Journal of Economic Growth, 2: 27–59 (1997)

Empirics for Growth and Distribution: Stratification, Polarization, and Convergence Clubs

DANNY T. QUAH

London School of Economics

This paper studies cross-country patterns of economic growth from the viewpoint of income distribution dynamics. Such a perspective raises new empirical and theoretical issues in growth analysis: the profound empirical regularity is an "emerging twin peaks" in the cross-sectional distribution, not simple patterns of convergence or divergence. The theoretical problems raised concern interaction patterns among subgroups of economies, not only problems of a single economy's accumulating factor inputs and technology for growth.

Keywords: conditional, convergence, distribution dynamics, income distribution, inequality, trade, twin peaks

JEL classification: C13, C33, F43, O30

1. Introduction

This paper describes some recent research on patterns of growth across countries: the facts that this paper seeks to document and to explain are given in the stylized features of Figure 1.

Figure 1 is a caricature that will inform my subsequent analysis. Below, I will describe which aspects of this figure have been established to be accurate and which remain conjecture. For the time being, however, it is useful simply to note some features in the caricature.

The horizontal axis in Figure 1 indexes time; the vertical axis, per capita incomes. Figure 1 records, for different time points, (the densities corresponding to) cross-country per capita income distributions. As drawn, the distribution at time t shows most countries having a medium level per capita income; there are few that are very rich, and few very poor.

Over time, cross-country income distributions fluctuate: Figure 1 makes explicit the distribution again at $t + s$. In general, there is one such object for each time period. Figure 1, therefore, is like a time-series plot, except that instead of recording the trajectory of a scalar or vector quantity—like GNP, money, or the price level—the figure comprises the trajectory of an entire distribution.

A first immediate question that pictures like Figure 1 address is whether poor countries are catching up with rich ones. That would happen if, for example, the sequence of distributions collapses over time to a degenerate point limit. But in general that need not occur, and there are other ways whereby the poor can catch up with the rich—as illustrated, for instance, by the criss-crossing arrows.

Figure 1, provocatively, shows the distribution at $t + s$ to have a twin-peaks property: there is a clustering together of the very rich, a clustering together of the very poor, and a

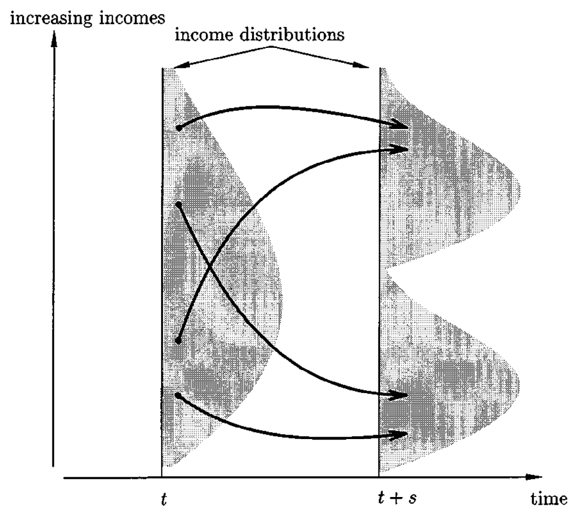


Figure 1. Emerging twin peaks.

vanishing of the middle income class. By contrast, these features were not present at the earlier time t: it therefore seems reasonable to call Figure 1 a picture of *emerging twin peaks*.

In this work, there is nothing special about there being precisely two peaks or modes in the time $t + s$ distribution: What is important instead is that such features have surfaced when previously they were absent. What also matters is that these features have a natural interpretation in terms of *polarization*: those portions of the underlying population of economies collecting in the different peaks may be said to be polarized, one group versus another. More generally, if more than two peaks emerged, it might be natural to call the situation *stratification*. With the underlying population being countries, the economic historian's notion of *convergence clubs*—of countries catching up with one another but only within particular subgroups (e.g., Baumol 1986)—is also apposite.

This paper concerns that body of research on cross-country economic growth that attempts to refine empirically and to understand theoretically such emerging twin-peaks properties.[1] Such a focus might seem excessively narrow: it is useful, therefore, to note how this work relates to other areas of research.

To study the dynamics of cross-section distributions of country incomes—as given in Figure 1—is to combine simultaneously elements of macroeconomic reasoning, microeconomic analysis, and econometric modeling. The researcher is concerned with macroeconomic performance—measured in national income and aggregate growth—but for a rich cross-section of individual cross-sectional units, all potentially interacting with one another. Thus, macroeconomic theories of growth are relevant, but so are microeconomic models of cross-sectional interaction. Figure 1 makes explicit that the success or failure of any one country makes qualitative economic sense only in context: What does a 5% annual growth rate mean—is it high or is it low—if no other economy grows at less than 10% per year? Or if no other economy has ever grown at more than 1% per year?

Arrows drawn in Figure 1 indicate a variety of *intradistribution dynamics*. Some countries rich at time $t+s$ had already been rich at time t; similarly, others poor at $t+s$ had already been poor at t. There is, therefore, *persistence*. However, there is also *churning* or *mobility*: some of those rich at $t+s$ had begun poor; some of those poor at $t+s$ had begun rich. From these and from the vanishing of the middle class between t and $t+s$, it is also natural to suppose a *separating*: some groups of economies originally close together in the middle class have subsequently separated, with some becoming much richer than others—even though they had begun close together.

Figure 1 thus contains a rich spectrum of dynamic behavior. Not only is the global, external shape of the distribution evolving—with twin peaks emerging, and stratification and polarization settling in—but also intradistribution mobility is simultaneously occurring. Some portions of the distribution display persistence in rich and poor states, others show overtaking dynamics, and yet others a slowing-down in growth so that they are themselves overtaken. Put differently, there are both *shape* and *mobility* dynamics in the distributions in Figure 1. An appropriate econometric analysis should capture these. Moreover, researchers might be interested not just in modeling such features in the historical record: they might seek also to project these measured tendencies forwards from the observed sample. In Figure 1, what if $t+s$ is some time in the future? The econometric analysis should provide a model that allows such calculations.

How does this generate new econometrics? Simply tracking the moments of the cross-sectional distributions in Figure 1 will typically shed no light on many of the characteristics I have just described. Similarly uninformative, for Figure 1, would be giving extensive tabulations of the univariate time-series behavior of each of the underlying cross-sectional units or indeed of documenting the multi-variate time-series characteristics of selected subsets of those cross-sectional units. Cross-sectional and panel data regressions, if all they do is capture the behavior of a conditional average, will be altogether unrevealing for the dynamics of the entire cross-section distribution.[2] What a researcher needs to do is to analyze those evolving distribution dynamics directly.

Formulating the problem of economic growth in the form of Figure 1 draws an equivalence between the analysis of growth and of distribution.[3] It is not that higher growth can cause or, alternatively, be driven by greater inequality, but rather the two are considered simultaneously. Note, however, that the distribution that is relevant here is the distribution of income *across* countries, not that *within* a given economy. Thus, the problem considered in this study differs from the classical set of questions prominently considered by Kuznets and subsequently refined in Benabou (1996b), Galor and Zeira (1993), Persson and Tabellini (1994), and many others.

From the perspective of economic growth empirics, the work described below relates to research using convergence predictions to distinguish endogenous and neoclassical growth. That literature is large, but helpfully summarized in Barro and Sala-i-Martin (1995) and Sala-i-Martin (1996). However, some papers have argued that that growth and convergence literature is uninformative for whether poor countries are catching up with rich ones, and unrevealing in general for the dynamics of the distribution of welfare across countries (e.g., Friedman 1992, Leung and Quah 1996, Quah 1993a, b, 1996b, c, f).[4] Such ambiguity, on the other hand, cannot taint the analysis following from Figure 1.

Finally, independent of macroeconomic analyses of aggregate growth, the study of distributions and their dynamics has long been a central part of economic analysis, not just of personal incomes (e.g., Atkinson 1995; Cowell, Jenkins, and Litchfield 1996; Durlauf 1996; Esteban and Ray 1994; Loury 1981; Schluter 1997; and Shorrocks 1978; among others) but also of many other economic categories including earnings, firm and industry shares, regional economic performance, and occupational categories (e.g., Lamo 1996, Lillard and Willis 1978, Lucas 1978, Konings 1994, Koopmans 1995, Quah 1996a, Singer and Spilerman 1976, and Sutton 1995).

While the current work shares ideas with all of these, it also differs in a number of significant ways. But those putative contributions will be easier to see at the end rather than the beginning of the paper.

The remainder of the presentation then is organized as follows: Section 2 puts empirical flesh on the caricature given in Figure 1: the shape and mobility dynamics sketched in Figure 1 are broken down further into density and Tukey box plots, and stochastic kernels. A simple illustrative model is given in Section 3—the model is highly stylized; its purpose is only to suggest the kinds of conceptual modeling issues that will be further helpful. Section 4 builds on those ideas and illustrates the role of conditioning in explaining the distribution dynamics documented in Section 2. Section 5 summarizes the conclusions from this study.

2. A First Empirical Analysis

Figure 2 plots the log of per capita incomes across 105 countries, all relative to the world average per capita income in each year. The underlying data are drawn from the well-known Summers-Heston (1991) dataset.

On the vertical axis in the figure, zero indicates equality with the world average. Time proceeds sequentially along the axis marked *Year*. Along the axis marked *Economy* are the different countries. The particular ordering on this axis gives no insight. Nor will it be used below. For the record, however, the ordering is alphabetical within continents, beginning in Africa with Algeria and Angola, and ending in Oceania with Vanuatu and Western Samoa.

To relate Figure 2 to Figure 1, observe that at each point on the *Year* axis, one can slice across the graph, parallel to the *Economy* axis, and recover the point-in-time cross-country income distribution. I have computed Figure 3.1 and Figure 3.4 doing exactly that.

2.1. Shapes

If I did no more than this, however, I would, of course, have lost important dynamic, intradistribution information—I will return to this point below. For the time being, this procedure gives a sequence of snapshots of the resulting income distributions across countries. Figure 3.1 and Figure 3.4 provide two views of these cross-sectional distributions. The first, Figure 3.1, is a sequence of kernel-smoothed densities taken at roughly decade-long

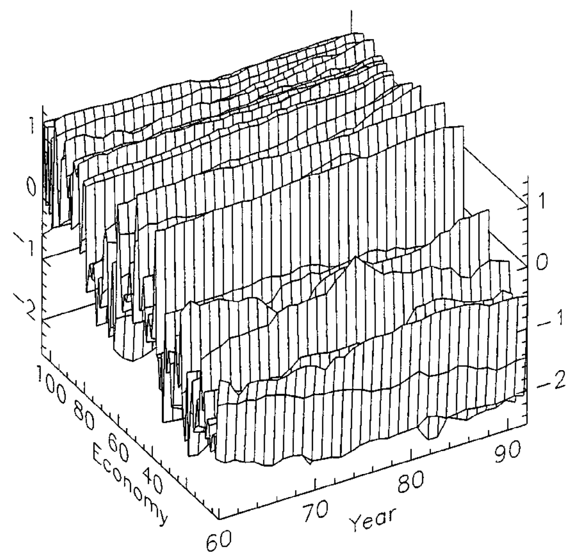


Figure 2. (Log) Relative per capita incomes across 105 countries.

intervals. The second, Figure 3.4, is a sequence of Tukey boxplots for the same underlying data and for the same time periods.

A quick word on inference is useful here: Figure 3.1 and Figure 3.4 record properties of the *population*. The data that go into these figures cannot be interpreted as a random sample. In the language of Efron and Tibshirani (1993), these figures are direct representations of a census: they are not pictures of a random sample from which statistical analysis can help us infer properties of the true underlying population. These pictures already *are* that population. Thus, from the perspective of statistical inference, classical random sampling assumptions do not hold. Below, when I turn to models of endogenous cross-sectional interaction, we will see that the departure here from a classical sampling framework goes even deeper.

The kernel-smoothed estimates in Figure 3.1 were obtained using a Gaussian kernel.[5] By how the data are defined, 1/2 on the horizontal axis indicates one-half the world average per capita income; 2 indicates twice the world average; and so on. Looking across three decades, we see that in 1961 a nascent twin-peakedness—the first mode at a little less than 1, the second at slightly greater than 2.5—was beginning to be visible. By 1988 that second peak had become pronounced. The relative income distance between the peaks doubled from about 1.5 in 1961 to more than 3 in 1988. Finally, for the observations extant, these tendencies appear monotone: the data show no reversals in the dynamics just described.[6]

For completeness, I give in Figure 3.2 and Figure 3.3 two other related snapshot density sequences. In Figure 3.2 the distributions are in natural logs of per capita income; in

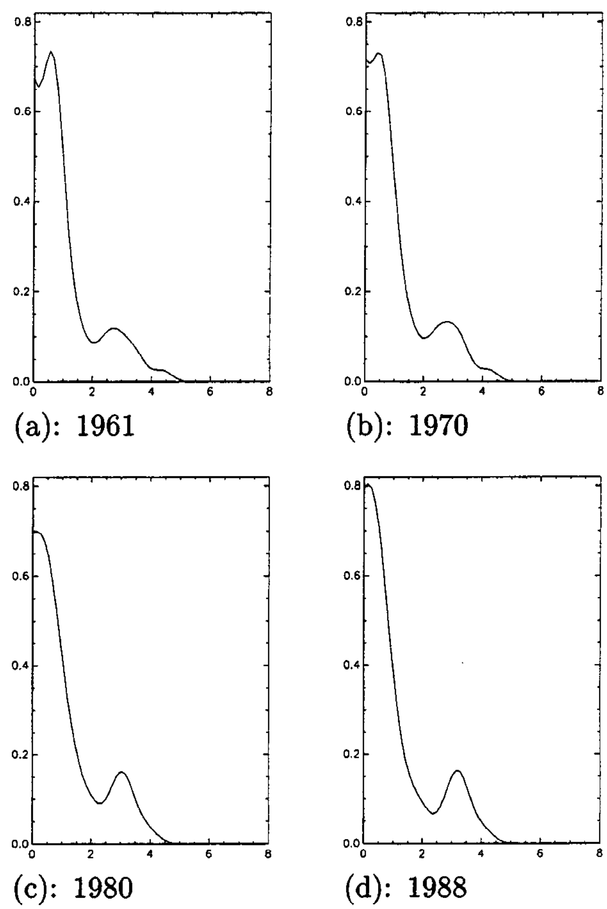


Figure 3.1. Densities of relative (per capita) incomes across 105 countries.

Figure 3.3 the distributions are weighted by the relative numbers of people in each economy. One convenient interpretation of the second is that it shows the distributions of individual incomes across people in the world, assuming that within each economy individual personal incomes are equally distributed, and thus equal to the level of per capita income. Properly interpreted, the "emerging twin peaks" character remains, but is modified. In logs, the peaks are closer together—as one would expect—but the rise of the richer peak at the expense of the poorer remains pronounced. Weighted by populations, the distribution sequence shows *three* peaks, rather than two: the rise of the higher peaks appears to be at the expense of the middle (valley) group. Thus, although details differ, the principal message of the "Emerging twin peaks" Figure 1 comes through in a range of perturbations on the empirical analysis.

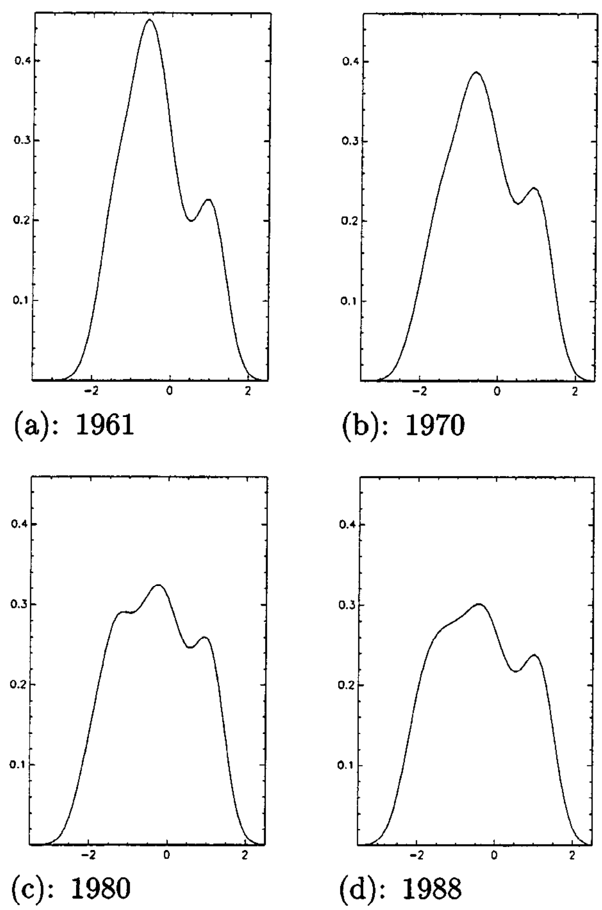


Figure 3.2. Densities of log relative (per capita) incomes across 105 countries.

Figure 3.4 comprises Tukey boxplots constructed from exactly the same data used in Figure 3.1. To understand these pictures, recall the construction of a Tukey boxplot.[7] The box in the middle of each boxplot describes central tendencies of a distribution: the thin line inside the box locates the median; the top and bottom edges are the 75th and 25th percentiles respectively. The middle 50% of the distribution is thus contained in the box; the height of the box—ignoring its vertical location—is the interquartile range.

In Figure 3.4 the middle box for each of the years grows in extant: thus, the middle 50% of the cross-section distribution can be covered only by progressively larger portions of income space. Put differently, the middle 50% of the distribution is spreading out; or, when taken together with the evidence in Figure 3.1, the middle-income class is vanishing—precisely as in Figure 1.

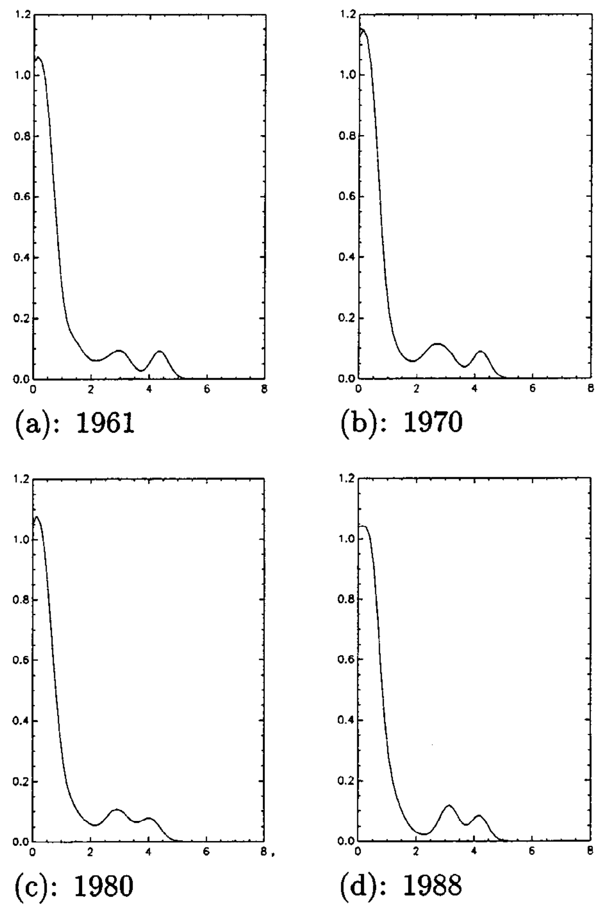


Figure 3.3. Densities of relative (per capita) incomes across 105 countries. Weighted by population.

Emanating from the middle box in each Tukey boxplot are rays reaching to *upper* and *lower adjacent values*. If the interquartile range is r, then the upper adjacent value is the largest income value observed no greater than the 75th percentile plus $1.5 \times r$. The lower adjacent value is similarly defined, extending downwards from the 25th percentile. Indicated by asterisks in Tukey boxplots are *upper* and *lower outside values*—observations that lie outside the upper and lower adjacent values. From a statistical perspective, these might be considered outliers—in the current application, however, these denote the macroeconomies that have performed extraordinarily well or extraordinarily poorly relative to the bulk of other macroeconomies. They represent real people and real countries, not just observations that might be useful to delete in a statistical analysis.

Figure 3.4 shows no extraordinarily poorly-performing economies—or, more accurately,

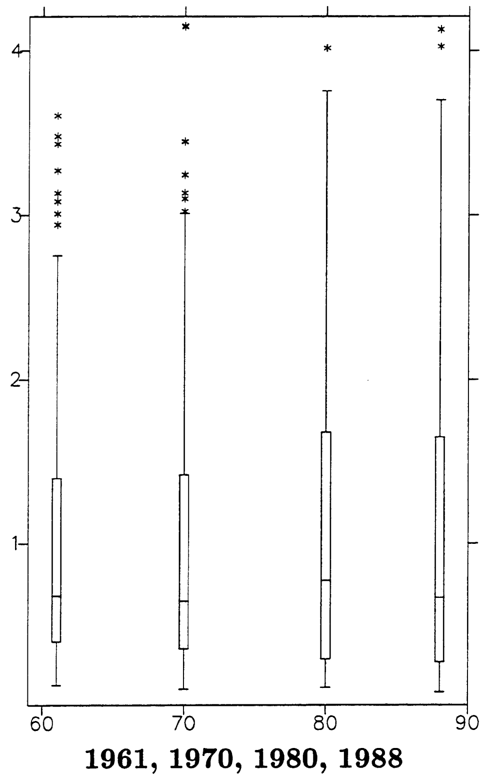


Figure 3.4. Boxplots, relative (per capita) incomes across 105 countries.

when economies performed especially badly, they were not alone. On the upside, by contrast, the early part of the sample showed several outstanding performers: there is a sprinkling of asterisks in the early boxplots. However, over time, parts of the rest of the world have caught up with these initially very rich economies, even as other parts of the world remained poor. Unlike the upper portion, the lower part of the boxplot has never risen, and, indeed, relative to the median shows a continuing decline.

These two descriptions Figure 3.1 and Figure 3.4 have fleshed out and confirmed the *shape* dynamics sketched in the twin-peaks picture Figure 1. We turn now to the *mobility* dynamics also depicted in Figure 1 but not yet examined in the data.

2.2. Mobility

An easy way to quantify churning or intradistribution dynamics in a sequence of distributions is to discretize the space of income values, and then simply count the observed transitions out of and into distinct discrete cells. For instance, in Figure 4—which reproduces the

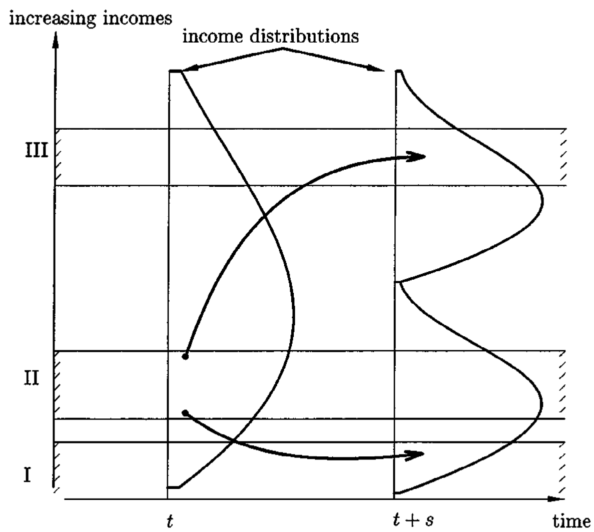


Figure 4. Discretization for intradistribution dynamics.

essential features of Figure 1—one might add up the number of transitions out of cell II into cells I and III respectively (and everywhere else), and then normalize those counts by the total number of observations. Using discrete cells that span the space of all possible realizations, one can then construct a transition probability matrix (as in, e.g., Quah 1993a, b).

It is well known, however, that such a discretization can distort dynamics in important ways when the underlying observations are continuous variables (see, e.g., Chung 1960). The solution is not to use a discretization at all, but to retain the original set of continuous income observations in quantifying intradistribution dynamics. Doing that is like allowing the number of distinct cells {I, II, III, ...} in Figure 4 to tend to infinity and then to the continuum. The corresponding transition probability matrix tends to a matrix with a continuum of rows and columns. In other words, it becomes a *stochastic kernel*, as graphed in Figure 5.1 (in Section 4 below, I provide a more precise technical derivation of a stochastic kernel).

The figure shows the stochastic kernel for 15-year transitions in our relative-income data, averaging over 1961 through 1988. From any point on the axis marked *Period t* extending parallel to the axis marked *Period* $t + 15$ the stochastic kernel is a probability density function: the projection traced out is nonnegative and integrates to unity. That projection is similar to a row of a transition probability matrix: such a row has all entries nonnegative and summing to 1. Roughly speaking, this probability density describes transitions over 15 years from a given income value in period t.

A graph such as Figure 5.1 shows how the cross-sectional distribution at time t evolves into that at $t + 15$. If most of the graph were concentrated along the 45-degree diagonal, then elements in the distribution remain where they began. If, by contrast, most of the

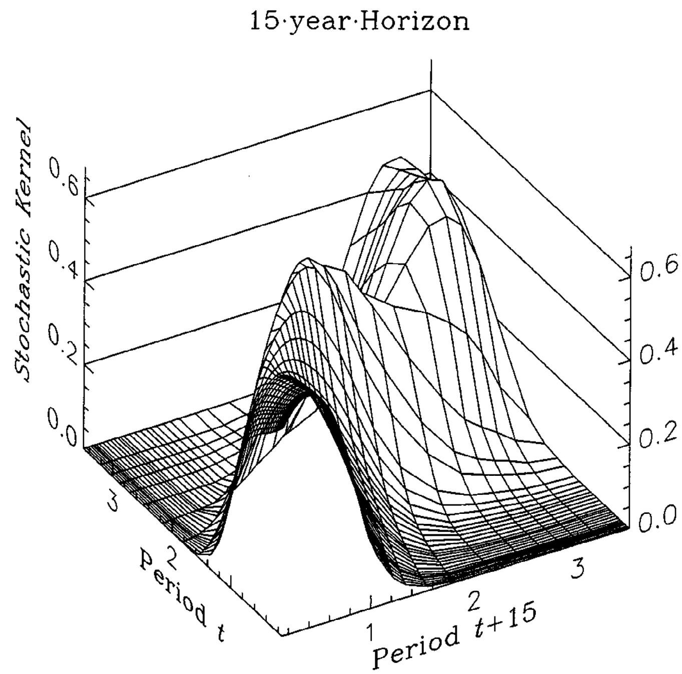
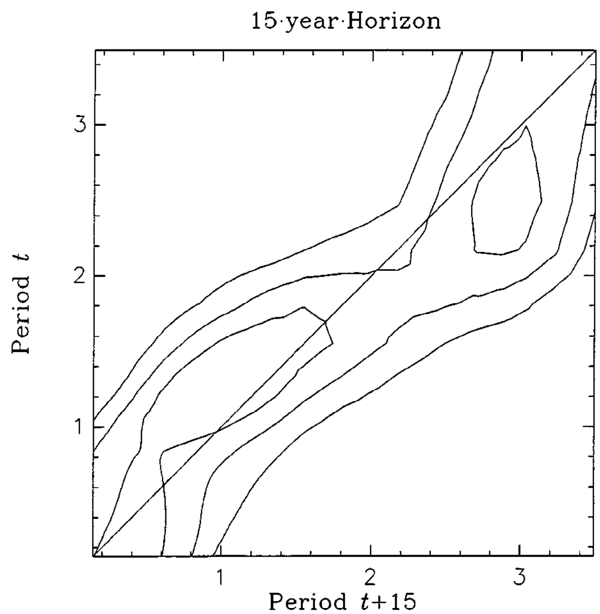


Figure 5.1. Relative income dynamics across 105 countries.

mass in the graph were rotated 90 degrees counter-clockwise from that 45-degree diagonal, then substantial overtaking occurs—the rich become poor, and the poor rich, periodically over 15-year horizons. If most of the graph were concentrated around the 1-value of the *Period* $t+15$ axis—extending parallel to the *Period t* axis—then over a 15-year horizon, the cross-section distribution converges towards equality. Generalizing this, if most of the mass located parallel to the *Period t* axis, with projections on period t-values equal to each other, then the kernel is one where a single (15-period) iteration takes any initial distribution to the same long-run distribution. Dynamics over longer horizons can be studied by recursively applying a given stochastic kernel.[8]

In Figure 5.1 a twin-peaks property again manifests. Over the 15-year horizon, a large portion of the probability mass remains clustered around the main diagonal. However, along that principal ridge, a dip appears in the middle-income portion while the kernel itself rises towards local maxima in both poor and rich parts of the income range. Contour plot Figure 5.2 makes this clearer. The two peaks in the stochastic kernel—because they sit (almost) on the 45-degree diagonal—correspond to what Durlauf and Johnson (1995) call “basins of attraction.” At the same time, however, while the middle-income class is vanishing, portions of the cross section do transit from high to low and from low to high: the stochastic kernel is positive almost everywhere, and communicates across the entire range of income values.



Contour plot at levels 0.2, 0.35, 0.5

Figure 5.2. Relative income dynamics across 105 countries.

Figures 5.3, 5.4, 5.5, and 5.6 provide comparable stochastic kernel representations on, respectively, the log and population-weighted versions of the income distributions. In both cases, the message from the unweighted per-capita income is amended but not overturned. The twin peaks are closer together in the log case (as in just the snapshot density sequence Figure 3.2), but the stochastic kernel again shows clearly the polarization dynamics. In the population-weighted case, the multiplicity of peaks is again evident.

To sum up this first pass through the data, we conclude that the data confirm most of the stylized features earlier given in Figure 1. There is a wide spectrum of intradistribution dynamics—overtaking and catching-up occur simultaneously with persistence and languishing—while overall the twin-peaks shape in the cross-sectional distribution emerges.

3. A Simple Model and Empirical Implications

What might explain these "emerging twin-peaks" regularities? In particular, what theoretical models and further empirical analyses will shed light on these stylized facts?

Given this assignment, it is hard to see what insights will obtain from a conventional approach: study standard "growth and convergence" models of representative economies, and then analyze such models using panel-data econometric methods that absorb heterogeneity into "individual effects." Sure, those techniques deal with data that show rich cross-section

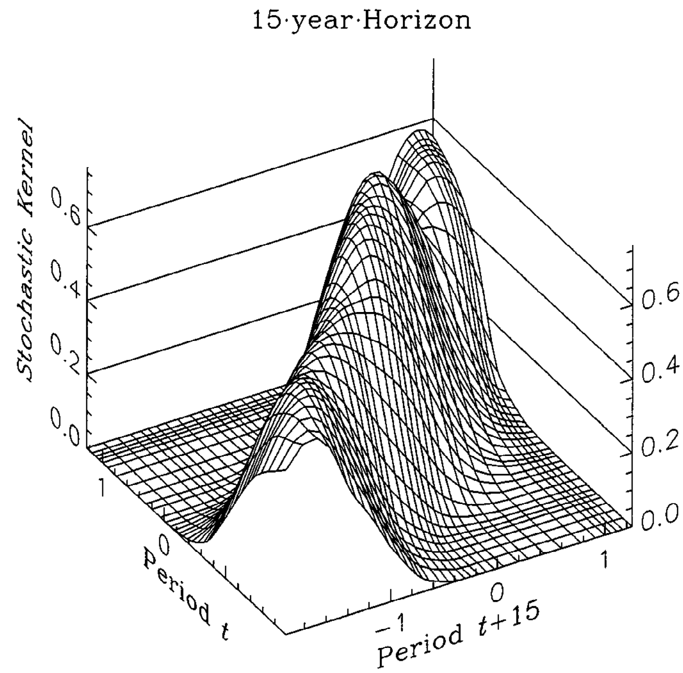
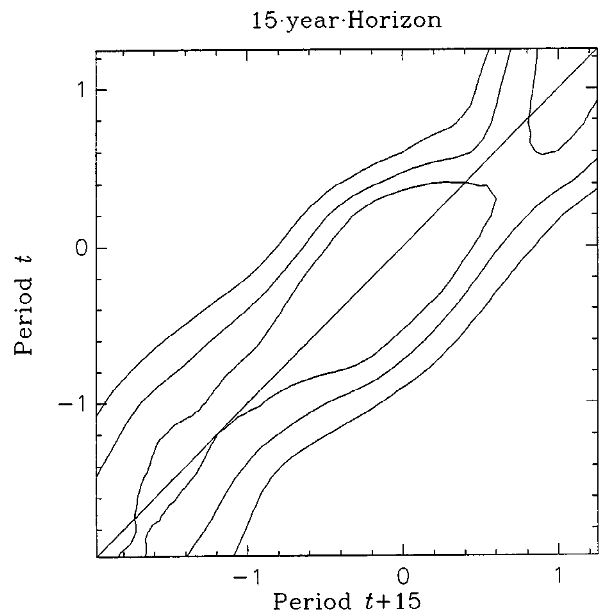


Figure 5.3. Log relative income dynamics across 105 countries.

and time-series variation (as in the uninformative Figure 2), but that fact alone does not recommend them.[9] Individual-effects panel data methods had been developed to take into account the inconsistency in estimating regression coefficients when unobserved heterogeneity is correlated with regressors (Chamberlain, 1984, makes this particularly clear). They were not designed to naturally provide a picture of how an entire distribution evolves. Those regression methods average across the cross section: they can give only a picture of the behavior of the conditional mean, not of the whole distribution. Moreover, sweeping out individual heterogeneities, in the current application, amounts to no more than resigning oneself to leave unexplained the (significant) differences across individual countries. But this is precisely what we wish to understand here.

Instead, it is reasonable to guess that more insightful for studying twin-peaks behavior will be theorizing directly in terms of the entire distribution, and permitting explicit patterns of cross-section interaction—clustering together into distinct clumps—to endogenously emerge. Examples of such models exist. They include those in Durlauf (1993), Ioannides (1990), Kirman, Oddou, and Weber (1986), Townsend (1983), and Quah (1996d). Here, I present a simplified version of the model in Quah (1996d); it is stripped down to an extent that allows insight into key issues but at some cost in rigor and economic motivation.[10]

Let $\mathcal{J}$ be the index set of economies, taken as fixed throughout the discussion. A *coalition* of economies is a subset $\mathcal{C}$ of $\mathcal{J}$. Each economy l in $\mathcal{J}$ is characterized by an economy-



Contour plot at levels 0.2, 0.35, 0.5

Figure 5.4. Log relative income dynamics across 105 countries.

specific stock h_l, which can be interpreted as human capital. This stock is used in two nonrival ways: first, it represents the potential for technical progress and ongoing growth—it is the source of useful ideas. Second, it produces nonstorable output for current consumption—it is an input in a production technology.

Production occurs from coalitions of economies forming to jointly produce a single nondurable consumption good. Denote the total output of coalition $\mathcal{C}$ by $Y_{\mathcal{C}}$. Assume that $Y_{\mathcal{C}}$ depends on the distribution of h_l across l in $\mathcal{C}$, and is increasing in each h_l. Assume also that out of total coalition output, economy l in $\mathcal{C}$ gets portion $\psi(Y_{\mathcal{C}}, h_l)$, with ψ increasing in both arguments, and satisfying exact product exhaustion:

$$\sum_{l \text{ in } \mathcal{C}} \psi(Y_{\mathcal{C}}, h_l) = Y_{\mathcal{C}}.$$

(Primitive assumptions implying these properties would be first, compensation according to marginal product and second, the CES technology

$$Y_{\mathcal{C}} = \left[\sum_{l \text{ in } \mathcal{C}} h_l^{\theta} \right]^{1/\theta}, \qquad 0 < \theta < 1,$$

with θ, describing the elasticity of substitution in the CES production function, giving isoquants between linear and Cobb-Douglas technologies. Quah (1996d) gives the natural interpretation of these properties as economies of scale deriving from specialization.)

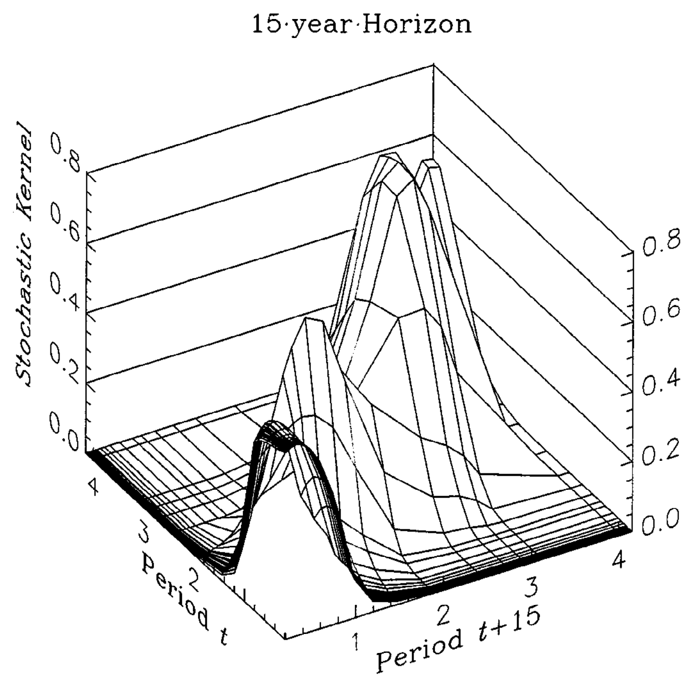


Figure 5.5. Relative income dynamics across 105 countries, weighted by population.

By these assumptions, enlarging the coalition always increases total output $Y_{\mathcal{C}}$. The compensation scheme ψ then ensures that all economies unanimously agree to be in the single grand coalition comprising the entire cross section. This, therefore, is a force for consolidation. If this were all there were at work, the cross-sectional interaction would be trivial: the only coalition that exists includes simultaneously all members of the cross section $\mathcal{J}$.

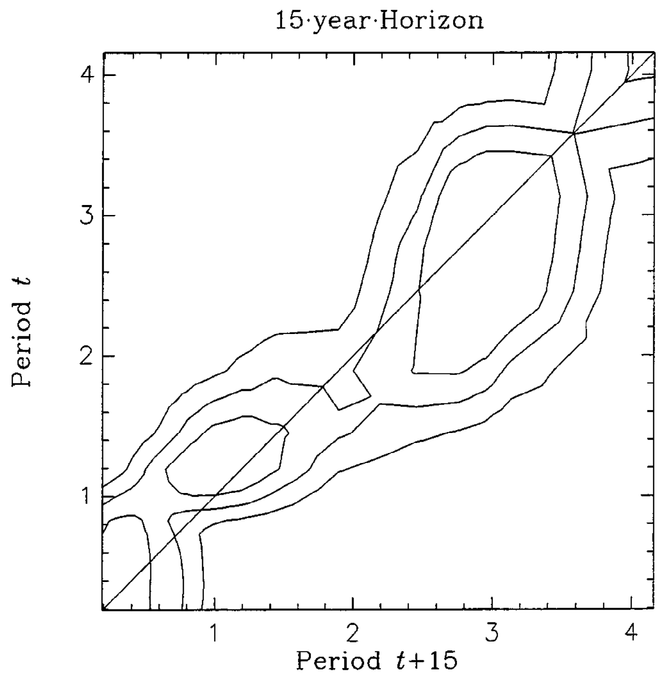
Turn now to the dynamics of h_l. Denote for each coalition $\mathcal{C}$ the average value of h_l's across $\mathcal{C}$ by $H_{\mathcal{C}}$. Suppose that human capital in economy l in $\mathcal{C}$ evolves as

$$\dot{h}_l = \tilde{\phi}(h_l, H_{\mathcal{C}}) \quad \text{for } l \text{ in } \mathcal{C},$$

with $\tilde{\phi}$ increasing in both arguments and homogeneous degree 1. This says simply that human capital in economy l accumulates not just from the human capital already extant in it, but also from the average human capital extant in the economies with which l interacts.[11] Dividing by h_l, this becomes the proportional growth equation:

$$\dot{h}_l/h_l = \tilde{\phi}(1, H_{\mathcal{C}}/h_l) \stackrel{\text{def}}{=} \phi(H_{\mathcal{C}}/h_l).$$

By construction ϕ is increasing in the ratio $H_{\mathcal{C}}/h_l$.



Contour plot at levels 0.2, 0.35, 0.5

Figure 5.6. Relative income dynamics across 105 countries, weighted by population.

It is now easy to see the force for fragmentation, and against the single grand coalition forming. Economies in higher average h coalitions have faster proportional growth rates. The problem with allowing a coalition to get too large is that the coalition then (typically) lowers its average $H_{\mathcal{C}}$: this would slow growth for all the economies already in the coalition. Economies already in good coalitions would, ceteris paribus, refuse to admit economies that lower the coalition average $H_{\mathcal{C}}$.

The force for consolidation (the compensation $\psi(Y_{\mathcal{C}}, h_l)$) is a *level* effect—it affects current consumption. The force for fragmentation (the growth $\phi(H_{\mathcal{C}}/h_l)$) is a *slope* effect—it affects future consumption. Parameterizing economies' discount rates for intertemporal consumption allows calibrating the tradeoff across level and slope effects, and thus provides a theory of endogenous coalition formation. Equilibrium is a set of coalitions $\{\mathcal{C}_1, \mathcal{C}_2, \mathcal{C}_3, \ldots\}$ such that no economy assigned to a coalition wishes to belong to a different coalition agreeing to admit it. Quah (1996d) describes an equilibrium that comprises nontrivial consecutive subsets of the cross section $\mathcal{J}$ of economies. Then, as shown in Figure 6.2, the distribution of incomes across economies within the same coalition converges towards equality; those across different coalitions separate and then diverge.

The equilibrium distribution dynamics that arise depends on the functions ϕ and ψ and, significantly, also on initial conditions in the distribution of h. If, the initial distribution were not that given in Figure 6.2, but instead that of Figure 6.1, then the model implies

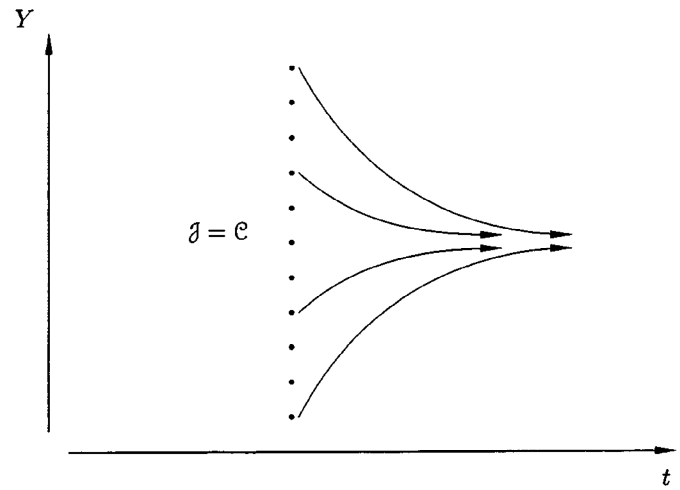


Figure 6.1. Convergence, one-world.

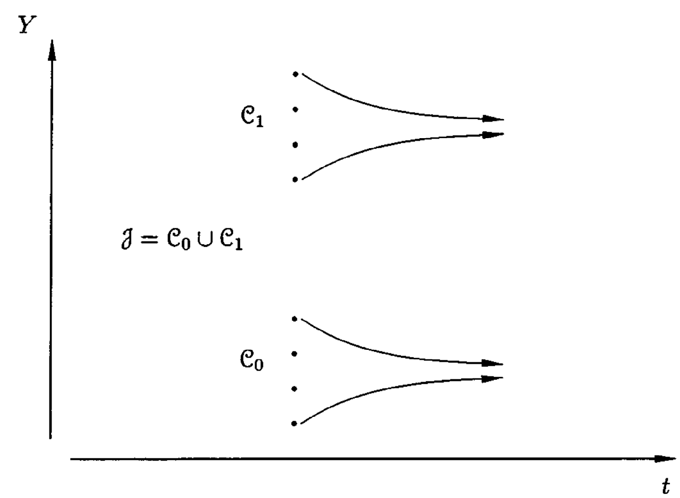


Figure 6.2. Convergence, two-coalition polarization.

convergence of the entire cross-section, not subsets, to a single degenerate point mass.

The ideas here can be enriched in a variety of ways: if what mattered were some multidimensional attribute and not just the single-dimensioned h, then equilibrium coalitions need not be consecutive (as in Figure 6.1 and Figure 6.2), but might intermesh in some form of a seamless web. If important stochastic disturbances perturbed each economy's development process, then overtaking and criss-crossing across coalitions and different parts of the income distribution might occur.

The basic insight, however, remains: by studying interactions across the cross-section, one gets a picture of how the entire distribution evolves through time. Clustering—the polarization or stratification that emerges in Figure 6.2—into convergence clubs then manifests as a central part of the economic reasoning. For empirical analysis, it is useful to note that these endogenous cross-section interactions result in violation of the classical random sampling assumptions that are traditionally adopted in work with cross-sectional data.

In general, the coalitions that form in equilibrium might be only implicit: no formal observable organization need be arranged to house them. An equilibrium such as Figure 6.2 is to a degree already consistent with the emerging twin-peaks stylized facts previously discussed. But can empirics shed further light on the dimensions along which such coalitions (implicitly) form? I turn next to this.

4. Conditioning

The emerging twin-peaks picture in Figure 1 is an instance of *unconditional dynamics*. What the analysis leading up to Figure 6.2 has done is provided a set of predictions largely consistent with those unconditional dynamics. Put another way, Figure 1 is a reduced form consistent with one particular structure—that described in the previous section. The goal of this section is to obtain independent, auxiliary evidence that might shed more light on the link between the underlying structure and the reduced-form distribution dynamics of Figure 1. Such an exercise might even be regarded as *explaining* distribution dynamics.

Note that to explain distribution dynamics, in the sense used here, is more involved than, say discovering a particular coefficient to be significant in a regression of a dependent variable on some right-hand side variables. What we seek is an empirical computation that helps us understand the law of motion in an entire distribution.

The key insight is to turn on its head the reasoning behind the "unconditional dynamics" already given. Just as stochastic kernels quantify how distributions evolve through time, they can also describe how a set of conditioning factors *alters* the cross-sectional distribution of income. Thus, to understand if a hypothesized set of factors explains emerging twin-peaks features, we can simply ask if the stochastic kernel transforming the unconditional distribution to a conditional one removes those same features.

To see how this works in detail, recall the technical derivation of a stochastic kernel. Since we are concerned here with real-valued incomes, the underlying state space is the pair $(\mathbb{R}, \mathfrak{R})$, i.e., the real line $\mathbb{R}$ together with the collection $\mathfrak{R}$ of its Borel sets. Let $\mathbf{B}(\mathbb{R}, \mathfrak{R})$ denote the Banach space of bounded finitely-additive set functions on the measureable space $(\mathbb{R}, \mathfrak{R})$ endowed with total variation norm:

$$\forall \mu \text{ in } \mathbf{B}(\mathbb{R}, \mathfrak{R}): \qquad |\mu| = \sup \sum_j |\mu(A_j)|,$$

where the supremum in this definition is taken over all $\{A_j : j = 1, 2, \ldots, n\}$ finite measurable partitions of $\mathbb{R}$.

Empirical distributions on $\mathbb{R}$ can be identified with probability measures on $(\mathbb{R}, \mathfrak{R})$; those are, in turn, just countably-additive elements in $\mathbf{B}(\mathbb{R}, \mathfrak{R})$ assigning value 1 to the entire space $\mathbb{R}$. Let $\mathfrak{B}$ denote the Borel σ-algebra generated by the open subsets (relative to total variation norm topology) of $\mathbf{B}(\mathbb{R}, \mathfrak{R})$. Then $(\mathbf{B}, \mathfrak{B})$ is another measurable space.

Note that $\mathbf{B}$ includes more than just probability measures: an arbitrary element μ in $\mathbf{B}$ could be negative; $\mu(\mathbb{R})$ need not be 1; and μ need not be countably-additive. On the other hand, a collection of probability measures is never a linear space: that collection does not include a zero element; if λ_1 and λ_2 are probability measures, then $\lambda_1 - \lambda_2$ and $\lambda_1 + \lambda_2$ are not; neither is $x\lambda_1$ a probability measure for $x \in \mathbb{R}$ except at $x = 1$. By contrast, the set of bounded finitely-additive set functions certainly is a linear space, and as described above, is easily given a norm and then made Banach.

Why embed probability measures in a Banach space as we have done here? A first reason is so that distances can be defined between probability measures; it then makes sense to talk about two measures—and their associated distributions—getting closer to one another. A small step from there is to define open sets of probability measures, and thereby induce (Borel) σ-algebras on probability measures. Such σ-algebras then allow modeling random elements drawn from collections of probability measures, and thus from collections of distributions. The data of interest when modeling the dynamics of distributions are precisely random elements taking values that are probability measures.

This framework allows a more rigorous description of the stochastic kernels already used above. Let F_t denote the distribution of incomes across economies at a given time t. Associated with F_t is a measure λ_t in $(\mathbf{B}, \mathfrak{B})$. If $(\Omega, \mathfrak{F}, \Pr)$ is the underlying probability space, then λ_t is the value of an $\mathfrak{F}/\mathfrak{B}$-measureable map $\Lambda_t\colon (\Omega, \mathfrak{F}) \to (\mathbf{B}, \mathfrak{B})$. The sequence $\{\Lambda_t\colon t \geq 0\}$ is then a $\mathbf{B}$-valued stochastic process.

How should the law of motion for such a process be modeled?

The simplest scheme for doing so is analogous to the first-order autoregression from standard time-series analysis:

$$\lambda_t = T^*(\lambda_{t-1}, u_t) = T^*_{u_t}(\lambda_{t-1}), \quad t \geq 1,$$

where T^* is an operator that maps the product of measures together with generalized disturbances u to probability measures; and $T^*_{u_t}$ absorbs the disturbance into the definition of the operator. (Why the * appears in T^* and $T^*_{u_t}$ will be clarified below.) This is no more than a stochastic difference equation taking values that are entire measures; equivalently, it is an equation describing the evolution of the distribution of incomes across economies.

To understand the structure of operators like $T^*_{u_t}$, it helps to use the following:

Stochastic kernel definition: *Let μ and ν be elements of* $\mathbf{B}$ *that are probability measures on* $(\mathbb{R}, \mathfrak{R})$. *A* ***stochastic kernel*** *relating μ and ν is a mapping $M_{(\mu,\nu)}\colon (\mathbb{R}, \mathfrak{R}) \to [0, 1]$ satisfying:*

(i) $\forall y$ *in* $\mathbb{R}$, *the restriction $M_{(\mu,\nu)}(y, \cdot)$ is a probability measure;*

(ii) $\forall A$ *in* $\mathfrak{R}$, *the restriction $M_{(\mu,\nu)}(\cdot, A)$ is $\mathfrak{R}$-measurable;*

(iii) $\forall A$ *in* $\mathfrak{R}$, *we have $\mu(A) = \int M_{(\mu,\nu)}(y, A)d\nu(y)$.*

To see why this is useful, first consider (iii). In an initial period, for given y, there is some fraction $d\nu(y)$ of economies with incomes close to y. Count up all economies in that group who turn out to have their incomes subsequently fall in a given $\mathfrak{R}$-measurable subset $A \subseteq \mathbb{R}$. When normalized to be a fraction of the total number of economies, this count is precisely $M(y, A)$ (where the (μ, ν) subscript can now be deleted without loss of clarity). Fix A, weight the count $M(y, A)$ by $d\nu(y)$, and sum over all possible y's, i.e., evaluate the integral $\int M(y, A) d\nu(y)$. This gives the fraction of economies that end up in state A regardless of their initial income levels. If this equals $\mu(A)$ for all measurable subsets A, then μ must be the measure associated with the subsequent income distribution. In other words, the stochastic kernel M is a complete description of transitions from state y to any other portion of the underlying state space $\mathbb{R}$.

Conditions (i) and (ii) simply guarantee that the interpretation of (iii) is valid. By (ii), the right hand side of (iii) is well-defined as a Lebesgue integral. By (i), the right hand side of (iii) is a weighted average of probability measures $M(y, \cdot)$, and thus is itself a probability measure.

How does this relate to the structure of $T^*_{u_t}$? Let $\mathbf{b}(\mathbb{R}, \mathfrak{R})$ be the Banach space under sup norm of bounded measurable functions on $(\mathbb{R}, \mathfrak{R})$. Fix a stochastic kernel M and define the operator T mapping $\mathbf{b}(\mathbb{R}, \mathfrak{R})$ to itself by

$$\forall f \text{ in } \mathbf{b}(\mathbb{R}, \mathfrak{R}), \quad \forall y \text{ in } \mathbb{R}: \qquad (Tf)(y) = \int f(x) M(y, dx).$$

Since $M(y, \cdot)$ is a probability measure, the image Tf can be interpreted as a forwards conditional expectation. For example, if all economies in the cross section begin with incomes y, and we take f to be the identify map, then $(Tf)(y) = \int x M(y, dx)$ is next period's average income in the cross section, conditional on all economies having income y in the current period.

Clearly, T is a bounded linear operator. Denote the adjoint of T by T^*. By Riesz Representation Theorem, the dual space of $\mathbf{b}(\mathbb{R}, \mathfrak{R})$ is just $\mathbf{B}(\mathbb{R}, \mathfrak{R})$ (our original collection of bounded finitely additive set functions on $\mathfrak{R}$); thus T^* is a bounded linear operator mapping $\mathbf{B}(\mathbb{R}, \mathfrak{R})$ to itself. It turns out that T^* is also exactly the mapping in (iii) of the Stochastic Kernel Definition, i.e.,

$$\forall \nu \text{ probability measures in } \mathbf{B}, \forall A \text{ in } \mathfrak{R}: \qquad (T^*\nu)(A) = \int M(y, A) d\nu(A).$$

(This is immediate from writing the left side as

$$\begin{aligned}
(T^*\nu)(A) &= \int \mathbf{1}_A d(T^*\nu)(y) = \int (T\mathbf{1}_A)(y)\, d\nu(y) && \text{(adjoint)} \\
&= \int \left[\int \mathbf{1}_A(x) M(y, dx) \right] d\nu(y) && \text{(definition of } T\text{)} \\
&= \int M(y, A)\, d\nu(y), && \text{(calculation)}
\end{aligned}$$

with $\mathbf{1}_A$ the indicator function for A.)

The definition of a stochastic kernel nowhere requires that ν and its image μ under T^* be sequential in time. Thus, a stochastic kernel M representing T^* can be used to relate any two different distributions, in particular an unconditional observed distribution, and one conditional on a set of explanatory factors.

To sharpen the focus, we return to the model of cross-sectional interaction given in Section 3. In that model, while under certain conditions stratification emerges, as in Figure 6.2, that feature is absent if we consider each economy only in relation to the other members of its implicit coalition. In particular, conditioning each economy's observations on the behavior of its *neighbors*, the distribution dynamics imply convergence to a degenerate point mass. This motivates the following:

Conditioning scheme definition: *For a collection of economies $\mathcal{J}$, a **conditioning scheme** $\mathfrak{S}$ is a collection of triples, one for each economy l in $\mathcal{J}$ at time instant t, with each triple comprising:*

(i) $\tau_l(t)$ *an integer lag,*

(ii) $C_l(t)$ *a subset of $\mathcal{J}$, and*

(iii) $\bar{\omega}_l(t)$ *a set of probability weights on $\mathcal{J}$ never positive outside $\mathcal{C}_l(t)$.*

The subset $\mathcal{C}_l(t)$ is the collection of economies associated with, or neighbors of, l at time t. Since the weights $\bar{\omega}_l(t)$ can be positive only on $\mathcal{C}_l(t)$, they determine the relative importance of the different economies in $\mathcal{C}_l(t)$ in the evolution of economy l at time t. Finally, $\tau_l(t)$ is the lag in time: it indicates the delay with which developments in the economies in $\mathcal{C}_l(t)$ affect l. Since $\tau_l(t)$ can be a positive constant, l's associated neighbors $\mathcal{C}_l(t)$ can include l itself.

Sometimes, it will be convenient to use an arbitrary scaled version of the probability weights $\bar{\omega}_l(t)$, i.e., an appropriate collection of nonnegative numbers having a nonzero sum. Such an alternative collection will be denoted $\omega_l(t)$. The probability weights $\bar{\omega}_l(t)$ of a conditioning scheme can always be constructed from $\omega_l(t)$, so that referring only to the latter is without loss.

If $Y = \{Y_l(t)\colon l \text{ in } \mathcal{J} \text{ and } t \geq 0\}$ denotes the original observations on (relative) per capita incomes, define the *conditional* version $Y \mid \mathfrak{S} = \tilde{Y}$ by

$$\tilde{Y}_l(t) \stackrel{\text{def}}{=} Y_l(t)/\hat{Y}_l(t)$$

where

$$\hat{Y}_l(t) \stackrel{\text{def}}{=} \sum_{j \text{ in } \mathcal{C}_l(t)} \bar{\omega}_j(t) Y_j(t - \tau_l(t)).$$

In words, $Y \mid \mathfrak{S}$ comprises per capita incomes relative to one's neighbors, appropriately weighted.[12]

Some special cases will help develop intuition. Fix $\tau_l(t) = \tau > 0$; take $\mathcal{C}_l(t) = \{l\}$, i.e., just the economy itself; and define $\omega_l(t) = \{1 \text{ on } \mathcal{C}_l(t) \text{ and } 0 \text{ elsewhere}\}$. Then the

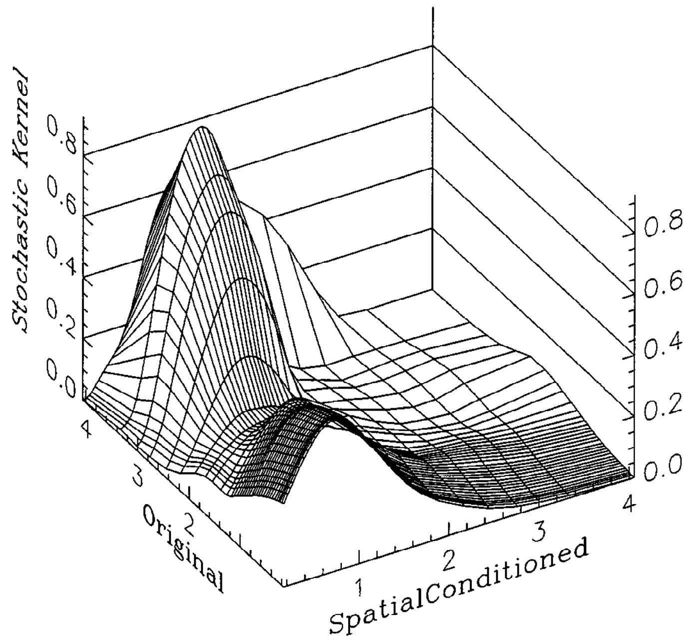


Figure 7.1. Stochastic kernel, relative (per capita) incomes across 105 countries.

stochastic kernel mapping the unconditional distribution in Y to the conditional $Y \mid \mathfrak{S}$ is just the usual Markov stochastic kernel for τ-period transitions. It is this, for $\tau = 15$, that is graphed in Figure 5.

For the second example, fix $\tau_l(t) = 0$; take $\mathcal{C}_l(t) = \mathcal{C}_l(0)$ to be the set of economies physically contiguous with l but excluding l itself; and let $\omega_l(t)$ but the set of population sizes at time $t - \tau_l(t) = t$ in each of the economies in $\mathcal{C}_l(0)$. Then $\hat{Y}_l(t)$ is the per capita income in all contiguous economies abutting on l; and $\tilde{Y}_l(t)$ is per capita income in l relative to that in the surrounding economies. In words $Y \mid \mathfrak{S}$ is per capita income relative to one's physical neighbors. The stochastic kernel mapping $Y \to Y \mid \mathfrak{S}$ would depart from the identity map if the convergence clubs of Section 3 were spatially concentrated: most interaction and exchange occurred within groups of countries physically close to one another.[13]

Figure 7.1 displays precisely that stochastic kernel. The graph here has the same interpretations as previously given for Figure 5.1 except that the axes are now *Original* and *Conditioned* rather than *Period* t and *Period* $t + 15$. (For completeness, I also give Figure 7.2 and Figure 7.3 showing the densities and boxplots for $Y \mid \mathfrak{S}$.) The most prominent change comparing stochastic kernels in Figure 7.1 and Figure 5.1 is the counterclockwise shift in mass to parallel the *Original* axis. Put differently, spatial factors account for a large part of the distribution of incomes across countries: rich economies are typically close

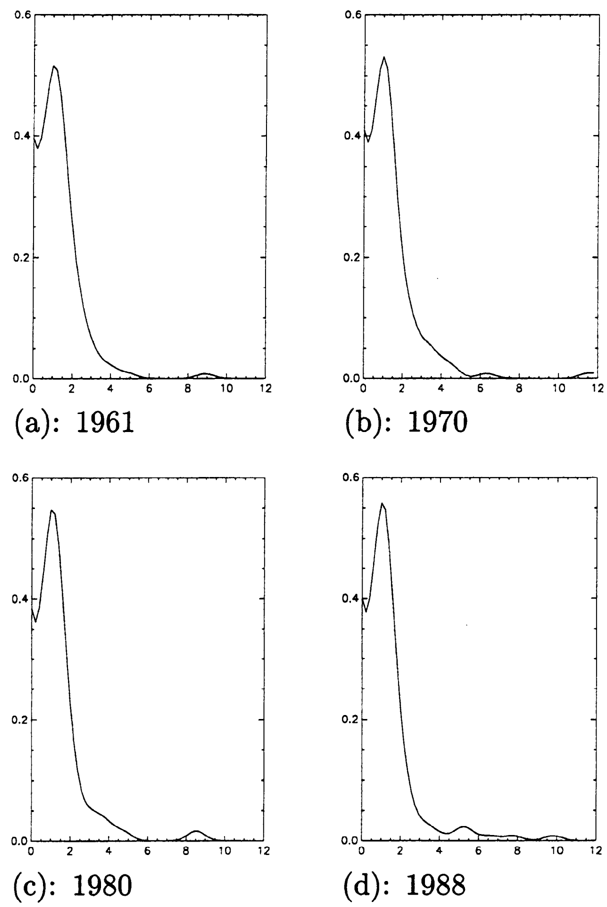


Figure 7.2. Densities of relative (per capita) incomes across 105 countries.

to—interact more with—other rich ones; similarly poor economies are typically close to other poor ones.

The snapshot densities for $Y \mid \mathfrak{S}$ displayed in Figure 7.2 and Figure 7.3 no longer show emerging twin-peaks features. In summary, it appears that the polarization earlier identified in the unconditional distribution-dynamics of cross-country incomes is well explained by physical geography. Not only are rich countries located close to other rich ones, such tendencies have magnified through time.

However, while physical geography has been just shown to play an important role in cross-country income distribution dynamics, it is likely international trade that most economists would identify as the interactions of Section 3. This is easily incorporated in the conditioning-schemes analysis, and constitutes the third example. Let $\tau_l(t) = 0$ so that

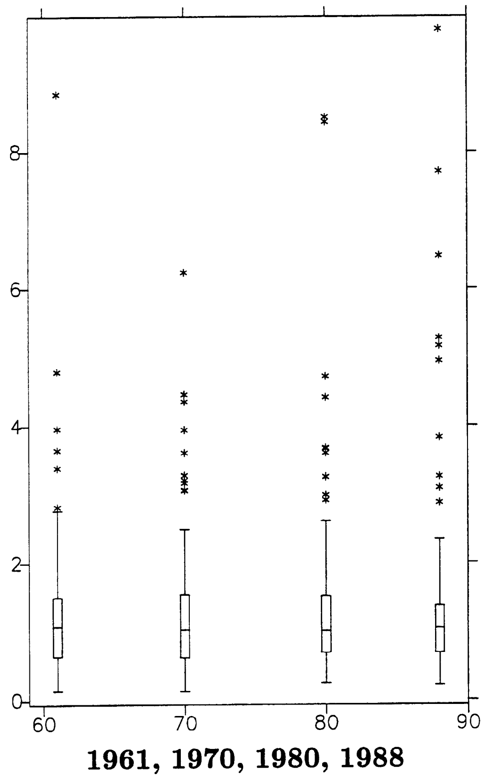


Figure 7.3. Boxplots, relative (per capita) incomes across 105 countries.

there is no delay. Fix T to be the end period of the observed sample, and take $\mathcal{C}_l(t) = \mathcal{C}_l(T)$ to be that set of l's trading partners at time T such that their total trade (imports plus exports) share is at least 50% of l's total trade at T. Finally, let $\omega_l(t) = \omega_l(T)$ be the measured trade shares of the different economies in $\mathcal{C}_l(t)$ out of l's total trade accounts. The stochastic kernel mapping $Y \to Y \mid \mathfrak{S}$ now describes the importance of trade in explaining cross-country income distribution dynamics.[14]

Figure 8.1 displays the stochastic kernel for $Y \to Y \mid \mathfrak{S}$ with $\mathfrak{S}$ trade conditioning. Here, the counterclockwise twist in the kernel towards the vertical is even more pronounced than in Figure 7.1: rich countries trade mostly with other rich ones; and, interestingly, the very poorest countries, mostly with rich ones again. The way to see this is to notice that the stochastic kernel in Figure 8.1 clusters about 1 on the high end of the *Original* scale, but about a value less than 1/2 on the low end. Figure 8.2 and Figure 8.3 give the densities and boxplots for $Y \mid \mathfrak{S}$. In Figure 8.2 the second peak that emerges in 1988 is at the average 1. Figure 8.3 shows the conditional distributions to be relatively compact (i.e.,

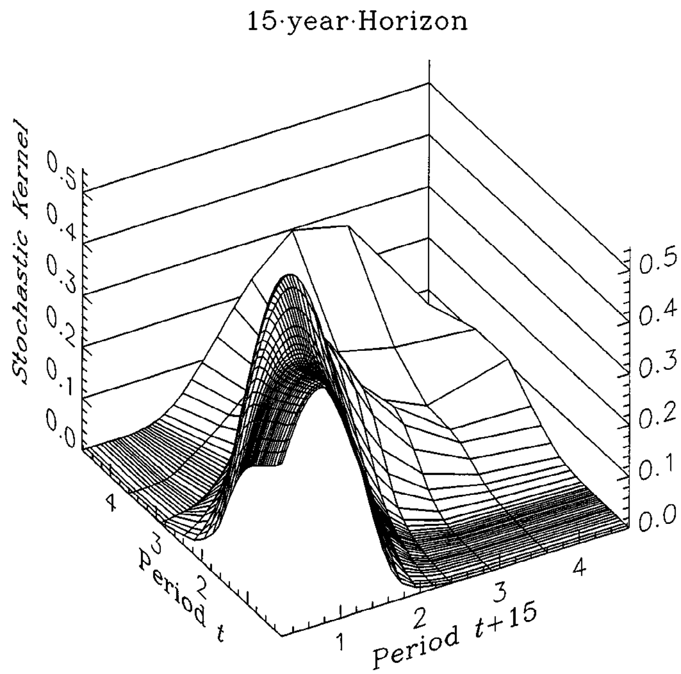


Figure 7.4. Spatial-conditioned relative income dynamics.

have their support relatively narrow); the vanishing of the middle-income class is no longer visible.

The conditioned income distributions give information on dynamics as well. Figure 7.4 and Figure 8.4 provide stochastic kernel representations on 15-year transitions in space- and trade-conditioned incomes. For space, Figure 7.4 shows a marked improvement in convergence possibilities—the poor catching up with the rich—except at the very highest income levels: the stochastic kernel is concentrated parallel to the Period t axis on the average value of 1. For trade, however, that increase in convergence dynamics is most obvious only for middle-income countries.

To summarize, this section has provided a set of empirical computations designed to explain the emerging twin-peaks dynamics earlier documented. Using the idea of an endogenously determined set of cross-section neighbors being important, this section developed the notion of conditioning schemes for empirical use. Applied to space and trade, we see a first glimpse of how important and large such cross-sectional interaction effects might be. The importance of trade here is not just expressed in blanket measures of how open an economy is; emphasized instead are patterns of who trades with whom.

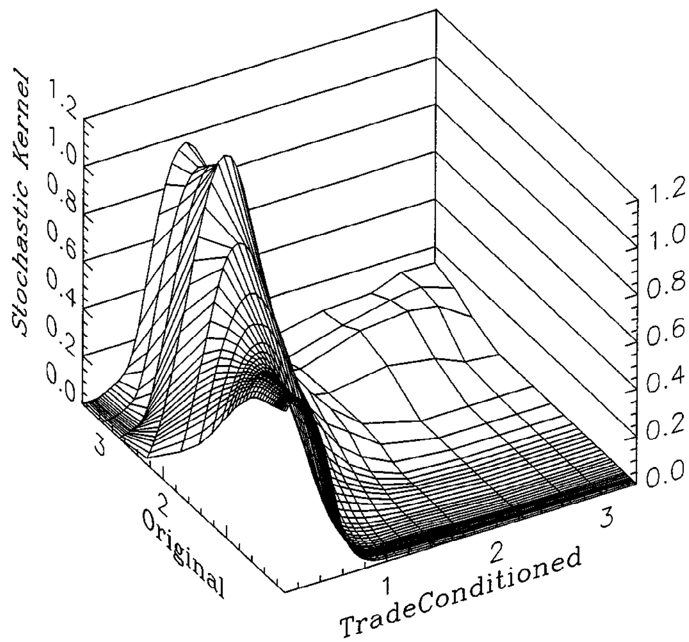


Figure 8.1. Stochastic kernel, relative (per capita) incomes across 105 countries.

5. Conclusions

This paper has analyzed patterns of economic growth across countries from the perspective of distribution dynamics. In doing so, it uncovered empirical regularities—emerging twin peaks, incipient polarization and stratification—that are hidden to traditional methods of empirical analysis.

Those distribution-dynamic features call for explanation. This paper has argued that such explanation is noteworthy in two significant respects: It will (i) differ from conventional models of growth and accumulation in the direction of theorizing in terms of the entire cross section distribution, and (ii) depart from standard techniques of econometric analysis—both in the empirical effects that have to be modeled and in permitting the cross-sectional interaction that the theoretical reasoning had suggested would matter.

The paper has developed a class of conditional distribution analyses using the idea of *conditioning schemes*. These showed the importance of space and trade—endogenous cross-sectional interaction more generally—for understanding cross-country patterns of growth. Unlike traditional studies on trade and growth, the study above emphasized not measures of openness to trade, but instead empirical patterns of who trades with whom. Although considerable progress has taken place, much remains to be done still in rigorous theoretical and empirical analyses of such cross-sectional dynamics.

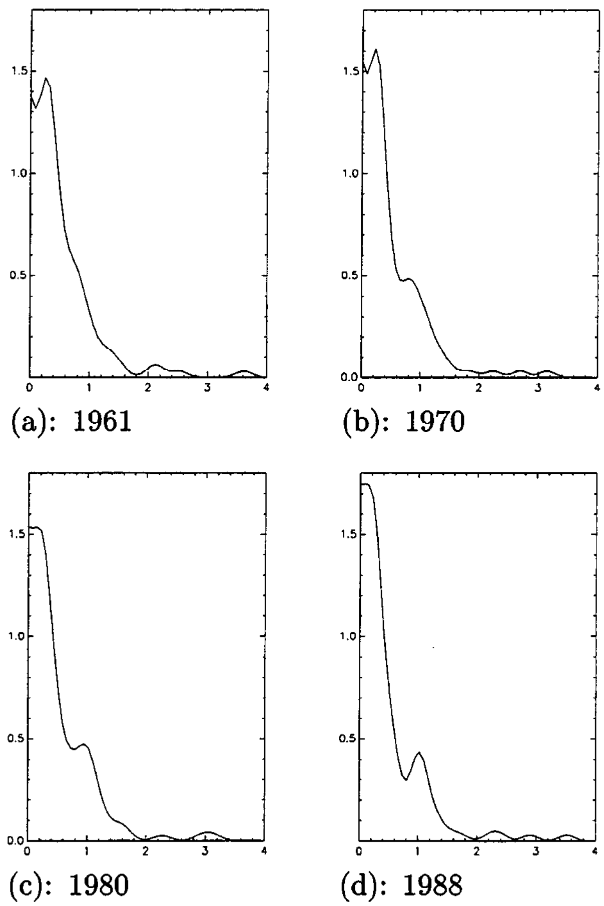


Figure 8.2. Densities of relative (per capita) incomes across 105 countries.

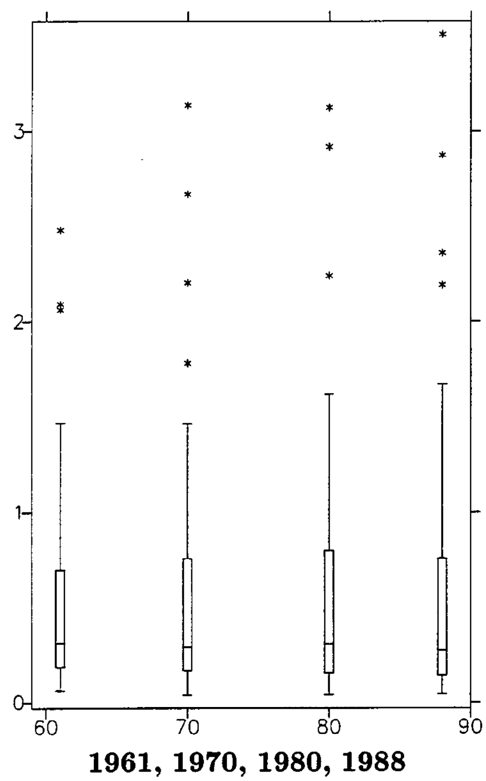


Figure 8.3. Boxplots, relative (per capita) incomes across 105 countries.

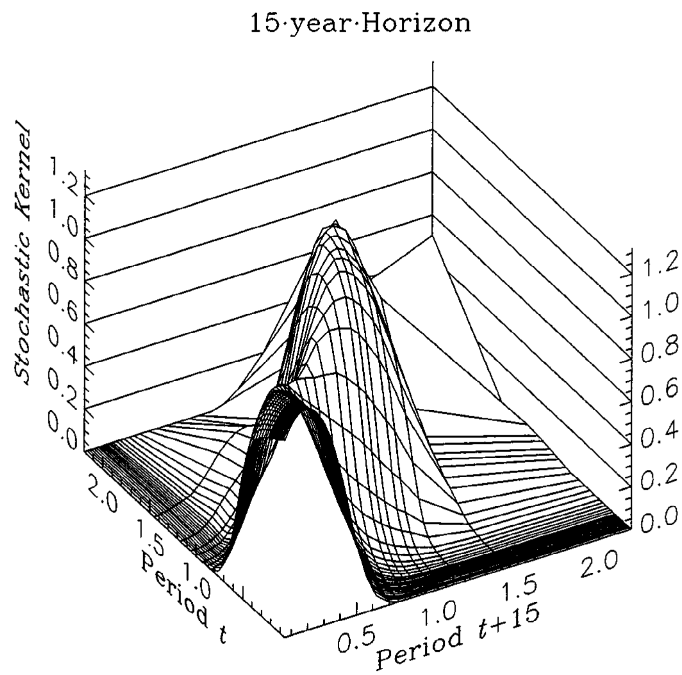


Figure 8.4. Trade-conditioned relative income dynamics.

Acknowledgments

I thank Luca Stanca for help, and conference participants at a MacArthur Foundation meeting and the Santa Fe Institute for constructive criticisms on related work. Discussions with Ken Arrow, Steven Durlauf, Yannis Ioannides, and Chuck Manski have been especially helpful. This paper was delivered as an Invited Lecture (Economics/Econometrics) at the 1996 Econometric Society European Meetings in Istanbul, where Rob Engle and Harald Uhlig provided many insightful comments. I have received helpful suggestions from numerous others, including Willem Buiter, Stephen Redding, Tom Sargent, Peter Sinclair, Jonathan Temple, and Adrian Wood. I thank the British Academy and the MacArthur Foundation for financial support. All calculations and graphs below were produced using the author's econometrics shell tSrF.

Notes

1. Even though I have just argued that interest should not be thus confined, I will continue to use the phrase "emerging twin-peaks" for two reasons: one, it is a convenient and evocative shorthand; and two, the cross-country data support the twin-peaks description.
2. The statements in this paragraph should not be taken as intending anything stronger than what they actually assert. They refer to the features of Figure 1, no more and no less.
3. Bliss (1996) and Quah (1996b) have also taken this perspective, although without the cross-sectional interaction that will figure prominently below.
4. Other papers relevant to this debate include Ben-David (1994), Bernard and Durlauf (1996), Canova and Marcet (1995), Desdoigts (1996), Durlauf and Johnson (1995), Galor (1996), and Jones (1997).
5. I took the data nonnegativity into account following the procedure and automatic bandwidth choice given in Silverman (2.10 and 3.4.2, 1986).
6. Bianchi (1995)—using bootstrap tests for multimodality related to ideas in Izenman and Sommer (1988) and Silverman (1981, 1983)—and Paap and van Dijk (1994)—using density mixture techniques—have provided statistical descriptions on this sequence of pictures, earlier given in Quah (1993b). Jones (1997) presents, in essence, the same picture. Cowell, Jenkins, and Litchfield (1996) have noted similar twin-peakedness in UK personal income distributions.
7. Although this is relatively unused in economics and econometrics, it is a familiar textbook object in statistics; see, e.g., Cleveland (1993).
8. Of course, there is no logical necessity why only 15-year horizons need be considered. Moreover, other statistics (i.e., real-valued functionals) of a stochastic kernel might also be considered, beyond just the visual description given in the text. For further exploration of these issues, see Durlauf and Johnson (1994), and Quah (1993a, b, 1996a, b) who studied kernels estimated over varying horizons, and ergodic characterizations, mobility indexes, and first passage-times calculated off estimated kernels. Desdoigts (1994), Lamo (1996), and Schluter (1997) have used related ideas in empirical research.
9. See, however, Islam (1995) and Nerlove (1996) for an opposing view.
10. Pursuing this cross-section interactions approach reflects a bias based on instinct, not necessarily anything more rigorous. There might well be simpler explanations for the twin peaks: e.g., distinct groups of economies having differing preferences, giving rise to differing investment rates, or even just underlying productivity shocks having a particular, bimodal distribution.
11. It is not essential that it be exactly the average $H_{\mathcal{C}}$ that affects $\tilde{\phi}$, just that it be some appropriate functional of the distribution of h's in the coalition $\mathcal{C}$.
12. If the denominator $\hat{Y}_l(t)$ were replaced by the exponential of the conditional expectation of $Y_l(t)$ conditioned on an information set $\mathcal{G}$, then $\tilde{Y}_l(t)$ is just the exponential of the expectations error $\log(Y_l(t)) - \log(\hat{Y}_l(t)) = \log(Y_l(t)) - E(\log(Y_l(t)) \mid \mathcal{G})$.

13. If we generalize beyond cross-country growth, this particular conditioning scheme provides a natural empirical counterpart to the theoretical effects described in Benabou (1996a), Durlauf (1996), Ioannides (1996), and others. It complements the empirical analysis of Brock and Durlauf (1995). Quah (1996e) had used a form of conditioning scheme to study aspects of globalization in Europe.
14. I have also experimented with using only imports or only exports in this definition and with T set to the beginning or middle rather than the end of the sample: not much changes in the conclusion. The conditioning scheme described here is more intricate than just measuring the openness of an economy—here, information on who trades with whom is used. Additional factor content data, as for example in Coe and Helpman (1995) or Eaton and Kortum (1996), might additionally be exploited to more sharply focus on the learning and technology components in cross-country interaction.

References

Atkinson, A. B.(1995). *Incomes and the Welfare State: Essays on Britain and Europe*. Cambridge: Cambridge University Press.

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

Baumol, W. J. (1986). "Productivity Growth, Convergence, and Welfare," *American Economic Review* 76, 1072–85.

Ben-David, D. (1994). "Convergence Clubs and Diverging Economies," Working Paper 922, CEPR.

Bénabou, R. (1996a). "Heterogeneity, Stratification, and Growth: Macroeconomic Implications of Community Structure and School Finance," *American Economic Review* 86, 584–609.

Bénabou, R. (1996b). "Inequality and Growth." In *Macroeconomics Annual*, Ben Bernanke and Julio Rotemberg (eds.), Cambridge, MA: NBER and MIT Press. Forthcoming.

Bernard, A. B., and S. N. Durlauf. (1996). "Interpreting Tests of the Convergence Hypothesis," *Journal of Econometrics* 71, 161–174.

Bianchi, M. (1995). "Testing of Convergence: A Bootstrap Test for Multimodality," Working Paper, Bank of England.

Bliss, C. J. (1996). "Long-run Wealth Distribution with Random Shocks," Working Paper 23, Oxford: Nuffield College.

Brock, W. A., and S. N. Durlauf. (1995). "Discrete Choice with Social Interactions," Working Paper, Madison: University of Wisconsin.

Canova, F., and A. Marcet. (1995). "The Poor Stay Poor: Nonconvergence Across Countries and Regions," Discussion Paper 1265, CEPR, November.

Chamberlain, G. (1984). "Panel Data." In Zvi Griliches and Michael D. Intriligator (eds.), *Handbook of Econometrics vol. II*. Amsterdam: Elsevier North-Holland.

Chung, K. L. (1960). *Markov Chains with Stationary Transition Probabilities*. Berlin: Springer-Verlag.

Cleveland, W. S. (1993). *Visualizing Data*. Summit, NJ: Hobart Press.

Coe, D. T., and E. Helpman. (1995). "International R&D Spillovers," *European Economic Review* 39, 859–887.

Cowell, F. A., S. P. Jenkins, and J. A. Litchfield. (1996). "The Changing Shape of the UK Income Distribution: Kernel Density Estimates." In John Hills (ed.), *New Inequalities: The Changing Distribution of Income and Wealth in the United Kingdom*. Cambridge: Cambridge University Press.

Desdoigts, A. (1994). "Changes in the World Income Distribution: A Non-parametric Approach to Challenge the Neoclassical Convergence Argument," PhD dissertation, Florence: European University Institute.

Desdoigts, A. (1996). "Determining Development Patterns and "Club" Formation Using Projection Pursuit Techniques," Working Paper, ECARE, ULB.

Durlauf, S. N. (1993). "Nonergodic Economic Growth," *Review of Economic Studies* 60, 349–366.

Durlauf, S. N. (1996). "A Theory of Persistent Income Inequality," *Journal of Economic Growth* 1, 75–93.

Durlauf, S. N., and P. Johnson. (1994). "Nonlinearities in Intergenerational Income Mobility," Working Paper, University of Wisconsin, Economics Department.

Durlauf, S. (1995). "Multiple Regimes and Cross-country Growth Behavior," *Journal of Applied Econometrics* 10, 365–384.

Eaton, J., and S. Kortum. (1996). "Trade in Ideas: Patenting and Productivity in the OECD," *Journal of International Economics* 40, 251–278.

Efron, B., and R. J. Tibshirani. (1993). *An Introduction to the Bootstrap*. New York: Chapman and Hall.

Esteban, J.-M., and D. Ray. (1994). “On the Measurement of Polarization,” *Econometrica* 62, 819–851.
Friedman, M. (1992). “Do Old Fallacies Ever Die?” *Journal of Economic Literature* 30, 2129–2132.
Galor, O. (1996). “Convergence? Inferences from Theoretical Models,” *Economic Journal* 106, 1056–1080.
Galor, O., and J. Zeira. (1993). “Income Distribution and Macroeconomics,” *Review of Economic Studies* 60, 35–52.
Ioannides, Y. M. (1990). “Trading Uncertainty and Market Form,” *International Economic Review* 31, 619–638.
Ioannides, Y. M. (1996). “Residential Neighborhood Effects and Inequality,” Working Paper, Economics Department, Tufts University.
Islam, N. (1995). “Growth Empirics: A Panel Data Approach,” *Quarterly Journal of Economics* 110, 1127–1170.
Izenman, A. J., and C. J. Sommer. (1988). “Philatelic mixtures and multimodal densities,” *Journal of the American Statistical Association* 83, 941–953.
Jones, C. I. “On the Evolution of the World Income Distribution,” *Journal of Economic Perspectives*. Forthcoming.
Kirman, A. P., C. Oddou, and S. Weber. (1986). “Stochastic Communication and Coalition Formation,” *Econometrica* 54, 129–138.
Konings, J. (1994). “Gross Job Flows and Wage Determination in the UK: Evidence from Firm-level Data,” PhD dissertation, LSE.
Koopmans, R. M. (1995). “Asymmetric Industry Structures: Multiple Technologies, Firm Dynamics, and Profitability,” PhD dissertation, LSE.
Lamo, A. R. (1996). “Cross-section Distribution Dynamics,” PhD dissertation, LSE.
Leung, C., and D. T. Quah. (1996). “Convergence, Endogenous Growth, and Productivity Disturbances,” *Journal of Monetary Economics*, 38, 535–547.
Lillard, L. A., and R. J. Willis. (1978). “Dynamic Aspects of Earning Mobility,” *Econometrica* 46, 985–1012.
Loury, G. C. (1981). “Intergenerational Transfers and the Distribution of Earnings,” *Econometrica* 49, 843–867.
Lucas, R. E., Jr., (1978). “On the Size Distribution of Business Firms,” *Rand Journal of Economics* 9, 508–523.
Nerlove, M. (1996). “Growth Rate Convergence: Fact of Artifact?” Working Paper, University of Maryland.
Paap, R., and H. K. van Dijk. (1994). “Distribution and Mobility of Wealth of Nations,” Working Paper, Tinbergen Institute, Erasmus University.
Persson, T., and G. Tabellini. (1994). “Is Inequality Harmful for Growth?” *American Economic Review* 84, 600-621.
Quah, D. T. (1993a). “Empirical Cross-section Dynamics in Economic Growth,” *European Economic Review* 37, 426–434.
Quah, D. T. (1993b). “Galton’s Fallacy and Tests of the Convergence Hypothesis,” *The Scandinavian Journal of Economics* 95, 427–443.
Quah, D. T. (1996a). “Aggregate and Regional Disaggregate Fluctuations,” *Empirical Economics* 21, 137–159.
Quah, D. T. (1996b). “Convergence Empirics Across Economies with (Some) Capital Mobility,” *Journal of Economic Growth* 1, 95–124.
Quah, D. T. (1996c). “Empirics for Economic Growth and Convergence,” *European Economic Review* 40, 1353–1375.
Quah, D. T. (1996d).“Ideas Determining Convergence Clubs,” Working Paper, Economics Department, LSE.
Quah, D. T. (1996e). “Regional Convergence Clusters Across Europe,” *European Economic Review* 40, 951–958.
Quah, D. T. (1996f). “Twin Peaks: Growth and Convergence in Models of Distribution Dynamics,” *Economic Journal* 106, 1045–1055.
Sala-i-Martin, X. (1996). “Regional Cohesion: Evidence and Theories of Regional Growth and Convergence,” *European Economic Review* 40, 1325–1352.
Schluter, C. (1997). “Topics in Distributional Analysis,” PhD dissertation, LSE.
Shorrocks, A. F. (1978). “The Measurement of Mobility,” *Econometrica* 46, 1013–1024.
Silverman, B. W. (1981). “Using Kernel Density Estimates to Investigate Multimodality,” *Journal of the Royal Statistical Society, Series B* 43, 97–99.
Silverman, B. W. (1983). “Some Properties of a Test for Multimodality Based on Kernel Density Estimates.” In J. F. C. Kingman and G. E. H. Reuter, (eds.), *Probability, Statistics, and Analysis*. vol. 79. *London Mathematical Society Lecture Note Series* Cambridge: Cambridge University Press.
Silverman, B. W. (1986). *Density Estimation for Statistics and Data Analysis*. New York, NY: Chapman and Hall.
Singer, B., and S. Spilerman. (1976). “Some Methodological Issues in the Analysis of Longitudinal Surveys,” *Annals of Economic and Social Measurement* 5, 447–474.
Summers, R., and A. Heston. (1991). “The Penn World Table (Mark 5): An Expanded Set of International Comparisons, 1950–1988,” *Quarterly Journal of Economics* 106, 327–368.

Sutton, J. (1995). "The Size Distribution of Business, Parts I and II," Discussion Paper EI/9, 10, STICERD, LSE.
Townsend, R. M. (1983). "Theories of Intermediated Structures," *Carnegie-Rochester Conference Series on Public Policy* 18, 221–272.