of years in the sample, due to rounding error. Cells showing 0 to two decimal places are left blank.

Table 3.2: Cross-country output per worker, relative to world average†
5-state fractile Markov Chain, 1962–84

	Quantile				
	[0.2]	[0.4]	[0.6]	[0.8]	[1.0]
(529)	0.95	0.05			
(552)	0.05	0.90	0.05		
(530)		0.05	0.90	0.05	
(551)			0.05	0.93	0.02
(552)				0.01	0.99

† See comments under Table 3.1.

Data Appendix

The data are derived from that given in Summers and Heston (1991). Real per capita income is taken to from `RGDPL` (Laspeyres index); population is `POP`. Countries in the sample were selected by first disallowing those not having continuously available data on these two variables for the period 1960–1985. I then also excluded Kuwait—a 3-dimensional graph of the variables easily shows the Kuwait observation to dominate every other feature of the data. The remaining 118 countries are listed below (integers immediately before the country names are the indexes in the Summers-Heston database):

1	(1)	Algeria	2	(2)	Angola
3	(3)	Benin	4	(4)	Botswana
5	(6)	Burundi	6	(7)	Cameroon
7	(8)	CapeVerdeIs	8	(9)	CentralAfrR
9	(10)	Chad	10	(12)	Congo
11	(13)	Egypt	12	(14)	Ethiopia
13	(15)	Gabon	14	(16)	Gambia
15	(17)	Ghana	16	(18)	Guinea
17	(19)	GuineaBiss	18	(20)	IvoryCoast
19	(21)	Kenya	20	(22)	Lesotho
21	(23)	Liberia	22	(24)	Madagascar
23	(25)	Malawi	24	(26)	Mali
25	(27)	Mauritania	26	(28)	Mauritius
27	(29)	Morocco	28	(30)	Mozambique
29	(31)	Niger	30	(32)	Nigeria
31	(33)	Rwanda	32	(34)	Senegal
33	(36)	SierraLeone	34	(37)	Somalia
35	(38)	SouthAfrica	36	(39)	Sudan
37	(40)	Swaziland	38	(41)	Tanzania
39	(42)	Togo	40	(43)	Tunisia
41	(44)	Uganda	42	(45)	Zaire
43	(46)	Zambia	44	(47)	Zimbabwe

45	(49)	Barbados	46	(50)	Canada
47	(51)	CostaRica	48	(53)	DominicanRep
49	(54)	ElSalvador	50	(56)	Guatemala
51	(57)	Haiti	52	(58)	Honduras
53	(59)	Jamaica	54	(60)	Mexico
55	(61)	Nicaragua	56	(62)	Panama
57	(65)	TrinidadTobag	58	(66)	USA
59	(67)	Argentina	60	(68)	Bolivia
61	(69)	Brazil	62	(70)	Chile
63	(71)	Colombia	64	(72)	Ecuador
65	(73)	Guyana	66	(74)	Paraguay
67	(75)	Peru	68	(76)	Suriname
69	(77)	Uruguay	70	(78)	Venezuela
71	(79)	Afghanistan	72	(81)	Bangladesh
73	(82)	BurmaMyanmar	74	(83)	China
75	(84)	HongKong	76	(85)	India
77	(87)	Iran	78	(88)	Iraq
79	(89)	Israel	80	(90)	Japan
81	(91)	Jordan	82	(92)	KoreaSouthR
83	(94)	Malaysia	84	(95)	Nepal
85	(97)	Pakistan	86	(98)	Philippines
87	(99)	SaudiArabia	88	(100)	Singapore
89	(101)	SriLanka	90	(102)	Syria
91	(103)	Taiwan	92	(104)	Thailand
93	(107)	Austria	94	(108)	Belgium
95	(109)	Cyprus	96	(110)	Denmark
97	(111)	Finland	98	(112)	France
99	(113)	GermanyWest	100	(114)	Greece
101	(116)	Iceland	102	(117)	Ireland
103	(118)	Italy	104	(119)	Luxembourg
105	(120)	Malta	106	(121)	Netherlands
107	(122)	Norway	108	(124)	Portugal
109	(125)	Spain	110	(126)	Sweden

111	(127)	Switzerland	112	(128)	Turkey
113	(129)	UK	114	(130)	Yugoslavia
115	(131)	Australia	116	(132)	Fiji
117	(133)	NewZealand	118	(134)	PapuaNGuinea

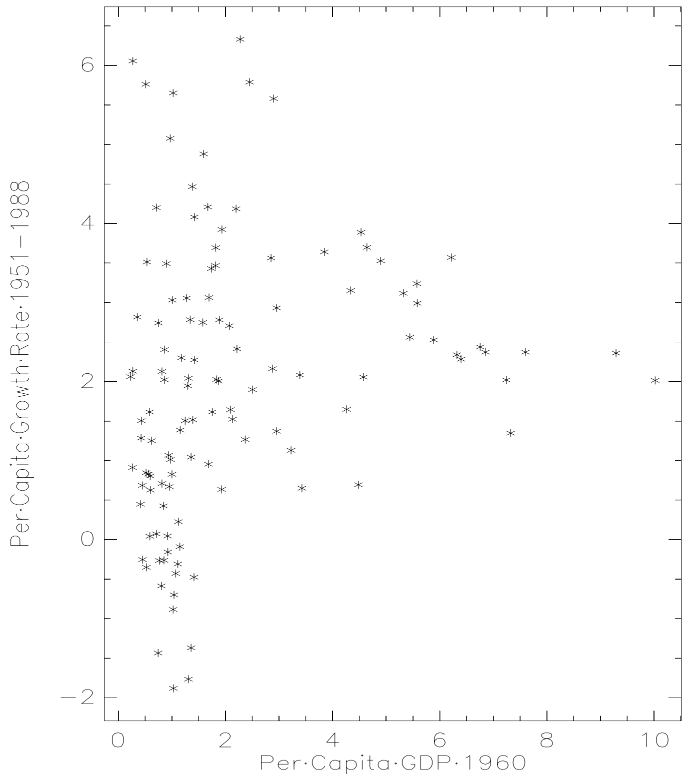


Figure 2.1: Average growth (1951–1988) versus initial levels (1960) [118 Summers-Heston countries]

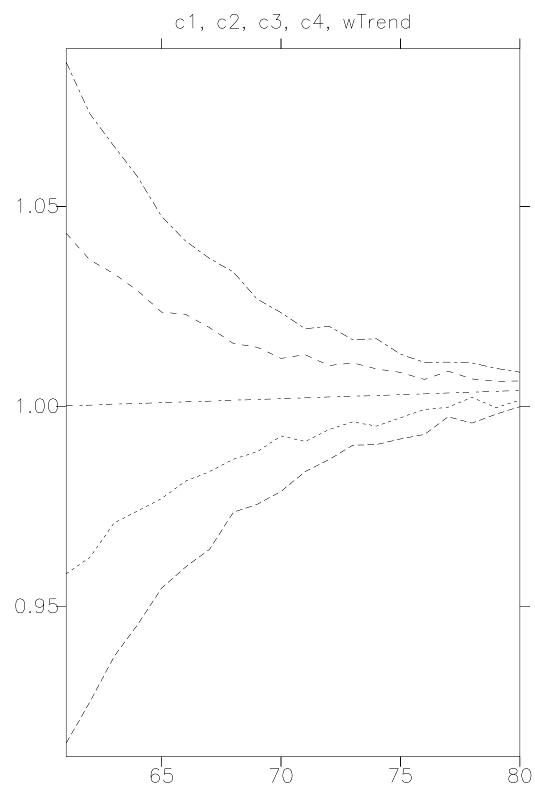


Figure 2.2: Income levels against time (time paths under hypothesized convergence)

Average growth against:

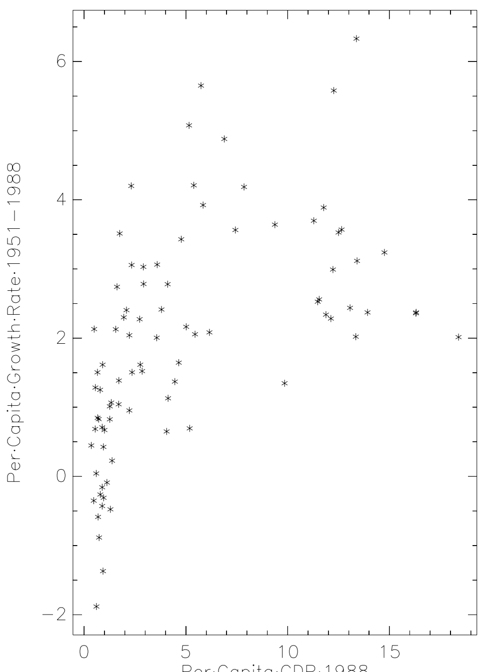


Figure 2.3: Final levels (1980)

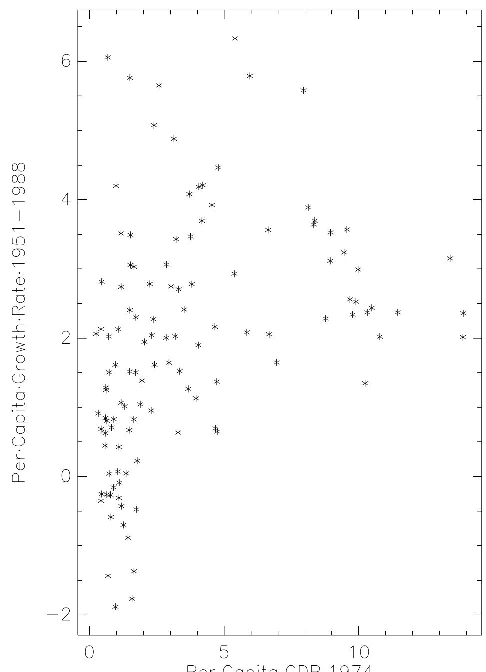


Figure 2.4: Intermediate levels (1974)

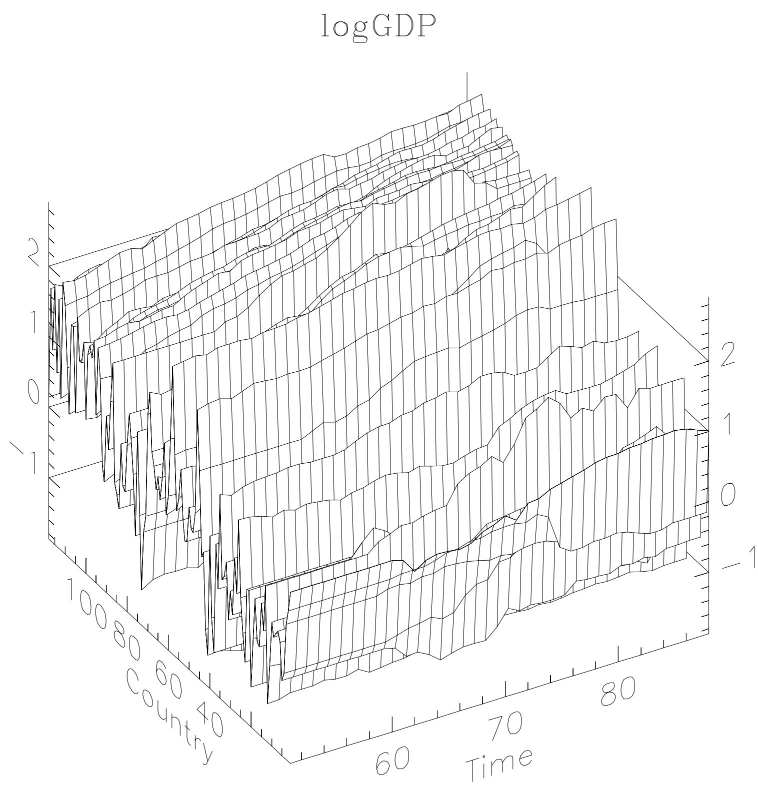
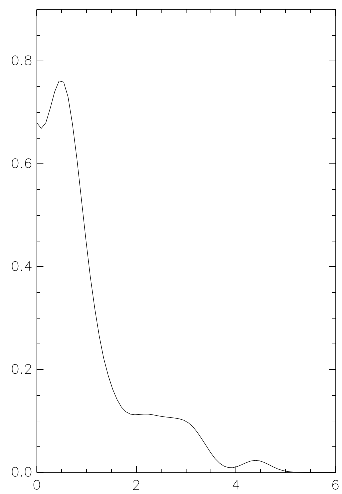
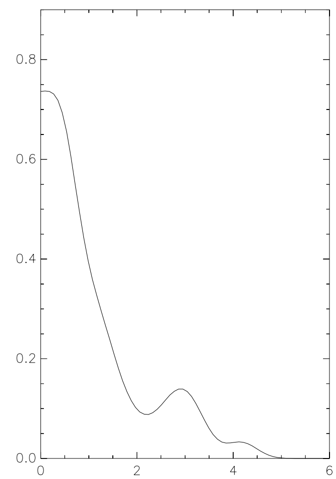


Figure 2.5: Log per capita income [118 Summers-Heston countries]

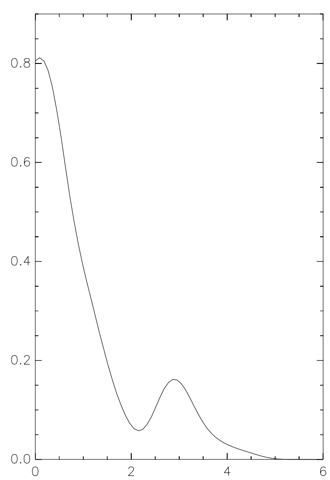
Figure 3.1: Density of Normalized Output per Worker



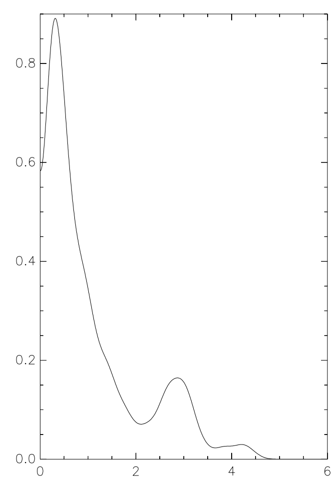
(a): 1962



(b): 1974



(c): 1985



(d): all years