



ACADEMIC
PRESS

Available online at www.sciencedirect.com

SCIENCE @ DIRECT®

Explorations in Economic History 40 (2003) 78–97

Explorations in
Economic History

www.elsevier.com/locate/eeh

Distribution dynamics: stratification, polarization, and convergence among OECD economies, 1870–1992[☆]

Philip Epstein, Peter Howlett,^{*} and Max-Stephan Schulze

Department of Economic History, London School of Economics, Houghton Street, London WC2A 2AE, UK

Received 28 November 2000

Abstract

Using a distribution dynamics approach, the growth experience of 17 OECD economies is investigated. After explaining the distribution dynamics approach, the empirical analysis examines both the observed period dynamics and the unique long-run equilibrium associated with three periods. This study suggests persistence and stratification, not convergence, characterized the pre-1914 regime, whereas convergence was the key feature of the post-war regime. However, a larger sample of OECD economies in the post-war period indicates that convergence was primarily a feature of the Golden Age and in the post-Golden Age period separation, polarization, and divergence came to the fore.

© 2002 Elsevier Science (USA). All rights reserved.

Keywords: Distribution dynamics; Economic growth; Convergence; OECD

[☆]The authors thank participants at the Third Conference of the European Historical Economics Society (Lisbon, 1999) and the Fourth World Congress of Cliometrics (Montreal, 2000) for their comments. Helpful suggestions by two anonymous referees and the editor of this journal are gratefully acknowledged. All remaining errors and omissions are the responsibility of the authors.

^{*}Corresponding author.

E-mail addresses: w.p.howlett@lse.ac.uk (P. Howlett), m.s.schulze@lse.ac.uk (M.-S. Schulze).

1. Introduction

Since the 1980s the debate about economic convergence has dominated empirical work about the dynamics of growth.¹ Economic historians have been attracted, in particular, by stories of club convergence.² However, the analytical foundations of most of the work in this area have rested on linear, or more usually log-linear, regression analysis. Thus, the results tend to be dependent on a conditional average, in which time is the dominant factor, a surprising outcome because space and distributional issues have long been important to both theorists and historians. A notable exception to the ‘regression school’ has been the work on distribution dynamics pioneered in a series of papers by Quah (1993, 1996, 1997). He believes that only by considering the issues of growth and distribution simultaneously can we understand their underlying dynamics. He has argued, for example, that there is no simple causal relationship between the concepts of β -convergence (where initially poor economies tend to grow faster over time than rich ones) and σ -convergence (where the dispersion in income levels across a set of economies tends to decrease over time) (Barro and Sala-I-Martin, 1991). Moreover, similar stories of global (or club) convergence may be driven by very different stories of individual economy mobility. This approach should appeal to economic historians—both because it can encompass a rich diversity of individual economy experience and because it emphasizes that same diversity.

Here we consider the experience of some of the leading OECD economies since 1870 within an explicit distribution dynamics framework. It is important to distinguish between the empirically observed dynamics of the international cross-sectional income distribution and its steady state solution (long-run equilibrium). Although the two are obviously related, the long-run equilibrium is not always apparent from the distributional characteristics demonstrated by the empirically observed dynamics during a given period. When discussing the latter, therefore, the focus is on mobility (the movement of economies between different income levels) and tendencies such as clustering (where groups of economies cluster around a certain income level) and stratification (where economies move into distinct income-level strata). Convergence

¹ Some of the more influential papers in this area include Solow (1956), Baumol (1986), Romer (1986), Barro and Sala-I-Martin (1991), Mankiw et al. (1992), Bernard and Durlauf (1995). Temple (1999) provides a review of more recent theoretical and empirical work by economists in this area. The empirical work he surveys is primarily concerned with the post-1960 period (using the Heston and Summers data). The most important historical research in this area has been done by Williamson and O’Rourke who focused on real wage (and other factor price) convergence, particularly in the period 1870–1913; this work is best captured in Williamson (1996) and O’Rourke and Williamson (1999).

² Examples of those who have told explicit club convergence stories are Tortella (1994) and Toniolo (1998). A summary of the literature can be found in Broadberry (1996). A notable exception to the club convergence approach taken by historians is Mills and Crafts (2000). A recent empirical article, using post-1960 data from the Penn World Tables, while pointing out the lack of convergence between rich and poor countries also identified so many convergence clubs (‘for the 112 countries in our sample ... we find 63 asymptotically perfect convergence clubs and 42 asymptotically relative clubs’) as to suggest that the term is effectively useless as an analytical tool (Hobijn and Franses, 2000).

(or divergence), on the other hand, is viewed as the outcome of a long run process and is captured in the shape of the steady state distribution.

Our main findings suggest that there is no strong evidence of distributional convergence among the leading 17 OECD economies before 1950. Furthermore, a more detailed study of the dynamics of an expanded set of 24 economies would indicate that post-World War II convergence was a characteristic only of the Golden Age (1950–1973) and that in the post-Golden Age economies diverged again, clustering around a pole of the relatively rich and another of the relatively poor.

2. Scope and data

Economic historians often discuss long-term growth in terms of regimes, epochs or phases. For example, Williamson (1996) in discussing factor price convergence among the OECD club identifies three epochs: (1) the late nineteenth century, characterized by fast growth, globalization, and convergence; (2) 1914–1950, which witnessed slow growth, de-globalization, and divergence; and (3) the post-1950 era, which has experienced fast growth, globalization and convergence. This periodization is similar to that proposed by Maddison (1995, pp. 59–87). Maddison argues that since 1820 there have been five distinct phases of development: 1820–1870, 1870–1913, 1913–1950, 1950–1973, and 1973–1992.³ The 1870–1913 phase is, according to Maddison, characterized as ‘a relatively peaceful and prosperous era’ in which ‘per capita growth accelerated in all regions and in most countries.’ This phase of growth eventually gave way to ‘an era deeply disturbed by war, depression, and beggar-your-neighbour policies...a bleak age, whose potential for accelerated growth was frustrated by a series of disasters.’ Next came ‘a golden age [1950–1973] of unparalleled prosperity’ in which income per head in all regions ‘grew faster than in any other phase.’ Finally, in the period after 1973 inflationary pressures, the breakdown of the Bretton Woods fixed exchange rate system, and the oil price shocks brought about ‘a sharp reduction in the pace of economic growth throughout the world...and the momentum of the golden age’ was lost. For analytical purposes we begin by following Williamson’s division of the period 1870–1992 into three sub-periods: 1870–1913 (the pre-1914 period), 1914–1950 (the trans-war period), and 1951–1992 (the post-war period).

For these periods we consider 17 advanced OECD economies (in their post-World War II boundaries) for which homogenous, annual long-run historical data are available: Australia, Austria, Belgium, Canada, Denmark, Finland, France, West Germany, Italy, Japan, Netherlands, New Zealand, Norway, Spain, Sweden, United Kingdom, and the United States. Maddison’s data set (1995) is the primary source

³ He further sub-divided the period 1913–1950 into four sub-periods (1913–1929, 1929–1938, 1938–1944, and 1944–1949).

for GDP per capita levels that are measured in 1990 international dollars.⁴ By definition this is a biased grouping, representing 17 of the most developed nations in the world at the end of the 20th century. If the theoretical literature is correct, the forces of convergence should be very strong among this particular group.

Maddison and much of the current historiography discuss the post-war period in terms of the 'Golden Age' (1951–1973) and 'post-Golden Age' (1974–1992). Data limitations mean that we cannot employ the full range of techniques we wish to use drawing on just the 17 advanced countries. Thus to be able to investigate the Golden Age and post-Golden Age periods we have considered a larger group of 24 economies. The additional economies are Czechoslovakia, Greece, Hungary, Ireland, Portugal, Switzerland, and Turkey. The addition of these seven economies will increase the variance compared to the group of 17 advanced economies, as they are a poorer group of economies. For example, in 1972 the average real GDP per capita in the seven additional economies was 53% of the average level of real GDP per capita in the 17 advanced economies.⁵

For each year we standardized the observations to the average level of real GDP per head among the 17 (or 24) advanced economies in that year (with the average taking a value of 1.00).

Convenient though they are, the Maddison estimates entail potentially substantial index number problems. In essence, these estimates involve the backward projection of 1990 purchasing power parity adjusted GDP levels using volume indices or growth rates of product per capita. Most recently, Prados de la Escosura (2000) pointed to the distortions that the use of single benchmarks such as PPP-adjusted 1990 dollars can introduce into the measurement of economies' relative income position over time. Taking up Eichengreen's suggestion, he makes the case for using a short-cut method to derive current price estimates of PPP-adjusted GDP for economies and periods in which aggregate PPPs are not available. He shows that this procedure yields different rankings than conventional constant price, fixed end-year estimates. This approach emphasizes that economic agents react to current rather than constant prices. However, Prados-type current price PPP-adjusted GDP estimates are so far available only for certain benchmark years, not on a continuous annual basis. Hence they cannot serve as an alternative data set in distribution dynamics analysis as presented here.

By relying on Maddison's GDP data, we focus on data that has been used most widely in the historical literature to examine the pattern and periodization of long-run (income) convergence.⁶ For unlike the Summers and Heston data set, much used

⁴ Our group of countries is the same as the group of 17 advanced economies defined by Maddison except that Switzerland (for which annual data is only available after 1900) is replaced by Spain, using the estimates of Prados de la Escosura (1995). We also include new GDP estimates for nineteenth century Austria (Schulze, 2000) and have closed the gaps in the Maddison series for Japan using the work of Ohkawa and his collaborators (Ohkawa et al., 1979; Japan Statistical Association, 1998).

⁵ In 1992, after the collapse of the communist regimes of Eastern Europe, this figure was reduced only slightly, to 52%.

⁶ For example, Broadberry (1996), Mills and Crafts (2000), Toniolo (1998), and Williamson (1996) for comparisons with his real wage data. Other influential papers that rely on Maddison's long-run data include Bernard and Durlauf (1995) and Baumol (1986).

in the empirical economics literature, Maddison's comparative estimates extend back into the nineteenth century. However, in an effort to gauge the extent to which the findings obtained here are sensitive to the choice of GDP indices used, we compared the computations for the Maddison set with those for the Summers and Heston data for matching years and economies. The results indicate no material difference.⁷

3. Some traditional empirics

For any given group of economies, σ -convergence implies that over time the variation in their incomes relative to their mean income will decline. Fig. 1 uses box-plots of the standardized income levels to illustrate the empirics of this in order to provide some background for the discussion of distribution dynamics.⁸ It is clear from Fig. 1a that there was significant σ -convergence across the 17 economies over the period as a whole and that this process accelerated in the post-war period.⁹ However, these changes in income dispersion were not the result of a simple linear force.

For example, between 1873 and 1893 the interquartile range actually increased and convergence in this period would appear to have been driven by a significant reduction at the upper and lower tails of the distribution. Further, it is noticeable that the process of σ -convergence seemed to have received little further impetus in the very late 19th and early 20th centuries, the years with generally faster economic growth than in the two preceding decades. The major change between 1893 and 1913 is that the interquartile range returned to its 1873 range. Perhaps surprisingly, neither the protectionist trade regimes of the 1920s, nor the great depression and the rise of protectionist trade blocs in the 1930s appear to have made much of an adverse impact on this process. Indeed, if anything the box-plots suggest that in the midst of the 1930s depression, the dispersion of income levels was less than it had been in 1913 (again, this is driven mainly by reduction at the upper and lower tails of the distribution). The main short-term impact of the Second World War was to increase dispersion, especially at the upper and lower ends of the distribution.¹⁰ This was a mere prelude, however, to the Golden Age, which saw a dramatic collapse in income

⁷ Details are available from the authors.

⁸ There are three main components in a box-plot: the box, the whiskers, and the outliers. The box represents the interquartile range (IQ), that is the distance between the 25th and 75th percentiles, with the line in the box showing the median income. The whiskers represent upper and lower adjacent values, whereby the upper (lower) value is the largest (smallest) data point less (greater) than or equal to the 75th (25th) percentile value plus $1.5 \times \text{IQ}$. The outliers are values that are more extreme than the upper or lower adjacent values and are plotted individually.

⁹ This is also reflected in a dramatic decrease in the ratio between the highest and lowest income observed in each year. In 1873 the ratio was 6.1, in 1893 and 1913 4.1 (and in each of these pre-1914 years this represented Australia/Japan), in 1933 2.5 (Denmark/Japan), in 1953 (after the war shock) 4.5 (USA/Japan), in 1973 2.1 (USA/Spain), and in 1992 1.7 (USA/Spain). The success of the Japanese economy was such that after having been the poorest of the 17 economies until 1958 by 1992 it was the second richest economy.

¹⁰ Its other, longer-term, impact was to greatly increase the lead of the USA over the other economies.

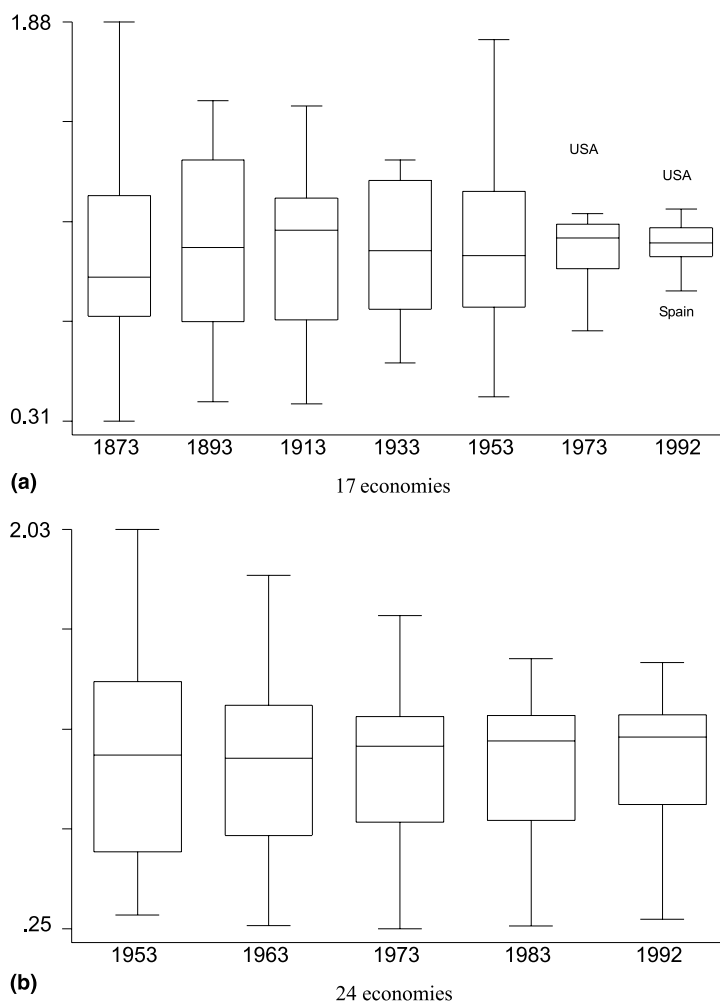


Fig. 1. The dispersion of income, selected years. *Note.* The vertical axis measures real GDP per capita, standardised on the average level of real GDP per capita of the economies in the sample in the given year: (a) 17 economies; (b) 24 economies.

dispersion. In this period there is unambiguous σ -convergence, a process that continued in the post-Golden Age. Finally, it should be noted that despite this convergence in the post-war period we also see the only evidence of outliers with the USA as a rich outlier in both 1973 and 1992 and Spain in 1992.

Fig. 1b shows that for the 24 economies there is also a sharp decline in income dispersion in the post-war period. However, this decline comes to an end in the mid-1970s and thereafter dispersion is relatively stable. This provides some empirical justification in treating the Golden Age and post-Golden Age as two separate regimes.

There is an important element of the historical story that is hidden in the box-plots (and in more conventional measures of σ -convergence, such as the coefficient

of variation of the income levels or the standard deviation of the logarithm of income), that of intra-distribution movement. In order to illustrate the importance of mobility within the distribution, we consider briefly the actual empirical distributions associated with the economies under discussion. Initially we distinguish five income states: state 1 representing the lowest level and state 5 the highest level of income (note that these income states also form the basis for the computation of transition probabilities as set out below).¹¹ Fig. 2a shows the actual distribution of the economies at some key dates. The horizontal axis shows the income states whilst the vertical axis measures the frequency of economies in the given income state.¹² Thus, the bold lines suggest the shape of the distribution. The matrices, on the other hand, represent mobility; they show the movements of individual economies across income states between each pair of dates. For example, of the five economies in income state 2 in 1870, by 1913 one economy had fallen into income state 1, one had risen to income state 3, and another had risen to income state 4, leaving two economies that remained in income state 2.

For the 17 economies, Fig. 2a reveals a variety of experiences across the different periods. It suggests that σ -convergence tells us too little about the dynamics of income distribution over both time and space and conceals important historical issues. For example, despite the strong case made in the historical literature for the significance of the forces of convergence between 1870 and 1913, the empirical distributions suggest that mobility was less common between those two dates than between 1913 and 1950 or 1950 and 1992. Thus, only seven of the 17 economies experienced mobility across income states whereas in the other two periods more than half of the economies experienced mobility. Mobility between 1870 and 1913 was also limited, being confined to a movement of only one income state (with the exception of Canada which rose from income state 2 to income state 4). It was also concentrated in the middle income states. Movement was most dramatic between 1950 and 1992 when two-thirds of the economies experiencing mobility moved more than one income state. Furthermore, there were also two cases of maximum mobility: Japan moved from income state 1 to income state 5 whereas New Zealand moved

¹¹ The choice of the number of income states is arbitrary but using three income states (rich, middle income, and poor economies) gives a rather stark view of the world and using seven income states (given the limited number of economies in our set) leaves the possibility of finding, in any particular year, an income state (or states) with no observations. However, we have carried out some sensitivity analysis to consider how alternative numbers of income states affect the results. They show little change. The income ranges representing the five states are not imposed by the researchers but are derived on purely empirical grounds. First, all the annual standardized income observations are treated effectively as a single cross-section. The observations in this 'cross-section' are then ranked from the lowest to the highest observation and split into five equal states: each state contains the same number of observations. This gives us the values for the partition for each state. It also means that the size and values of the states is different for each sub-period, since the population is different for each sub-period. Alternative definitions of the income states are also possible and will be investigated in future research. We use Quah's time series random field (TSRF) which is an econometric shell that permits the calculation of transition probability matrices and ergodic distributions.

¹² In effect, the bold lines are histograms collapsed to their midpoints.

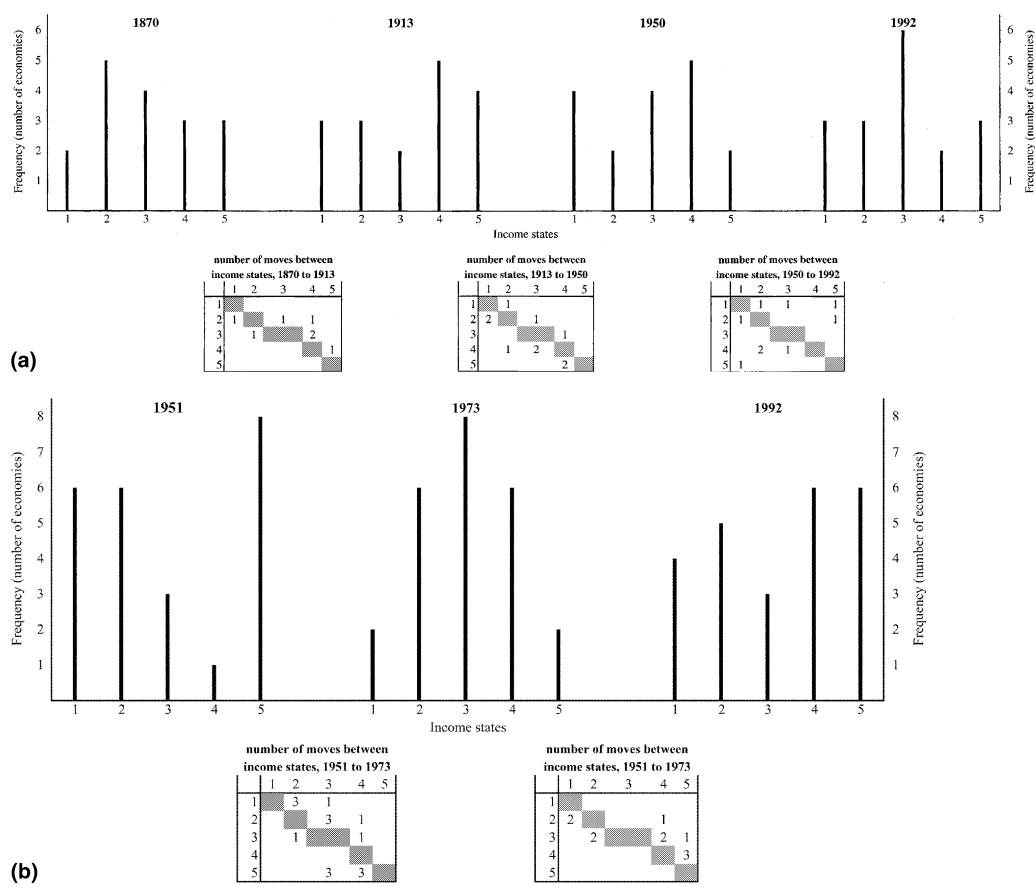


Fig. 2. (a) Empirical distribution and mobility by income state for 17 OECD economies: 1870, 1913, 1950, and 1992. (b) Empirical distribution and mobility by income for 24 OECD economies: 1951, 1973, and 1992. *Note.* The figure shows the empirical distribution of the 17 economies by income state for the selected years. The matrices show the number of economies which changed income state between each pair of selected years.

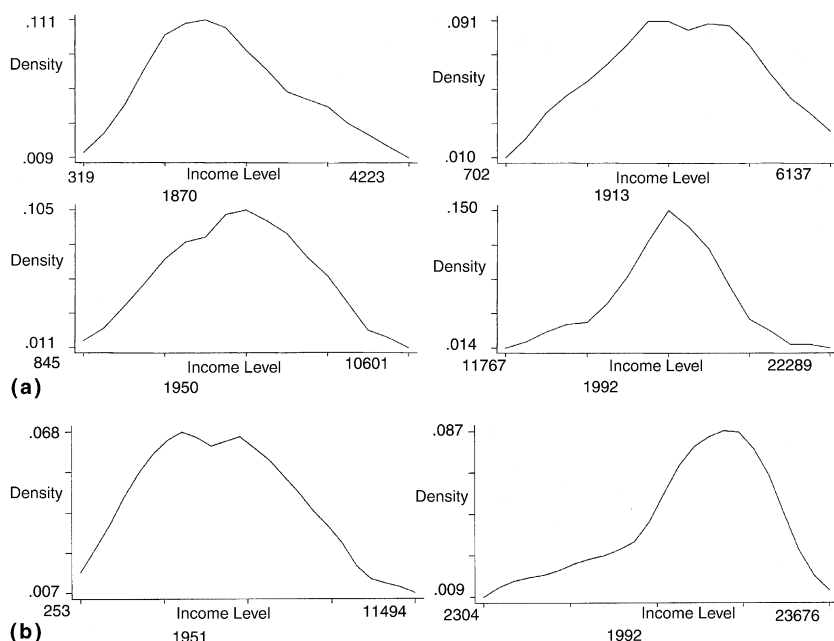


Fig. 3. (a) Densities of real GDP per capita in 17 OECD economies, 1870–1992. (b) Densities of real GDP per capita in 24 economies, 1951–1992.

in the opposite direction. In considering the direction of movements in Fig. 2a we find that between 1870 and 1913 twice as many economies moved to a higher income state than moved to a lower income state. This contrasts with comparisons of 1913 to 1950 and of 1950 to 1992. In both of these cases, economies that moved to a lower income state out-numbered those that moved to a higher income state.

Turning to the 24 economies, Fig. 2b shows that there was more movement between 1951 and 1973 than there was between 1973 and 1992 (respectively, 16 and 11 economies). Yet in both periods more economies moved to a higher income state than moved to a lower income state and most of these moves were of one income state only. However, between 1951 and 1973 six of the seven downwardly mobile economies were originally in income state 5 and between 1973 and 1992 all of the downwardly mobile economies were originally in income states 2 or 3. Thus studied in conjunction with the evidence on income dispersion in Fig. 1b, the mobility patterns reflected in Fig. 2b suggest that the notion of convergence, and the forces that underlie it, is more complex than traditional stories reveal.

Fig. 2 also shows how the shape of the distribution changes across each of the periods. Another way of considering these snapshots of the empirical distributions is through the graphical representation of a kernel density estimator.¹³ This has the

¹³ For an explanation of the kernel density estimator see the Technical appendix. We have used the Epanechnikov kernel.

advantage over a histogram representation in that it gives a smooth estimate. Fig. 3 presents kernel densities for the 17 and the 24 economies. Fig. 3a shows that 1992 is the clearest peak. This was the period in which the variation in incomes was also smallest. In Fig. 3b we also see a move from a relatively flat distribution in 1950 to a more clearly peaked (if skewed) distribution in 1992.

Although the empirical distributions (and their associated kernel estimates) have some merit, they must be used with great care. For example, we would need to be sure that the years chosen for comparison were not atypical or affected by short-term shocks. Furthermore, if we were trying to assess the impact of a particular regime on the distribution dynamics or trying to assess what the long-run equilibrium associated with a particular regime would be, the approach taken above is far too limited and too crude. It tells us something only about relative income positions in a particular year but nothing about the properties of the regime (or inherent tendencies that may or may not make for income convergence) as can be deduced from the analysis of the full set of annual income data. Herein lies the rationale for adopting distribution dynamics analysis. This is discussed below in terms of the mobility and the long-term characteristics of the dynamic distribution.

4. Dynamic distribution: representation

The main issue here is whether relatively rich and poor countries remain relatively rich and poor countries over time. One way of measuring this would be to define a set number of income states (for example, rich, middle income, and poor) and then to count the number of transitions out of one income state into another income state. This procedure could then be formalized into a transition probability matrix which allows us to discuss the degree of mobility and persistence within the distribution.¹⁴ Another approach, which we adopt here, involves the estimation of stochastic kernels. A stochastic kernel (and its related contour plot) is a graphical representation of the transition probabilities which has the advantage that it does not rely on a fixed number of discrete states but instead estimates a generalized form of the transition probability matrix which renders the state space continuous. It so helps avoid the potential problem of the results being sensitive to the (arbitrary) number of discrete income states chosen. Like the transition probability matrix, the stochastic kernel can be generated for any length of transition period. We have chosen 5-year transitions throughout on practical grounds.¹⁵

¹⁴ For an example of this approach see Epstein et al. (1999).

¹⁵ Over a shorter, 1-year transition period, for example, mobility can be expected to be very limited and emerging patterns would be more difficult to trace. A significantly longer horizon, say 10 or 15 years, would reduce the number of observations available for the estimation of the kernel. Here, the use of 5-year transitions aids consistency, for they can be readily computed for all sub-periods considered (including the shorter Golden Age and post-Golden Age sub-periods).

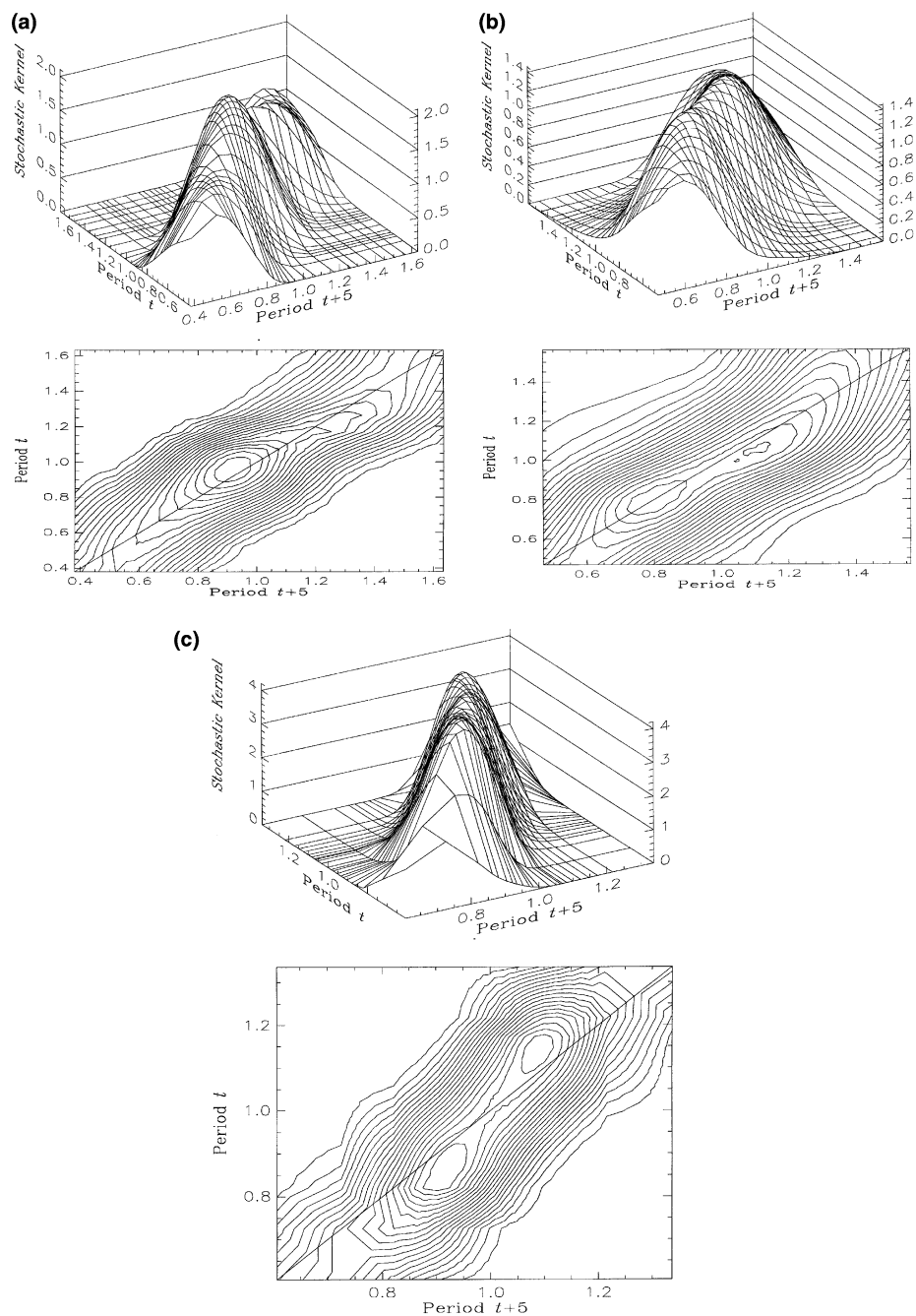


Fig. 4. (a) Seventeen advanced economies, 1870–1913. (b) Seventeen advanced economies, 1914–1950. (c) Seventeen advanced economies, 1951–1992.

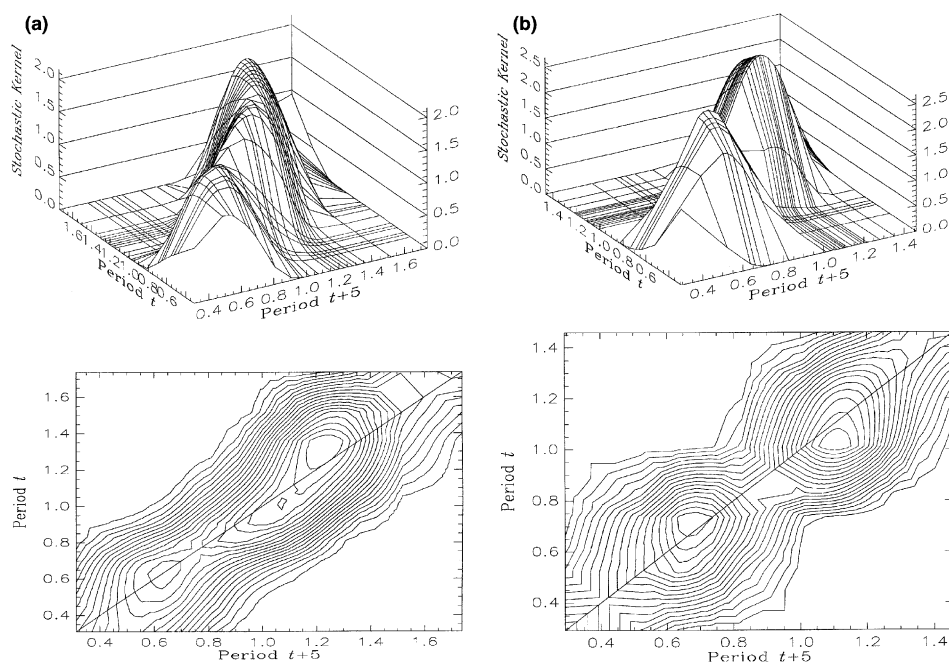


Fig. 5. (a) Twenty-four economies, 1951–1973. (b) Twenty-four economies, 1974–1992.

Fig. 4 considers the stochastic kernel and contour plot for 5-year transitions in our relative income data for the 17 economies, averaging over each of the three regime periods. Fig. 5 presents the same for the 24 economies in the two post-war periods.¹⁶ Thus in each case, the relative income of each economy in any given year t is periodically compared to its relative income in year $t + 5$ over each year in the sample period under review.

How do we interpret the stochastic kernel? The stochastic kernel provides evidence about mobility and persistence in the empirically observed distribution. Figs. 4 and 5 show how the cross-sectional distribution at time t evolves into that at time $t + 5$. The horizontal axes (for period t and period $t + 5$) give relative income, with 1.0 representing the standardized average level of income. Thus, a movement from right to left along the period t horizontal axis, or from left to right along the period $t + 5$ horizontal axis, represent increasing relative income. Slicing vertically through the kernel from any point on the period t axis extending parallel to the period $t + 5$ axis gives a probability density function that describes transitions over 5 years from a given relative income in period t .¹⁷ This is captured on the vertical axis whose scale,

¹⁶ We also compared these 5-year to the 1-year transition stochastic kernels for each regime and found them to be robust (in the sense that there was no significant difference between the two different kernels).

¹⁷ The profile traced out by this slice is non-negative and integrates to unity and is similar to a row of a transition probability matrix.

however, is unbounded. Thus two characteristics of the stochastic kernel help to reveal patterns of distributional mobility: its location and the shape of its surface.

Mobility and persistence can first be assessed by asking how the stochastic kernel lies relative to the 45° diagonal. This is more conveniently captured in a contour plot, a view from above on the stochastic kernel where contours have been drawn at the indicated relative income levels and projected onto the base of the graph. If most of the stochastic kernel were concentrated along this diagonal then mobility is low and there is little change in the cross-section distribution: economies' relative income in period $t + 5$ has not changed significantly since period t (the relatively rich remain rich and the relatively poor remain poor)—an example of persistence.¹⁸ If, on the other hand, most of the mass of the stochastic kernel had rotated around the 45° diagonal then this would indicate substantial changes in the distribution and a high degree of mobility.¹⁹ A counter-clockwise movement around the diagonal would represent a situation in which, relatively speaking, the rich were becoming poorer and the poor were becoming richer, periodically over 5-year horizons, thus indicating a tendency towards income equalization. At the extreme, this might take the form of over-taking with rich countries becoming poor and poor countries becoming rich. A clockwise movement would indicate the reverse: that the rich were becoming richer and the poor were becoming poorer, thus suggesting that forces of divergence were potentially more powerful.

The surface shape of the stochastic kernel (or contours in the corresponding contour plot) tells us about probabilities of transition from given relative incomes in t to different relative incomes in $t + 5$. In other words, a peak reflects a (comparatively) large number of observed transitions from one particular part of the distribution to another. It thus provides evidence on clustering over a 5-year horizon. There may be more than one peak if different economies' transitions cluster in different parts of the distribution (or around different income poles). For example, in the classic twin peaks story, polarization would be expressed as clustering of transitions around a low income pole and a high income pole.²⁰ Furthermore, if this were also associated with a dip in the middle of the stochastic kernel this would suggest that separation was an important underlying characteristic: middle income economies move into either high or low income parts of the distribution.

5. Dynamic distribution: empirics

What does Fig. 4 reveal about the distribution dynamics of the 17 advanced economies under the three different regimes? For the pre-1914 period, the main message is

¹⁸ Obviously, with relatively short transition periods one would expect to find that most of the stochastic kernel would be concentrated along the 45° diagonal.

¹⁹ Given the time horizon the kernel has been estimated for, we would not expect there to be major rotations.

²⁰ Polarization is a complex concept but Esteban and Ray (1994) explain those complexities clearly and methodically.

distributional persistence. Most of the mass of the stochastic kernel is concentrated along the 45° diagonal, which suggests that mobility was not a significant factor over the 5-year horizon. There is some limited counter-clockwise movement in the tail-ends of the distribution with the initially richest economies displaying a comparatively higher probability of moving into a lower relative income segment of the distribution than remaining where they began. The reverse holds for the poorest economies. In terms of shape there is a distinction between the pre-1914 period and the later two periods in that the former is characterized by a single peak and the latter by twin peaks. However, the single peak in the pre-1914 period is centred close to the 1.0 value of both the period t and $t + 5$ axes. In other words, middle income economies display a high probability to maintain their initial middle income position within the cross-section distribution over 5-year transitions. In summary, the relative income dynamics observed for the pre-1914 period are dominated by persistence with weak tendencies of movement only from the very tailends towards the middle of the distribution.²¹ As such, this evidence on mobility is not entirely incompatible with a Williamson-type story of convergence in the late 19th century, though it does not add up to strong support. For that to be forthcoming, the data should show a far more pronounced pattern of mobility.

The trans-war period displays a higher degree of relative income mobility than the years before the First World War. However, the lower of the two peaks of this stochastic kernel is centred on the 45° diagonal and thus suggests that economies with below average incomes (centred around the 0.8 levels indicated on the horizontal axes) showed a high propensity over a 5-year horizon to stay in that part of the cross-section distribution where they began. The far less pronounced upper income peak (near the 1.0 levels) exhibits a clockwise movement, indicating that, broadly speaking, average income economies tended to become richer over a 5-year horizon. Moreover, mobility in this period is mainly concentrated in the tails of the distribution and exhibits counter-clockwise movement. Thus, overall the trans-war period is one of conflicting signals, perhaps not surprisingly given the traumas of Maddison's 'bleak age.'

The pattern of distributional change in the half century since the Second World War is characterized by polarization and emerging twin peaks (or clusters of observed transitions) that are clearer and further apart from each other than in the trans-war period, suggesting the formation of distinctive clubs of the comparatively rich and poor. Moreover, there is significant mobility as both these peaks exhibit a counter-clockwise movement around the 45° diagonal which indicates that within the two 'clubs' (of those economies that were initially just below and those that were initially well above the average of incomes) there is a tendency towards income equalization over a 5-year transition period. It also implies that the group of the relatively rich is becoming relatively poorer, whereas the lower income club is getting relatively richer.

²¹ Cf. the discussion of the box-plots (Fig. 1) in Section 3 above.

However, the expanded set of 24 economies reveals a richer, more nuanced picture of distributional change for it permits distinguishing between the Golden Age and post-Golden Age years (Fig. 5). It shows that although twin peaks are apparent in the Golden Age this characteristic was even stronger in the post-Golden Age.²² The pronounced dip in the middle of the stochastic kernel in the latter period also suggests that this is partly the result of separation with the middle income economies gravitating strongly towards one pole or the other. In other words, the ‘middle ground’ is thinning out over a 5-year horizon. Indeed, the contour plot for the post-Golden Age period is indicative of what have been called ‘basins of attraction.’²³ However, emptying of the middle income sections does not preclude transitions from low to high income parts of the distribution, and vice versa: the stochastic kernel is positive almost everywhere and covers the whole income range. The other important difference between the two sub-periods is the change in the location of the stochastic kernel. In the Golden Age period the kernel is, broadly, shifting counter-clockwise relative to the 45° line (especially around the upper income pole), in the post-Golden Age period it is shifting clockwise. Thus, the clustering in the post-Golden Age shows a relatively strong tendency towards polarization, where the relatively rich get richer and the poor become poorer.

6. Long-run equilibrium distributions

So far, the discussion has focused on mobility and persistence in the cross-country income distribution during the three main periods under review. However, if we believe that each of the periods identified here do indeed represent different regimes, then it is also important to establish the long-run equilibrium of each regime so as to be able to address the issue of convergence. These steady states are not apparent from the distributional characteristics reflected in the empirically observed dynamics during a given period.

To find the steady state solution of a regime we turn to the discrete analysis of the transition probabilities matrix, i.e., the discrete analogue of the stochastic kernel.²⁴ First, for each of the data sets and periods under consideration we estimate 1-year transition probability matrices, using five income states. These matrices give the probabilities of economies to move from one income state to another, on average in any 1 year across the period. Second, drawing on the well-known feature of

²² This is consistent with other studies of the post-war which have used a similar distribution dynamics approach. For example, Bianchi (1997) examined the distribution of GDP in 119 countries in 1970, 1980, and 1989 and found ‘increasing evidence for bimodality’ (p. 393), concluding that ‘the empirical evidence... appears to support the view of clustering and stratification of growth patterns over time, in sharp contrast to the convergence hypothesis’ (p. 408). Similarly, Paap and van Dijk (1998), in their examination of real GDP per capita of 120 countries over the period 1960–1989, also discovered evidence of bimodality between a relatively large group of poor countries and a small group of rich countries and further argued that the mean real GDP per capita of the two groups were diverging.

²³ Durlauf and Johnson (1995).

²⁴ On the derivation of the income states see Footnote 13.

Table 1
Ergodic distributions

	Income States				
	1	2	3	4	5
17 advanced economies, 1870–1992					
1870–1913	.171	.152	.167	.245	.266
1914–1950	.218	.205	.202	.201	.174
1951–1992	.156	.203	.259	.228	.153
24 advanced economies, 1951–1992					
1951–1973	.038	.182	.313	.319	.148
1974–1992	.103	.153	.212	.241	.291

any transition probability matrix that, through continuous iteration, it will eventually yield a unique long-run steady state condition, we derive the ergodic distribution. This is a central concept in the analysis since it permits gauging the ‘convergence’ properties of historical regimes. By addressing the question of what the outcome would be if the dynamic system represented by the transition probability matrix were allowed to evolve unrestrictedly (i.e., beyond the length of the actual historical period whose empirical data it is incorporating), it allows us to estimate the extent to which particular regimes were conducive to long-run convergence processes.²⁵ The ergodic distribution suggests what the shape of the long-run equilibrium distribution would look like.²⁶

Table 1 shows the ergodic distributions for the 17 and 24 economies associated with each period. The numbers in the table report the equilibrium proportion of economies falling in either of the five relative income states. According to this evidence, the long-run equilibrium of the pre-1914 regime was not one characterized by convergence. The steady state distribution shows two distinct plateaux, the first of which over income states 1, 2, and 3, the second over income states 4 and 5. This would suggest that in the long-run stratification was at least as strong a force as persistence.²⁷

For the trans-war regime the ergodic distribution is effectively flat across the first four income states (each of them counting for about 20–21% of the observations in

²⁵ As stated previously this was done using TSRF. A more detailed explanation of the calculation and use of transition probability matrices and ergodic distributions is provided in the appendices of Epstein et al. (1999).

²⁶ Note, though, that the imposition of discrete states involves some loss of responsiveness to changes in the underlying data. For example, using the stochastic kernel allows us to trace transitions within the distribution over shorter distances. Movements that fall within the boundaries of discrete income states cannot be captured by the transition matrices, but such movements could be traced in the stochastic kernel estimates. Thus tendencies such as clustering, separation, and polarization are more readily picked up in the stochastic kernel. In other words, transition matrices offer a less detailed and precise depiction of the period dynamics than stochastic kernels and, therefore, the ergodic distributions are affected by the coarser underlying state grid.

²⁷ It should be noted that persistence was the prevalent feature both in the corresponding 1-year transition probability matrix for this period and in the stochastic kernel estimated for 5-year transitions.

the distribution) and then tails off slightly in income state 5. However, the shape of this distribution, which points neither to convergence nor divergence, cannot be interpreted as merely an outcome of persistence. Rather, it seems to have been associated with a considerable degree of intra-distributional mobility.²⁸ This is consistent with the history and historiography of this period. It is a period that experienced three of the most significant shocks in the 20th century (the two World Wars and the Great Depression). Furthermore, the impact of these shocks, and the economic reaction to each shock, differed across individual economies. Thus, it is not surprising that no clear long-run convergence or divergence can be detected in this period.²⁹

Finally, the post-war regime offers almost a textbook example of distributional convergence in that there is a clear peak in the middle of the distribution and the lowest points are at the extremes of the distribution.

Turning to the 24 economies, the long-run equilibrium of the Golden Age regime shows a clear peak in the middle of the distribution (income states 3 and 4 account for almost two-thirds of the distribution). In contrast, the long-run equilibrium of the post-Golden Age regime shows increasing density positively associated with a movement up the income state scale. This is an example of a uni-modal steady state distribution that does not reflect convergence. Indeed, the stochastic kernel indicated that the empirically observed clustering was in fact associated with separation and polarization where the relatively rich get richer and the poor became poorer.

7. Implications

Much of the recent economic history debate on long-run convergence has been shaped by the work of Williamson (1996) and O'Rourke and Williamson (1999). The focus there is primarily on the relationship between economic growth, globalization, and factor price convergence. Williamson argues that the open economy forces of mass migration and trade ('globalization') made for rapid real wage convergence in the pre-1914 period, while the inter-war years were characterized by slow growth, 'de-globalization,' and divergence. Growth after 1950 was fast again and associated with globalization and convergence. 'Thus history offers an unambiguous positive correlation between globalization and convergence.'³⁰ The evidence presented here would suggest that history's offerings are more ambiguous than allowed by this bold claim.

²⁸ The sum of the off-diagonal values in the transition matrix for the trans-war period is the highest amongst the three sub-periods, indicating a high degree of the economies' mobility between income states. Whilst there is a lot of upward and downward movements within the distribution and there are rank order changes, these do not translate into a long-run equilibrium distribution displaying pronounced uni- or multi-modality. Note that the stochastic kernel, too, revealed a high degree of relative income mobility in this period.

²⁹ In an unpublished paper (available from the authors on request), we attempted to abstract from the war shocks and considered only the period 1922–1938. The ergodic distribution for this regime exhibited clear signs of uni-modality (that is, convergence). This suggests that the World Wars, rather than the Great Depression, are the important shocks in terms of the distribution dynamics of the regime.

³⁰ Williamson (1996, p. 277).

Drawing on a widely used standard data set, yet adopting a new analytical perspective by exploiting recent advances in non-parametric econometrics, we argue that the timing and incidence of convergence processes over the last 120 or so years is not as straightforward as has been suggested.³¹ However, the regimes we examined were, in terms of their distribution dynamics, distinctive from one another. This suggests that historians such as Maddison and Williamson are right to treat these periods as different epochs or phases of development. Our findings add up to a significant qualification of what can perhaps be described as a consensus view on the broad phasing of convergence since the late 19th century. Broadberry (1996) surveys the long-run evidence for convergence among industrialized economies, drawing on Maddison's national income data, Williamson's real wage data and his own work on labour productivity in manufacturing. He argues, first, that this collective body of evidence is broadly consistent with local rather than global convergence and, second, that the convergence process has not been smooth and continuous. However, the process of convergence, under way during the nineteenth century, was interrupted during the trans-war period and resumed again thereafter with a particularly rapid pace during 1950–1973. This temporal pattern of convergence is not confirmed by an analysis whose focus is on the dynamics of changes in the international income distribution. In particular, the Golden Age stands out as the only one of these regimes that was commensurate with long-run convergence.

Three features arise from our analysis. First, the data we have analyzed do not suggest that the regime prevalent before 1914 was consistent with strong convergence. Rather, the period dynamics captured in the stochastic kernel point to a relatively low degree of distributional mobility, while the long-run equilibrium of this regime was characterized by forces of persistence and stratification. Second, though indicative of a high degree of intra-distributional mobility, the period dynamics for the trans-war period provide conflicting signals. This lack of a clear tendency is reinforced by the relatively flat shape of the ergodic distribution. The steady state of this regime thus points to neither convergence nor divergence in real per capita incomes. Finally, the analysis of the distribution dynamics of the 24 economies in the post-war period potentially poses some fundamental questions. It suggests that convergence was a feature of the Golden Age but that it was a temporary phenomenon, or rather one specific to that regime. Once problems were encountered the regime gave way to one in which separation, divergence, and polarization came to the fore. Both regimes saw the emerging formation of 'clubs,' but in the Golden Age the relatively rich got poorer and the relatively poor got richer, while the post-Golden Age witnessed polarization where the relatively rich got richer and the poor were losing out in relative terms. This begs the question as to how far the convergence of the Golden Age was an artefact of post-war reconstruction and catch-up.

³¹ To be fair, O'Rourke and Williamson (1999, pp. 5–28) acknowledge explicitly that the convergence properties of GDP per worker hour and GDP per capita are different from those of real wages and that the open economy mechanisms they argue were driving late 19th century convergence operated only indirectly on GDP per capita. They do maintain, though, that convergence was not limited to real wages and labour markets, but also extended to GDP per capita, albeit at markedly slower rates.

Technical appendix

The disadvantage of the discrete, n -state Markov model is that the number of states is chosen arbitrarily. The kernel-density estimator is a generalization of the discrete model, in which the n tends to infinity, rendering the state space continuous. The histogram is a discrete analogue of the kernel estimate of single densities. Data are divided into disjoint class intervals with the bar centred at the midpoint of the interval. The height of each bar reflects the number of observations in the interval. As an extension of the histogram, the kernel density estimate permits the class intervals to overlap. In effect the interval, which in this case is called a ‘window,’ slides along the range of the observations, and centre point estimates are made, the width of the window being known as ‘bandwidth.’ Unlike the histogram, the kernel weights each observation by its distance from the centre point. The weighted observations are summed to give the height of the ordinate at each point at which the kernel is being estimated. Estimates are smoother, and therefore more easily comparable with known parametric distributions.

The transition probability matrix can be seen as a histogram of joint densities. The stochastic kernel is a generalization of this.³² It can be represented either as a two-dimensional contour plot or as an orthogonal projection onto a surface in three dimensions. Each point of the surface is interpreted as a probability. The stochastic kernel estimate is simply the continuous analogue of the transition probability matrix.

Formally, following Quah (1997) Section 4, let μ and ν be probabilities, and let A be a window; then the stochastic kernel is the mapping $M_{(\mu,\nu)}$ which satisfies

$$\mu(A) = \int M_{(\mu,\nu)}(y, A) d\nu(y)$$

subject to certain restrictions. That is, for a given A , the count $M(y, A)$ is weighted by $d\nu(y)$ and summed over all possible values of y , giving the fraction of economies ending up in state A regardless of their initial state. The restrictions, which are discussed in Quah (1997), guarantee that $\mu(A)$ is a Lebesgue integral, that is, a weighted sum of data points in the window A . The stochastic kernel gives a complete description of transition probabilities from state y to any other state.³³

References

- Barro, R.J., Sala-I-Martin, X., 1991. Convergence across states and regions. *Brookings Papers*, 107–182.
- Baumol, W., 1986. Productivity growth, convergence and welfare: what the long-run data show. *American Economic Review* 76, 1072–1085.
- Bellman, R., 1960. *Introduction to Matrix Analysis*. McGraw-Hill, New York.
- Bernard, A.B., Durlauf, S.N., 1995. Convergence in international output. *Journal of Applied Econometrics* 10, 97–108.

³² See Epstein et al. (2000) for a theoretical and empirical discussion of the transition probability matrix. Applications of the stochastic kernel are discussed in Quah (1997) and Durlauf and Quah (1998).

³³ Further related readings include Doob (1953), Bellman (1960), Chung (1960), and Stokely et al. (1989).

- Bianchi, M., 1997. Testing for convergence: Evidence from non-parametric tests. *Journal of Applied Econometrics* 12, 393–411.
- Broadberry, S.N., 1996. Convergence: what the historical record shows. In: van Ark, B., Crafts, N.F.R. (Eds.), *Quantitative Aspects of Post-War European Growth*. Cambridge University Press, Cambridge.
- Chung, K.L., 1960. *Markov Chains with Stationary Transition Probabilities*. Springer, Berlin.
- Doob, J.L., 1953. *Stochastic Processes*. New York.
- Durlauf, S., Johnson, P.A., 1995. Multiple regimes and cross-country growth behaviour. *Journal of Applied Econometrics* 10, 365–384.
- Durlauf, S., Quah, D., 1998. The new empirics of economic growth. Centre for Economic Performance, London School of Economics and Political Science, Discussion Paper No. 384.
- Epstein, P., Howlett, P., Schulze, M.-S., 1999. Income distribution and convergence: the European experience, 1870–1992. Department of Economic History, London School of Economics and Political Science, LSE Working Papers in Economic History 52/99.
- Epstein, P., Howlett, P., Schulze, M.-S., 2000. Distribution dynamics: stratification, polarization and convergence among OECD economies, 1870–1992. Department of Economic History, London School of Economics and Political Science, LSE Working Papers in Economic History 58/00.
- Esteban, J.-M., Ray, D., 1994. On the Measurement of Polarization. *Econometrica* 62, 819–851.
- Hobijn, B., Franses, P.H., 2000. Asymptotically Perfect and Relative Convergence of Productivity. *Journal of Applied Econometrics* 15, 59–82.
- Japan Statistical Association, 1998. *Historical Statistics of Japan*, vol. 2. Japan Statistical Association, Tokyo.
- Maddison, A., 1995. *Monitoring the World Economy, 1820–1990*. OECD, Paris.
- Mankiw, N.G., Romer, D., Weil, D., 1992. A contribution to the empirics of economic growth. *Quarterly Journal of Economics* 107, 407–438.
- Mills, T.C., Crafts, N.F.R., 2000. After the golden age: a long run perspective on growth rates that speeded up, slowed down and still differ. *Manchester School* 68, 68–91.
- Ohkawa, K., Shinohara, M., Meissner, L., 1979. *Patterns of Japanese Economic Development*. Yale University Press, New Haven.
- O'Rourke, K.H., Williamson, J.G., 1999. *Globalization and History*. MIT Press, Cambridge.
- Paap, R., van Dijk, H.K., 1998. Distribution and mobility of wealth of nations. *European Economic Review* 42, 1269–1293.
- Prados de la Escosura, L., 1995. Spain's gross domestic product, 1850–1993: quantitative conjectures. Appendix. Universidad Carlos III, Working Paper 96-06, Economics Series 02.
- Prados de la Escosura, L., 2000. International comparisons of real product, 1820–1990: an alternative data set. *Explorations in Economic History* 37, 1–41.
- Quah, D.T., 1993. Empirical cross-section dynamics in economic growth. *European Economic Review* 37, 426–434.
- Quah, D.T., 1996. Twin peaks: Growth and convergence in models of distribution dynamics. *Economic Journal* 102, 1045–1055.
- Quah, D.T., 1997. Empirics for growth and distribution: stratification, polarization, and convergence clubs. *Journal of Economic Growth* 2, 27–59.
- Romer, P., 1986. Increasing returns and long run growth. *Journal of Political Economy* 94, 1002–1037.
- Schulze, M.-S., 2000. Patterns of growth and stagnation in the 19th century Habsburg economy. *European Review of Economic History* 4, 311–319.
- Solow, R.M., 1956. A contribution to the theory of economic growth. *Quarterly Journal of Economics* 70, 65–94.
- Stokey, N.L., Lucas, R.E., Prescott, E.C., 1989. *Recursive Methods in Economic Dynamics*. Harvard University Press, Cambridge, MA.
- Temple, J., 1999. The new growth evidence. *Journal of Economic Literature* 37, 112–156.
- Toniolo, G., 1998. Europe's Golden Age, 1950–1973: speculations from a long-run perspective. *Economic History Review* 51, 252–267.
- Tortella, G., 1994. Patterns of economic retardation and recovery in south-western Europe in the nineteenth and twentieth centuries. *Economic History Review* 47, 1–21.
- Williamson, J.G., 1996. Globalization, convergence, and history. *Journal of Economic History* 56, 277–306.