Working Paper No. 52/99

Income distribution and convergence: the European experience, 1870-1992

Philip Epstein, Peter Howlett, Max-Stephan Schulze

Department of Economic History
London School of Economics 1999

Introduction

Economic convergence has emerged as one of the key debates in the theoretical and historical literature over the last decade.[1] Galor identified three forms of long run per capita income convergence: absolute convergence, whereby convergence occurs independently of the initial conditions facing each economy; conditional convergence, whereby convergence occurs among economies which have identical structural characteristics, independently of their initial conditions; and club convergence, whereby convergence occurs only if the structural characteristics are identical and initial conditions are also similar.[2] Of these, the absolute convergence hypothesis has been discredited whereas there is empirical support for both the conditional convergence and club convergence hypotheses. The club convergence hypothesis, in particular, has much to offer to economic historians. It stresses the importance of both the initial conditions facing each economy and the structural and institutional features of the economy (e.g. preferences, technologies, rates of population growth, government policies, etc.).

However, much of this debate has focused on growth and has ignored issues of distribution. This is a serious weakness in the convergence literature because only by understanding the interaction between different economies can the causal mechanisms underlying the process of convergence, which by definition is a concept of relative behaviour, be understood. Most recently, though, an approach focusing on distribution dynamics, rather than linear (or log-linear) regression, has been pioneered by Danny Quah.[3] We use this method to examine the pattern of income growth and convergence in Europe since the mid-nineteenth century.

---

[1] Williamson, J. G., 'Globalization, Convergence, and History', *Journal of Economic History* 56 (1996), pp.277-306; Broadberry, S.N., 'Convergence: What the Historical Record Shows', in B. van Ark and N. Crafts (eds.), *Quantitative Aspects of Post-War European Growth* (Cambridge, 1996), pp.327-346.

[2] Galor, O., 'Convergence? Inferences from Theoretical Models', *Economic Journal* 106 (1996), pp.1056-69.

[3] See, for example: Quah, D. T., 'Empirical cross-section dynamics in economic growth', *European Economic Review* 37 (1993), pp.426-34; Quah, D. T., 'Twin peaks: growth and convergence in models of distribution dynamics', *Economic Journal* 106 (1996), pp.1045-55; Quah, D. T., 'Empirics for growth and

The strength of this approach is that it can help to answer a range of important questions that have been at the forefront of the work of many economic historians. Quah summarises these questions thus:

> 'Are the currently leading economies always the first to push back technology frontiers and does new technology then always filter passively to poorer economies? Are there costs of adoption that lead to leap-frogging, where it is the temporarily follower economies that jump to being leader, because they find it easier to exploit new discoveries? Or do persistent advantages accrue to the leader, richer countries simply by virtue of their already being leader and richer? Do poorer economies need to overcome poverty-trap barriers before they can hope to catch up with richer ones.'[4]

A brief summary of the theoretical and historical literature on convergence will be followed by a description of the data to be utilised (we will focus on income growth in Europe since 1870) and an empirical consideration of F-convergence in our period. The rest of the paper then outlines the Quah methodology and applies it to our data. We will focus on three separate issues:

1. the empirical distribution of income across the economies being studied in different sub-periods;
2. the transition probability matrix, which allows us to discuss issues of persistence and mobility within the distribution;
3. the ergodic distribution, which allows us to discuss the long run equilibrium of different regimes and hence polarisation and stratification

We should stress that this paper is a preliminary historical application of the Quah technique of distribution dynamics and, as such, its findings will be generally descriptive rather than analytical in nature.

---

distribution: stratification, polarization, and convergence clubs', *Journal of Economic Growth* 2 (1997), pp.27-59.

[4] Quah, 'Twin peaks', p.1047.

Convergence: the traditional approach

An important distinction in the literature is that between *ß*-convergence and *F*-convergence.[5] *ß*-convergence implies that over time, for a given group of economies, growth rates will converge and that initially poor countries grow faster than rich countries, in other words, there is an inverse relationship between the growth rate and the initial level of per capita income (or productivity). Referring to Galton's regression fallacy, the appropriateness of the standard technique of regressing growth rates on initial income or productivity levels as a means of testing for *ß*-convergence has been questioned by Friedman.[6] Simple or unconditional *ß*-convergence is generally seen as too crude a measure and, therefore, the convergence process is often modified to be conditional on additional structural characteristics.[7] Conditioning on other structural characteristics such as, for example, education levels and government expenditure allows gaps between different economies, or clubs of economies, to remain persistent (despite convergence within a club).[8] *F*-convergence refers to a reduction in the

---

[5] This distinction was first made by Barro and Sala-Martin in their influential study of the income behaviour of the states of the USA (1880-1988) and of 73 regions of Western Europe (1950-1985). R.J. Barro and X. Sala-I-Martin, 'Convergence across states and regions', *Brookings Papers On Economic Activity* (1991), pp.107-82.

[6] Friedman, M., 'Do old fallacies ever die? *Journal of Economic Literature* XXX (1992), pp.2129-2132.

[7] For example, in their classic paper, Mankiw, Romer and Weil found 'no tendency for poor countries to grow faster on average than rich countries' although there was 'a significant tendency towards convergence in the OECD sample'. N.G. Mankiw, D. Romer and D.N. Weil, 'A contribution to the empirics of growth', *Quarterly Journal of Economics* 107 (1992), p.425.

[8] As Broadberry has noted conditional convergence analysis has introduced an ironic tone: 'there are forces propelling an economy towards the steady-state level of [income] and the steady-state growth rate, but the steady-state may differ between economies. In other words, there may be catch-up forces at work, but they may be offset by other effects keeping economies apart. Conditional convergence regressions, then, may actually be more useful in establishing factors making for divergence than in establishing convergence'. Broadberry, 'Convergence', p.328.

dispersion of income levels over time, i.e. the variance of income across the economies will decline. *ß*-convergence does not necessarily imply *F*-convergence primarily because shocks which have a localised or differential impact may increase income dispersion and thus offset the effects of *ß*-convergence. Moreover, income differentials between two economies can increase in absolute terms even if they decrease in percentage terms.[9] We will not be discussing *ß*-convergence but we will use *F*-convergence (as measured by the coefficient of variation) as our starting point to illustrate some characteristics of the income data used here.

Economic historians have found the notion of club convergence to have some appeal and we will illustrate this by considering the work of Tortella, Toniolo, Broadberry and Williamson. Tortella has argued that there were at least two growth clubs in Europe in the nineteenth and twentieth centuries. He makes a distinction between the Latin European economies (Spain, Portugal and Italy) and the north-western economies (Britain, France and Germany).[10] The Latin economies, according to Tortella, share certain common features, related to culture (language, religion, the legal system) and geographical endowment (which influences the pattern of agricultural production and diet). These, in turn, manifested themselves in economic characteristics that help to identify a pattern of modernisation that was distinct from the north-western economies. The economic characteristics were: relatively low agricultural productivity; a specific system of land tenure; low literacy and school enrolment; trade barriers which acted as obstacles to industrialisation; chronic budget deficits in the late nineteenth century; and rule by dictators for long periods of time in the twentieth century. Tortella argues that convergence in the latter part of the twentieth century was mainly due to the shock caused by the influx of cheap grain. This brought about a movement of resources from low productivity agriculture (grains

---

[9] Broadberry, 'Convergence', p.329. He also notes mathematical reasons for why convergence does not necessarily imply *F*-convergence.

[10] Tortella, G., 'Patterns of economic retardation and recovery in south-western Europe in the nineteenth *Economic History Review* 47 (1994), pp.1-21.

and legumes) to high productivity agriculture (wine, fruit and vegetables), and from agriculture into industry.

Toniolo identifies a European convergence club, including the eastern European economies, in the Golden Age (1950-73).[11] The emergence of the European club in this period, he argues, was due to four factors that enhanced social capability.[12] Those factors were: a relatively favourable human/physical capital ratio; a match between transferable technology and the emerging pattern of demand (the focus being on Fordism); what Toniolo refers to as the 'solution of the stability problem' which is his synthesis of the Maier and Milward positions; and, borrowing from the work of Eichengreen on post-war institutions and labour market solutions, the creation of new institutions.[13]

In a recent survey of three major data sets on national income, manufacturing productivity , and real wages, respectively, Broadberry argues for the presence of local rather than global convergence during the late 19th and 20th centuries and identifies convergence clubs, even within the group of advanced industrialized economies.[14] Whilst emphasizing the difference in comparative productivity levels and

---

[11] Toniolo, G., 'Europe's golden age, 1950-1973: speculations from a long-run perspective', *Economic History Review* 51 (1998), pp.252-67. His western European economies are Austria, Belgium, Denmark, Finland, France, Germany, Italy, the Netherlands, Norway, Sweden, Switzerland and the UK; his eastern European economies are Bulgaria, Czechoslovakia, Hungary, Poland, Romania, the USSR, and Yugoslavia. The western European economies are that we use in our group of 12 economies below.

[12] The influential notion of social capability is most closely associated with the work of Moses Abramovitz. See, for example, his 'Catching up, forging ahead, and falling behind', *Journal of Economic History* 46 (1986), pp.385-406

[13] See: Maier, C., *In search of stability: explorations in historical political economy* (Cambridge, 1987); Milward, A., *The European rescue of the nation- state* (London, 1992); Eichengreen, B., 'Institutions and economic growth: Europe after World War II', in N. Crafts and G. Toniolo (eds.), *Economic growth in Europe since 1945* (Cambridge, 1996), pp.38-72.

[14] Broadberry, 'Convergence', pp.329-43.

trends between manufacturing and the whole economy, he points to the long-run persistence in the labour productivity gap between European and North American manufacturing. This is explained as largely an outcome of different resource constraints and demand conditions that lead to the use of different production techniques with different levels of labour productivity. Within Europe, Broadberry distinguishes between a 'Northern' and 'Southern' convergence path. In the Northern economies, where manufacturing productivity had fluctuated around British levels in the 20th century, catching-up with Britain at the whole economy level was an outcome of the expansion in the size of industry and, even more so, the result of the shift of resources out of low productivity agriculture. In France and Italy with markedly lower pre-war productivity levels in manufacturing, convergence to average European productivity levels after 1945 has been achieved both in manufacturing and at the whole economy level.

Probably the most influential economic history research programme concerned with long run convergence is that of Williamson. In discussing the OECD club he identifies three epochs: the late nineteenth century was characterised by fast growth, globalisation, and convergence; 1914-50 witnessed slow growth, de-globalisation, and divergence; and the post-1950 era has experienced fast growth, globalisation and convergence.[15] Much of his work in this area has been primarily concerned with the issue of globalisation and its relationship to convergence (by convergence, he means *F*-convergence). In the pre-1914 period he argues that mass migration accounted for about 70% of the convergence in real wages and most of the rest was accounted for by trade forces: it 'follows that migration and trade restrictions associated with war and

[15] Williamson, J. G., 'Globalization', pp.277-306. In this paper he actually examines 17 countries using his real wage data set - the OECD countries of European origin (Australia, Belgium, Canada, Denmark, France, Germany, Great Britain, Ireland, Italy, the Netherlands, Norway, Portugal, Spain, Sweden and the USA) plus Argentina and Brazil. He compares his real wage results with results based on GDP per capita taken from Maddison but based on a different sample of countries. His Maddison sample does not include Argentina, Brazil, Ireland, Portugal, Spain or the USA; but it does include Austria, Finland and Switzerland.

policy must go a long way in explaining why convergence stopped after 1914.'[16] Thus, he argues that since the late nineteenth century there has been a strong correlation between globalisation and convergence and before 1914 this was a causal relationship (that is, globalisation led to convergence). However, he also points out the limitations of this sort of convergence analysis: 'much of the convergence since 1870 disappears when the net is widened to include Eastern Europe, and if it were widened still further to include the Third World, convergence would totally evaporate'.[17]

For the pre-1914 period, like Tortella, Williamson (using his real wage data) distinguishes the experience of northern Europe (Denmark, Sweden, Norway and Ireland) from that of Portugal and Spain. Indeed, he argues that 'what dominated European experience was not Britain's failure (which hastened convergence) but the failure of the Latin economies (which retarded convergence)'.[18]

Data description

In this preliminary study, we focus on the European economies. Our analysis is restricted by the availability of homogenous long run historical data. Thus, we first examine 12 western European economies for the period 1870 to 1992 and an expanded group of 24 European economies for the period 1955 to 1992. In each case, the period is divided into sub-periods. These sub-periods have been chosen to conform to the stylised story told by Williamson and to avoid the effects of war shocks. Moreover, for the post-World War II era they reflect the current historiography, which tends to discuss the period in terms of the 'Golden Age' and post-'Golden Age'. Table 1 summarises this information.

The data have been taken from the well-known and widely used Maddison data set. These data are continuous, cover the whole of the twentieth and much of the

---

[16] Williamson, 'Globalization', p.295.

[17] Williamson, 'Globalization', p.279.

[18] Williamson, 'Globalization', pp.285-6.

nineteenth century, and have been much used in the recent historical literature. Drawing on this data set will also help to assess the extent to which a different methodological approach, rather than different data, sheds new light on existing interpretations.[19] Maddison thus provides us with annual estimates of the level of real GDP per capita in 1990 Geary-Khamis dollars for each economy.[20] For each year, we then standardised these data to the average level of real GDP per head in that year for the group of economies being examined (with the average taking a value of 1.00).[21]

F-convergence

The simplest measure of *F*-convergence is the coefficient of variation: for any given group of economies, *F*-convergence implies that over time the variation in their income**s** relative to their mean income will decline. Thus convergence should be reflected in a decline of the coefficient of variation over time. Figures 1 and 2 show this measure for our two groups of economies in order to provide some background for the discussion of distribution dynamics.

Looking at the 12 countries sample, four observations on the dispersion in *levels* of per capita incomes in Europe can be made. First, during the late nineteenth century up to 1913 dispersion gradually declined after an initial increase from the mid-1870s to the late 1880s. Second, the First World War and its aftermath interrupted this process: there was a pronounced yet temporary rise in income dispersion during the war,

---

[19] It can also be compared to the results found in the convergence debate in the economics literature which has typically drawn on the Summers and Heston data set (Summers, R., and Heston, A., ‘The Penn World Table (Mark 5): an expanded set of international comparisons, 1950-1988’, *Quarterly Journal of Economics* 106 (1991), pp.327-68).

[20] Maddison, A., *Monitoring the world economy, 1820-1990* (Paris, 1995), appendices D-1a, D-1b and D-1c, pp.194-201.

[21] For example, the richest of the 12 economies in 1870 was the UK whose level of real GDP per capita was 65 % greater than the average for all 12 economies in that year and thus its standardised value was 1.65.

followed by a prolonged period of virtually no change with an almost constant coefficient of variation of about .25, equivalent to that prevailing just before the outbreak of the First World War. Only during the later 1930s was there a further, albeit limited reduction in the degree of income dispersion. Third, during the Second World War the variance in per capita income across the European economies rose dramatically, to decline again in the immediate post-war years. Finally, from the 1950s to the 1990s there was a virtually uninterrupted diminution of variance in incomes. This evidence would suggest that during the pre-World War I period a process of gradual convergence got under way that was, however, interrupted by two world wars and the great depression. The period from 1913 to 1950 as a whole was one of *divergence* rather than convergence. After the end of the Second World War, per capita income differentials among European economies became progressively smaller, with a particularly fast reduction in income dispersion during the Golden Age' (1950-1973). In this sense one can say that post war growth was accompanied by convergence.

However, it should be emphasised that any statement on the presence or absence of convergence depends critically on the range of countries included in the sample. This point is well illustrated in Figure 2 which depicts the coefficient of variation in per capita incomes for the extended sample of 24 economies (1955-1992). It is important to note that, first of all, throughout the post-Second World War period the degree of income dispersion among these economies was much higher than it was for the 12 economies in the late nineteenth century, most of the inter-war period, and especially the post-1945 period. Moreover, it shows a markedly different pattern over time when, again, being compared to the 12 economies. Whilst there was a similarly pronounced decline during the 'Golden Age', there was a sharp rise in income dispersion from the late 1970s indicating divergence, this being the outcome of the accelerating relative (and in some cases absolute) decline in economic performance of the former Eastern bloc countries.

An Intuitive Introduction to Distribution Dynamics

Quah has argued that models of distribution dynamics can yield many insights into the process of growth and convergence.[22] The model of distribution dynamics he uses

[22] The following discussion is based on Quah, 'Empirics for growth and distribution', pp.27-29.

emphasises two main features: *shape* and *mobility*. The starting point is the empirical observations from the cross-sectional distribution of income at a point in time. Certain characteristics about the distribution can be noted, most obviously its shape. The change in the cross-sectional distribution can then be compared over time, to see whether the underlying characteristics change.

Figure 3 considers some of the possible outcomes of a change in the distribution of income over time. It assumes that in period t the distribution of incomes approximates a rather flat normal distribution and then it considers three possible outcomes in period $t + s$. In the first case, the shape of the distribution has not changed. If the economies that were rich (relative to the distribution) remain rich, and economies that were poor remain poor, then this would be an example of *persistence*. In the second case, we still have a normal distribution in $t + s$ but the peak is now far more pronounced: this is an example of *convergence*. It suggests a certain degree of mobility since some economies that had originally been in the tails of the distribution have moved into the middle of the distribution. In the third case, the distribution has become bi-modal in period $t + s$, Quah refers to this as a twin peak distribution. It is an example of *polarisation* in which the poor economies have clustered together around one peak and the rich economies have clustered together around another peak. It is also another example of persistence in the case of the originally rich and poor economies. The originally middle income economies, however, have shown a high degree of mobility. They have effectively disappeared, migrating to either the rich or the poor peak. Emerging twin peak behaviour can also be thought of as a process of *stratification* and is compatible with ideas about convergence clubs.

We have thus identified three stylised shapes for the distribution:

1) a flat distribution with no obvious peaks, which is consistent with a story of non-convergence;

2) a uni-modal distribution, which is consistent with a story of convergence; and

3) a multi-modal distribution, which is consistent with a story of club convergence (in our example we have shown the particular case of the twin peak distribution).

Obviously, there are many possible shapes of cross-section income distributions. Most observed empirical distributions will not take the exact form of one of these three stylised distributions but will be an intermediate solution. This means that the interpretation of these distributions must be undertaken with great care. Although we can derive meaningful inferences by observing the shape of the distribution over different time periods or across regimes, this shape may conceal other important information. In particular, comparing changes in the shape does not tell us anything about the experience of individual economies. It is possible for any individual economy to move from one part of the distribution to another. Thus, a rich country might become poor and a poor economy might become rich without necessarily affecting the shape of the cross-section distribution. We therefore need to consider the issue of mobility. In Figure 3 (a), if the rank order of the economies did not change between periods *t* and *t* + *s*, then we could characterise this as an example of persistence. Figure 4 (a) replicates this example of no changes in the shape of the distribution, but now shows mobility within the distribution: relative to the distribution, one economy is rich in period *t* but is poor in period *t* + *s*, whereas the other economy which is poor in period *t* is rich in period *t* + *s*. In the discussion above of the emerging twin peaks, we also noted the case of the original middle income economies moving into either the rich club or the poor club over time. This is an example of *separation* and illustrated in Figure 4 (b).

Therefore, another important characteristic of these distributions is the nature and degree of mobility. Examining the intra-distribution dynamics not only provides information on the rank of countries within the distribution but can also quantify the changes.[23] The degree of mobility is not dependent on the shape of the distribution. Insignificant changes in the shape of the distribution could be associated with very high degrees of mobility; likewise, very low degrees of mobility may induce significant changes in the shape of the distribution.

Empirical Distributions and Mobility

Before setting out the technical aspects of the Quah approach we want to consider briefly the actual empirical distributions associated with our two groups of economies, to illustrate the importance of mobility. The technique allows us to identify five income states (state 1 representing the poorest level of income and state five the highest level

---

[23] Quah, 'Twin peaks', p.1050.

of income) and the derivation of these states will be explained below. Tables 2 and 3 show the actual position of our economies at the beginning and end of each sub-period. Countries that experienced a move of more than two income states are shown in bold.

Consider the 12 economies in the post-war period. We saw that this period was marked by significant *F*-convergence but Table 2 also makes clear that convergence was only part of the story. Concentration on convergence ignores some important historical issues. For example, the literature on the UK in the post-war period is dominated by the story of relative economic decline.[24] A simple focus on convergence cannot address this issue but distribution dynamics can. In the context of Table 2 the movement of three income states experienced by the UK between 1955 and 1992 is very unusual. Thus, a distribution dynamics approach clearly identifies the sharp relative decline experienced by the UK in this period.[25] Table 2 also reveals another interesting story which does not appear in the current literature. The only other economy to experience an upward move of three income states in any of the three sub-periods was Austria between 1955 and 1992. This is the mirror image of the UK experience – why then does the Austrian experience not feature more prominently in the comparative literature?[26]

Finally, the period of greatest convergence, 1955-1992, was also the period (in terms of Table 2) of greatest mobility. In comparing the first and last years of each sub-period, the number of countries that experienced a change in their income states

---

[24] A useful survey, and critique of this literature is provided in: Supple, B., 'British economic decline Floud, R., and McCloskey, D. (eds.), *The economic history of Britain since 1700. Volume 3: 1939-1992* (Cambridge, 1994), pp.318-46. For an interesting perspective on how the analysis of 'declinism' became so prominent in postwar Britain see: Tomlinson, J., 'Inventing "decline": the falling behind of the British economy in the postwar years', *Economic History Review* 44 (1996), pp.731-57.

[25] It is also noteworthy that despite the recent literature on the relative recovery of the UK in the late 1980s, the economy was, in 1992, almost at the bottom of this European league.

[26] A recent example, though, is Boehm, B., and Punzo, L.F. (eds.), *Economic performance: a look at Austria and Italy* (Heidelberg, 1994). See also Matthews, A., and Sagarra, E. (eds.), *Economic performance in two small European economies: Ireland and Austria compared* (Dublin, 1988).

was the same in 1870-1913 and 1922-38. In the first period, four economies moved into a higher income state whilst two moved to a lower income state. Comparing 1922 to 1938 we find that three economies moved into a higher income state whilst three moved to a lower income state. In the post-war period, this comparison revealed that only three economies experienced no change, with six moving to a higher income state and three to a lower income state. It is perhaps surprising that in the turbulent decades of the interwar period the degree of movement shown in Table 2 was much less than in the post-war period and the same as that experienced in the pre-1914 period. In addition, the movement that did occur in the interwar period was modest: in all cases, it was by only one income state.

Although these empirical distributions have some merit, they must be used with great care. For example, in terms of Tables 2 and 3 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 in Tables 2 and 3 is too limited and too crude. In addition, good social science methodology implies that (at least as a starting point) the researcher attempts to use as full an information set as possible. The Quah approach to distribution dynamics provides us with a systematic and methodologically sound analytical framework with which to interrogate our full information set.

The first stage of this approach involves using the information on the actual annual changes in income levels of all the economies in the period under investigation to construct a *transition probability matrix*. This matrix can then be used (either in its original form or in iterated versions) to discuss the degree of mobility and persistence within the distribution. The second stage involves deriving the unique long run equilibrium distribution of the transition probability matrix, this is known as the *ergodic distribution*. The ergodic distribution in effect reveals what is the long run equilibrium shape of the distribution of the regime. If , for instance, the ergodic distribution entails a pronounced single peak or twin peaks, this would point to long-run convergence or polarisation, respectively. We can also compare the ergodic distributions of different regimes. For example, in terms of the stylised story told by Williamson we would expect that the ergodic distributions for 1870-1913 and 1955-1992 would exhibit clear single peaks (reflecting convergence). The ergodic distribution for 1922-1938 should, in comparison, have a much flatter peak or even exhibit polarisation given the large degree of fragmentation of the international economy and the preponderance of different currency and trading blocs in the 1930s.[27]

The Probability Transition Matrix

A stochastic process is a system that evolves over time according to probabilistic laws.[28] Such systems are defined by one or more states and the dynamics are described by the *transition probabilities* that the process will move from one state to another over time. Cox and Miller give an example of a two-state process which describes the likelihood that a rainy day will be followed by a dry day, a dry day by rain, etc. in a certain Middle Eastern city (the rainfall example). Such a process is an example of a *Markov chain*. We shall denote the two states, dry and wet weather, as 0 and 1. The probability that a dry day will be followed by a wet day is given by a, the probability of a wet day being followed by a dry one by ß. Thus, the probability of two consecutive dry days is 1 – a, and the probability of two consecutive wet days is 1 - ß. The transition probability matrix (**P**) for this process is written thus:

$$\mathbf{P} = \begin{array}{c} \\ 0 \\ 1 \end{array} \begin{array}{c} \begin{array}{cc} 0 & 1 \end{array} \\ \begin{bmatrix} 1 - a & a \\ ß & 1 - ß \end{bmatrix} \end{array} \qquad (1)$$

Thus, an important feature of any transition probability matrix is that each of its rows sums to unity. For the rainfall example the observed probabilities were:

[27] See Feinstein, C.H., Temin, P., and Toniolo, G., *The European economy between the wars* (Oxford, 1997).

[28] The following is based on Cox, D.R. and Miller, H.D., *The theory of stochastic processes* (London, 1965, reprinted 1990), chapters 1 and 3.

$$
\mathbf{P} = \begin{array}{c c} & \begin{array}{cc} 0 & 1 \end{array} \\ \begin{array}{c} 0 \\ 1 \end{array} & \left[\begin{array}{cc} 0.750 & 0.250 \\ 0.338 & 0.662 \end{array} \right] \end{array} \qquad (2)
$$

The empirical probability that there would be two consecutive dry days was 0.75 (and hence the probability of a dry day being followed by a wet day was 0.25) whilst the probability of two consecutive wet days was 0.338. The matrix **P** can also be iterated (multiplied by itself) to show what the transition probabilities will be at any future time *n*. For example, in this case:

$$
\mathbf{P}^{5} = \begin{array}{c c} & \begin{array}{cc} 0 & 1 \end{array} \\ \begin{array}{c} 0 \\ 1 \end{array} & \left[\begin{array}{cc} 0.580 & 0.420 \\ 0.568 & 0.432 \end{array} \right] \end{array} \qquad (3)
$$

This shows that if day 1 is dry, the probability that day 6 will be dry is 0.58 or if day 1 is wet, the probability that day 6 will be dry is 0.432.

Our study involves a more complex transition probability matrix in that there are five states rather two. The five states represent different income ranges, with state 1 representing the lowest income range and state 5 the highest income range. The first step is to determine the income ranges representing the five states. These are not imposed by the researchers but are derived on purely empirical grounds.[29] First, all the annual standardised 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

[29] We use TSRF (Time Series Random Field) which is an econometric shell developed by Professor Danny Quah. This calculates transition probability matrices and ergodic distributions. The operation of the programme, as it pertains this material, is further explained in appendix 2.

number of observations. This gives us the values for the partition for each state.[30] For example, for the 1870-1913 sub-period the lower and upper bounds for state 1 were, respectively, 0.473 and 0.672. Thus, it contained economies whose standardised income level in any particular year was more than 47.3% but less than 67.2% of the average income in that year. State 5 contained economies whose standardised income level in any particular year was more than 125.7% but less than 171.1% of the average income in that year.

Having defined the range of each state, we return to the empirical income observations in each year. Each of these annual empirical observations can now be related to a particular state. We then calculate the transition probability matrix by estimating the observed one-year transitions from one state to another for every year in the period under consideration. The matrices for the 12 economies and the 24 economies are shown in Tables 4 and 6.[31] Reading across a row in a matrix gives the probability that, on average in any one year across the period, an economy in a particular state moved to another state; all rows will sum to 1. Thus, reading down the diagonal of the matrix gives the probability that an economy in a particular income state in any given year remained in that same income state in the following year. This provides us with a measure of persistence. The off-diagonal values give the probability that an economy in a particular income state in any given year moved to a different income state in the following year.

Consider the transition probability matrix for the 12 economies in the post-war period given in Table 4. If we read along the row for state 2 we find that, for any given year, the probability that an economy in that income state would be in income state 2 in the following year was 78 per cent. The probability that such an economy would be

[30] 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 is also possible and will be investigated in future research.

[31] The number of observations shown in the transition probability matrices are less than the total number of observations on each economy in each sub-period and are not spread evenly across the five states because in calculating the transition from one year to the next we lose observations at the beginning and end of the period.

in the lower income state 1 in the following year is 6 per cent, whereas the probability that it would be in the higher income state 3 is 16 per cent. Thus, we can conclude that, for one year transitions, the degree of persistence for economies in income state 2 is 78 per cent whilst the degree of mobility is 22 per cent. In addition, it is more than twice as likely that such an economy will move to a higher income state rather than a lower income state.

Tables 4 and 6 also make clear that their dominant characteristic is persistence. Across all the matrices in Tables 4 and 6, the lowest diagonal value is 0.70 and it is not uncommon for the diagonal values to be above 0.90.
This should not be surprising, since we are measuring one-year transitions. Thus, all the movements shown in these tables are moves of one state only: all the one-year transitions involve an economy remaining in the same income state or moving to either the next lowest or next highest income state.

In the longer term, however, mobility becomes a more pertinent characteristic. As we saw with rainfall example, it is possible to iterate the matrix to find out the transition probabilities after *n* years. Table 5 shows the results for the 12 economies in the post-war period after 10 iterations. Thus, the probability that an economy that was in income state 1 in year 1 would still be in income state 1 in year 11 is 52 per cent, whilst there is a 27 per cent probability that it would be in income state 2.[32] The diagonal in Table 5 shows a much lower degree of persistence when compared to the one-year transition probability results: with the exception of economies which start in income state 1, the probability of moving to a different income state is higher than that of not moving. Mobility rather than persistence is the dominant characteristic in Table 5. The other obvious feature of Table 5 is that mobility is not restricted to moving a single class. For example, there is a remote probability that an economy in income state 1 in year 1 could, over the ten year period, be elevated into income state 5.

What does Table 4 reveal about one year transition probabilities of the regimes represented by the three different sub-periods? First, the degree of persistence

---

[32] Again one must exercise caution in discussing these results. Year 1 is 1955 but it would be misleading to then state that year 11 is 1966. This is because the result is the outcome of the iteration of the transition probabilities and is thus a statistical outcome not an empirical outcome.

appears to be greater in 1870-1913 compared to the other two periods: in four out of the five income states the 1870-1913 diagonal values are greater than those in the other two periods.[33] For 1870-1913, there is far greater mobility in income states 2 and 3 compared to the other income states. A similar pattern emerges in 1955-92 in that persistence declines as you move from the outer income states to the middle income state. This pattern of persistence is disrupted in the interwar period. The chance of moving from income state 1 to income state 2 is twice as high in 1955-92 compared to 1870-1913 and is four times more probable in 1922-1938. However, at the top end of the income scale an economy was almost twice as likely to slip down to income state 4 in the latter two sub-periods compared to 1870-1913. For 1870-1913, an economy in income state 2 was far more likely to experience upward mobility rather than downward mobility (as was the case in the post-war period also) whilst the reverse was true for an economy in income state 3. In the interwar period, an economy in income state 4 was far more likely to experience upward mobility.

Turning to Table 6 we find that for the 24 economies the degree of persistence was even stronger than that observed for the 12 economies and in all cases it was, surprisingly, higher after 1973 than before. In terms of mobility, the most striking results relate to income states 3 and 4. In the Golden Age, an economy in income state 3 was twice as likely to move up an income state than to move down an income state whereas the reverse was true for an economy in income state 4. In the post-1973 period, an economy in either income state 3 or income state 4 was three times more likely to move up one income state than to move down one income state. We shall see later that the period 1973-1992 was characterised by polarisation, but the mobility figures shown in Table 5 do not immediately suggest that the cause of this polarisation were the difficulties, and eventual collapse, of the eastern European economies.

The Ergodic Distribution (the Long Run Equilibrium)

[33] In comparison with 1922-38, only income state 3 registers a lower degree of persistence; in comparison with 1955-92, income state 2 is the outlier.

It can be shown that under certain conditions the Markov chain will eventually converge to a steady state. That is, ‘it calculates the unconditional probability of an economy ending up in any particular state.’[34] Thus, another important feature of any transition probability matrix is that it will yield a unique long run equilibrium condition (the ergodic distribution).

To illustrate this, we return to the rainfall example given by Cox and Miller. In this example the process converges (to three decimal places) after ten iterations:

$$\mathbf{P}^{10} = \begin{array}{c} \\ 0 \\ 1 \end{array} \begin{array}{c} \begin{array}{cc} 0 & \quad\quad 1 \end{array} \\ \begin{bmatrix} 0.575 & 0.425 \\ 0.575 & 0.425 \end{bmatrix} \end{array} \qquad (4)$$

This is effectively the equilibrium condition: the probability of day 11 being dry is the same whatever the initial state (each row is identical). The process has converged after ten iterations so that further iterations will yield exactly the same values, but to a greater number of significant figures. Mathematically, $\mathbf{P}^{10}$ is said to be a matrix of rank one, that is the row vector (0.575, 0.425). This can be interpreted as a histogram; that is, in the long run 57.5% of days will have rain and 42.5% of days will be dry. The iterative method for finding the steady-state solution can be tedious, especially when dealing with larger matrices. Fortunately, there is a standard method for computing this analytically (without recourse to iteration) which is explained in appendix 1.

Shape dynamics results

Figures 5 and 6 show the ergodic distributions for our two groups of economies by sub-period. (Values for these distributions are shown in Tables 4 and 6, and summarised in Appendix 3.) The ergodic distribution suggests what the shape of the long run equilibrium distribution would look like.

[34] Proudman, J., Redding, S. and Bianchi, M., ‘Is international openness associated with faster economic growth?’, *Bank of England Working Paper Series* 63 (1997), p.25.

For the 12 economies, it is not clear that the period 1870-1913 is a regime characterised by convergence. There is a sharp single peak (of 0.29, centred on income state 4) but other than that the distribution looks rather flat. This suggests that stratification was at least as strong a force as persistence. It should be noted that persistence was a prevalent feature in the one-year transition probability matrix for this period. The ergodic distribution for the interwar period also has an identifiable single peak (associated with income state 5) but not an obvious story of convergence. However, it is clear that this is not a regime characterised by divergence. The long run equilibrium is characterised by two plateaux (income states 1 and 2 representing the first plateau and income states 3 and 4 the second) which suggests stories of stratification rather than convergence. The post-war period offers possibly the most obvious case of convergence, although the peak is not pronounced.

For the 24 economies, the period of the Golden Age does suggest convergence. Indeed, it is similar to the long run equilibrium distribution for the 12 economies in the post-war period, if anything showing a more pronounced pattern of convergence. In the post-Golden Age era, however, the ergodic distribution exhibits strong polarisation and we see emerging twin peaks (based on income states 1 and 5). This is not merely a product of the collapse of the communist regimes after 1989 as the data are dominated by observations for the pre-1989 period.

Conclusion

The purpose of this exploratory paper is to gauge the extent to which a new approach based on income distribution dynamics might be of use to economic historians interested in long-run income and productivity convergence. In this first application of Quah's technique in historical research, his method has been used to examine the evidence on convergence on the empirical basis of Maddison's widely used per capita income data. Although this work is only in its initial stages, the preliminary findings on distribution shape dynamics, for example, point to complex patterns of stratification, persistence and polarisation that traditional growth regressions do not uncover. In short, this approach to the systematic analysis of income distribution across economies has a lot to offer to the historian.

Appendix 1: a Technical Explanation of the Steady-State Solution

This appendix explains the analytical method for computing the steady-state solution for any transition probability matrix.[35]

For certain transition matrices the steady-state can be found, algebraically, by decomposition into linear combinations of steady-state and transient matrices. this is known as the 'spectral' or diagonal representation of such systems, and is derived from the 'eigenvalues' of the matrix.[36] For the 2×2 matrix

$$\mathbf{P} = \begin{bmatrix} a & b \\ c & d \end{bmatrix}$$

the determinant is defined as the diagonal cross-multiplication of elements $ad - bc$. The *eigenvalues* $\lambda = (\lambda_1, \lambda_2)$ of **P** are defined by the determinant

$$\left| \mathbf{P} - l\,\mathbf{I} = \mathbf{0} \right|$$

where **I** is the identity matrix and **0** is the zero vector.

For the probability matrix of the rainfall example,

$$\begin{bmatrix} 1-a & a \\ b & 1-b \end{bmatrix},$$

the eigenvalue determinant is

[35] This material is based on Cox and Miller, *Theory of stochastic processes*, pp.

[35] Cox and Miller, *Theory of stochastic processes*, pp. 9-83.78-83.

[36] Cox and Miller, *Theory of stochastic processes*, pp. 9-83.

$$|\mathbf{P} - \lambda\mathbf{I}| = \begin{vmatrix} 1-\alpha-\lambda & \alpha \\ \beta & 1-\beta-\lambda \end{vmatrix} = \mathbf{0}\,.$$

This yields the *characteristic* polynomial in λ

$$(1 - \alpha - \lambda)(1 - \beta - \lambda) - \alpha\beta = 0$$

whose roots are the eigenvalues $\lambda_1 = 1$, $\lambda_2 = 1 - \alpha - \beta$. (In all probability matrices the largest eigenvalue is 1.)

In any stochastic system the main condition for the existence of a steady-state is that the eigenvalues of **P** are distinct. When this condition is satisfied there is a matrix **Q** from which is obtained the diagonal or spectral representation of **P**, of the form

$$\mathbf{P} = \mathbf{Q}\begin{bmatrix} \lambda_1 & 0 \\ 0 & \lambda_2 \end{bmatrix}\mathbf{Q}^{-1}\,.$$

In the example,

$$\mathbf{P} = \mathbf{Q}\begin{bmatrix} 1 & 0 \\ 0 & 1-\alpha-\beta \end{bmatrix}\mathbf{Q}^{-1}\,.$$

It can be shown that

$$\mathbf{Q} = \begin{bmatrix} 1 & \alpha \\ 1 & -\beta \end{bmatrix},$$

$$\mathbf{Q}^{-1} = \frac{1}{\alpha+\beta}\begin{bmatrix} \beta & \alpha \\ 1 & -1 \end{bmatrix}$$

whence we derive, by simplification,

$$\mathbf{P}^{\boldsymbol{n}} = \frac{1}{a+b}\begin{bmatrix} b & a \\ b & a \end{bmatrix} + \frac{(1-a-b\ )^{\boldsymbol{n}}}{a+b}\begin{bmatrix} a & -a \\ -b & b \end{bmatrix}.$$

The first matrix on the right hand side is constant: that is, it does not alter on further iteration. Because its rows are identical it is a *row vector*, whose terms β/(α+β), α/(α+β) are the steady–state proportions of states 1 and 2 (in the example, ‘dry’ and ‘wet’) of the system. Here β/(α+β) = 0.575, α/(α+β) = 0.425. Since $1-a-b<1$, the second term on the right hand side is transient, and goes to zero rapidly as *n* increases. The speed of convergence is governed by the magnitude of the *second* largest eigenvalue, which in the rainfall example is 0.412.

Hence

$$\mathbf{P}^{n} = \begin{bmatrix} 0.575 & 0.425 \\ 0.575 & 0.425 \end{bmatrix} + (0.412)^{n}\begin{bmatrix} 0.425 & -0.425 \\ -0.575 & 0.575 \end{bmatrix} \qquad (5)$$

Since 0.412^{10} = 0.00014, the transient component is zero to three significant figures after ten iterations, with $\mathbf{P}^{n}$ = (0.575, 0.425).

Appendix 2: the Operation of the Quah Program

Table A1 shows a fragment of the actual data on which the program operates, in this case for Austria, Belgium, Denmark and Finland during the pre–1914 period, normalized on the average for each year. These economies are shown merely for convenience. The software operates in three steps.

Step 1: the time series are ordered by value as a cross section. In the example, the lowest and highest values are shown in Table A1 below in bold italics: 0.319 and 1.392. All 176 observations fall within these limits. They are divided into five groups of approximately 35–36 observations in ascending order, each group being about one–fifth of the total number. The boundaries of these classes, or states are set to the limits of each group of observations as follows:

State	Interval
1	0.310, 0.514
2	0.514, 0.692
3	0.692, 0.855
4	0.855, 0.962
5	0.961, 1.392

Step 2: the program now treats each economy as a time series. It records any movement across state boundaries, either up or down, during the sample period. In the example, Belgium remains in State 5 until 1898, and then moves from State 5 to State 4 between 1898 and 1899; Austria, from State 4 to State 3 between 1882 and 1883. These transition probabilities are summarized in matrix form.

Step 3: from the transition matrix the steady state is computed by the methods discussed in Appendix 1.

Table A1. An example of the Maddison time series data:

four economies, 1870–1913

Table A1. An example of the Maddison time series data: four economies, 1870–1913

Year	Austria	Belgium	Denmark	Finland
1870	0.944228	***1.391917***	1.011205	0.55999
1871	0.989219	1.372591	0.997165	0.552215
1872	0.960366	1.334565	0.96954	0.536916
1873	0.957456	1.330523	0.966602	0.53529
1874	0.928352	1.290078	0.93722	0.519018
1875	0.919996	1.278466	0.928784	0.514347
1876	0.92211	1.281404	0.930919	0.515529
1877	0.918225	1.276005	0.926997	0.513357
1878	0.908664	1.262718	0.917344	0.508011
1879	0.926997	1.288195	0.935852	0.518261
1880	0.89658	1.245926	0.905144	0.501255
1881	0.894244	1.242681	0.902787	0.49995
1882	0.873479	1.213824	0.881823	0.48834
1883	0.850948	1.182514	0.859077	0.475744
1884	0.851448	1.183209	0.859582	0.476023
1885	0.850795	1.182301	0.858922	0.475658
1886	0.841659	1.169606	0.849699	0.47055
1887	0.824263	1.145431	0.832136	0.460825
1888	0.820818	1.140644	0.828659	0.458899
1889	0.800783	1.112803	0.808433	0.447698
1890	0.793369	1.1025	0.800948	0.443553
1891	0.788481	1.095707	0.796013	0.44082
1892	0.782137	1.086892	0.789609	0.437273
1893	0.782278	1.087087	0.789751	0.437352
1894	0.759875	1.055956	0.767134	0.424827
1895	0.747418	1.038644	0.754558	0.417863
1896	0.729118	1.013214	0.736083	0.407632
1897	0.717598	0.997205	0.724453	0.401191

Table A1. An example of the Maddison time series data: four economies, 1870–1913

1898	0.70261	0.976377	0.709322	0.392812
1899	0.690605	0.959694	0.697202	0.3861
1900	0.686178	0.953542	0.692732	0.383625
1901	0.689587	0.95828	0.696175	0.385531
1902	0.684544	0.951272	0.691083	0.382711
1903	0.6716	0.933285	0.678016	0.375475
1904	0.667861	0.928088	0.67424	0.373384
1905	0.655408	0.910783	0.661669	0.366422
1906	0.641027	0.890799	0.64715	0.358382
1907	0.626029	0.869958	0.63201	0.349997
1908	0.628668	0.873624	0.634673	0.351472
1909	0.61799	0.858785	0.623893	0.345503
1910	0.612786	0.851553	0.618639	0.342593
1911	0.596059	0.82831	0.601753	0.333242
1912	0.58148	0.808049	0.587034	0.325091
1913	0.571223	0.793797	0.57668	***0.319357***

Bibliography

Abramovitz, M., ‘Catching up, forging ahead, and falling behind’, *Journal of Economic History* 46 (1986), pp.385-406.

Barro, R.J. and Sala-I-Martin, X., ‘Convergence across states and regions’*, Brookings Papers* (1991), pp.107-82.

Broadberry, S. N., ‘Convergence: What the historical record shows’, in B. van Ark and N.F. R. Crafts (eds.), *Quantitative aspects of post-war European growth* (Cambridge, 1996), pp.327-346.

Cox, D.R., and Miller, H.D., *The theory of stochastic processes* (London, 1965, reprinted 1990).

Eichengreen, B., ‘Institutions and economic growth: Europe after World War II’, in N. Crafts and G. Toniolo (eds.), *Economic growth in Europe since 1945* (Cambridge, 1996), pp.38-72.

Feinstein, C.H., Temin, P., Toniolo, G., *The European economy between the wars* (Oxford, 1997).

Galor, O., ‘Convergence? Inferences from Theoretical Models’, *Economic Journal* 106 (1996), pp.1056-69.

Maddison, A., *Monitoring the world economy, 1820-1990* (Paris, 1995).

Maier, C., *In search of stability: explorations in historical political economy* (Cambridge, 1987).

Mankiw, N.G., Romer, D., and Weil, D., ‘A contribution to the empirics of economic *Quarterly Journal of Economics*, 107 (1992), pp.407-38.

Milward, A., *The European rescue of the nation- state* (London, 1992).

Bank of England Working Paper Series 63 (1997).

Quah, D.T., 'Empirical cross-section dynamics in economic growth', *European Economic Review* 37 (1993), pp.426-34.

Quah, D.T., 'Twin peaks: growth and convergence in models of distribution dynamics', *Economic Journal* 106 (1996), pp.1045-55.

Quah, D.T., 'Empirics for growth and distribution: stratification, polarization, and *Journal of Economic Growth* 2 (1997), pp.27-59.

Summers, R., and Heston, A., 'The Penn World Table (Mark 5): an expanded set of international comparisons, 1950-1988, Quarterly Journal of Economics 106 (1991), pp.327-68.

Supple, B., 'British economic decline since 1945' in R. Floud and D. McCloskey (eds.), *The economic history of Britain since 1700. Volume 3: 1939-1992* (Cambridge, 1994), pp.318-46.

Tomlinson, J., 'Inventing "decline": the falling behind of the British economy in the *Economic History Review* 44 (1996), pp.731-57.

Toniolo, G., 'Europe's golden age, 1950-1973: speculations from a long-run *Economic History Review* 51 (1998), pp.252-67.

Tortella, G., 'Patterns of economic retardation and recovery in south-western Europe in the nineteenth and twentieth centuries', *Economic History Review* 47 (1994), pp.1-21.

Williamson, J.G., 'Globalization, Convergence, and History', *Journal of Economic History* 56 (1996), pp.277-306.

van Ark, B. and Crafts, N.F.R. (eds.), *Quantitative aspects of post-war European growth* (Cambridge, 1996).