Did Globalization Lead to Segmentation? Identifying Cross–Country Growth Regimes in the Long–Run, 1870–2003*

Gianfranco Di Vaio†‡ Kerstin Enflo§

March 1, 2009

Abstract

Economic historians have stressed that income convergence was a key feature of the 'OECD–club' and that globalization was among the accelerating forces of those process in the long–run. This view has however been challenged, since it suffers from an *ad hoc* selection of countries. In the paper, a mixture model is applied to a sample of 64 countries to endogenously identify the existence of cross–country growth regimes over the period 1870–2003. Results show that growth patterns were segmented in two worldwide regimes, the first one being characterized by convergence, and the other one denoted by divergence. This outcome is consistent with the predictions of recent growth theories. Interestingly, when three historical epochs are analyzed separately (1870–1913; 1913–1950; and 1950–2003), the dynamics which dominate over the whole period emerged only during the Second Global Wave. Therefore, history does not provide unambiguous evidence about globalization and growth.

Keywords: Globalization; Economic growth; Income convergence; Multiple regimes; Mixture models.

JEL Classifications: C52, N10, O47.

*A previous version of the paper was presented at the 3rd Italian Congress of Econometrics and Empirical Economics, *Università Politecnica delle Marche*, Ancona, January 30–31, 2009; and at the 3rd Sound Economic History Workshop, *Gothenburg University*, October 3–4, 2008. We wish to thank Jan Bohlin, Paul Sharp, Jacob Weisdorf and all the participants. We are also grateful to Jörg Baten, Oded Galor, Paolo Malanima and Joseph Zeira for useful comments and suggestions. Usual disclaimers apply.

†Corresponding author. Faculty of Economics, LUISS 'Guido Carli' University. Address: Viale Romania 32, 00197 Roma, Italy. Tel.: +39 06 85225949. Fax: +39 06 5225919. E-Mail: gdivaio@luiss.it.

‡Department of Economics, Finance and Statistics. University of Perugia. Address: Via A. Pascoli 20, 06123 Perugia, Italy.

§Department of Economic History, Lund University. Address: Box 70 83, 220 07 Lund, Sweden.

1 Introduction

Starting with the pioneering studies by Abramovitz (1986) and Baumol (1986), there has been a general consensus that convergence in output per worker, or per capita, took place among the industrialized economies since 1870 (see, among others, Maddison 1987; Feinstein 1988; Broadberry 1993; Tortella 1994; and Toniolo 1998). Economic historians have stressed that this process was fuelled by the two trade booms of the periods 1870–1913 and post–1950, while the inter–war years 1913–1950 were characterized by increased protectionism, slow growth and divergence. It has therefore been argued that history provides an unambiguously positive relationship between globalization and convergence (Williamson 1996, p. 277).

However, the previous evidence relies on the choice of a small and *ad hoc* set of developed countries, that now belong to the so–called 'OECD–club' and that actually converged ex post.[1] In fact, those nations which have not converged since 1870 have been excluded from the sample, due to their present relative poverty. This calls also for a sample selection problem in long–term convergence studies, as it has already been shown by De Long (1988) several years ago. A way to solve these issues is to conduct the convergence analysis on a larger sample, by identifying the growth regimes without prior restrictions.

In order to bring light into this debate, mixture models appear to be useful analytical tools, since they allow us to endogenously identify unknown clusters (i.e. 'clubs') in the data, avoiding the imposition of ex ante selection criteria. The aim of this paper is to analyze long–term growth experiences using an unrestricted sample of 64 nations since 1870, to see whether we can identify the existence of different convergence patterns and whether the OECD–club really stands out during the two phases of globalization. The model we test is in the framework of the beta–convergence hypothesis (Barro and Sala–i–Martin 2004), for which we use data on the initial level, and on the growth rate, of GDP per capita.

We find that the period 1870–2003 is characterized by the segmentation of cross–country growth behavior. Over the long–run, the model identifies two regimes; the first one basically consisting of the OECD–club and being characterized by high growth rates and convergence of per capita income, while the other one, denoted by slow growth and divergence, comprising the rest of the nations. The identification of two growth regimes appears to be in line with recent theoretical predictions about the asymmetrical effects of trade on per capita incomes (Galor and Mountford 2006; 2008).

When the sample is split up into three historical epochs of global and anti–global waves, we do not find evidence of an early converging OECD–club between 1870 and 1913, as previously suggested by Williamson (1996). Instead, it appears that the converging dynamism of the advanced economies only emerged after World War II. These results highlight that assessing economic convergence using an exogenously selected

[1] The choice is usually due to the lack of historical data related to productivity, such as employment or hours worked.

sample may bring misleading results about the growth patterns in the long–run.

In addition, our results are partly analogous to those recently obtained by Epstein et al. (2003), who find that convergence was the key feature for the industrialized economies only during the post–1950 era. Contrary to Mills and Crafts (2000), however, we do find that convergence occurred among the OECD–countries over the long–run. Therefore, the positive causality between globalization and convergence turns out to be historically ambiguous, if the intra–periods dynamics are distinctly analyzed.

The paper is structured as follows: Section 2 focuses on the sources of long–term convergence; Section 3 explains the econometric specification; Section 4 presents the data; Section 5 describes the results; and Section 6 contains some concluding remarks.

2 Sources of long–run convergence

The long–term view to convergence has been central to the writings of economic historians such as Abramowitz (1986) and Baumol (1986). Their studies build on the data set collected by Maddison (1982; 1995) that provides data about GDP per worker and per capita for a large part of the world's countries from 1870. Williamson (1996) focused on convergence in real wages and other factor prices over the long–run. The general picture that have emerged from these studies is that there have been three distinct eras in global history: 1870–1913, 1913–1950 and post–1950, and that convergence was a general feature during the two trade–booms in the late 19th century and after the Second World War.[2]

Many historical studies have speculated especially about the First Global Wave (1870–1913), since it was characterized by globalization in trade, capital flows and migration (Estevadeordal et al. 2003). Taylor and Williamson argue that the period saw dramatic convergence, 'about as dramatic as it has been over the past century and a half', among the present OECD–countries, or an even wider sample of nations (1997, p. 27). This convergence, they assert, was to a large extent accounted for by the massive migration flows between Europe and the New World which helped erase productivity gaps in labor productivity and wages. They estimate such large effects of migration on convergence that it must have been offset by countervailing forces, such as capital accumulation chasing after immigrants and natural resources exploitation.

Taylor (1999) develops a model that explains labor productivity convergence between seven OECD–countries in the late nineteenth century by factor accumulation patterns. The model emphasizes the massive flows of capital and labor to the resource–abundant New World, while it down–plays the importance of technological transfers and human capital. Thus, the sources of convergence during the First Global Wave appear to be different from the post–war period, where technological transfers and human capital have been more broadly emphasized.

[2] Although the post–war period is often divided into two distinct phases: 1950–1973 and the period post–1973, with this latter being demarked by stagflation, slower growth rates and the break–down of the international economic framework established at Bretton–Woods.

A large part of the evidence for convergence during the first epoch of globalization has rested on data on real wages and output per worker. Taylor and Williamson (1997) however simultaneously analyze convergence in GDP per worker and per capita, and in wages. They call upon agnosticism in what variable that provides the 'correct' convergence criterion, although they emphasize that the dynamics of wage and output measures should remain distinct and that the choice of a particular variable should depend on the question under consideration (1997, p. 32). For example, they find that convergence in GDP per capita was slower and less influenced by migration compared to convergence in wages and labor productivity. This result is due to offsetting forces inherent in the algebra of their model, in which labor supply losses suppressed output in the Old World while increasing labor productivity and wages (1997, p. 43). Similarly, O'Rourke and Williamson (1999) acknowledge that the open–economy mechanisms behind convergence in the late 19th century only influenced GDP per capita indirectly. Still, they maintain, that convergence did not only appear in labor markets, but that was also extended, at slower rates, to GDP per capita.

Recently however, Epstein et al. (2003) have questioned whether the period 1870–1914 really was a phase of unconditional convergence fuelled by globalization, even within the OECD–club. Using distribution dynamics methods (Quah 1993; 1996) applied to Maddison's GDP per capita data, they find that the long–run equilibrium of the pre–1914 period was characterized by forces of stratification rather than convergence and argue that this latter was primarily a feature after World War II. In addition, O'Rourke and Williamson (1997) stress that there were large varieties in growth experiences within the Old World pre–1914. For example, although Ireland, Italy and the Scandinavian countries went through a spectacular catch–up with the industrial core, Spain and Portugal lagged behind. The authors also show that globalization was by far the dominant force accounting for these differing economic outcomes and suggest several explanations, covering the failure of capital flows to seek out cheap labor, diversities in schooling, and factor market isolation.

Given the recent debate about convergence and globalization during the First Global Wave, this paper will explicitly test whether we can endogenously identify an early converging OECD–club for this period.

3 Identifying cross–country growth regimes

In order to classify the cross–country growth regimes we make use of a mixture of linear regression. The main feature of this model consists in the ability to uncover heterogeneous patterns of growth in the sample, without imposing *a priori* or *ad hoc* assumptions on the adherence of each country to a specific regime. Mixture models, in the form of mixture densities, have been employed to test the existence of poverty traps or convergence clubs (see, for instance, Paap and Van Dijk 1998), after the pioneering work by Quah (1996) which identified the so–called 'twin peaks' in the income distribution worldwide. This kind of models has been then increasingly applied also to fit the distribution of regional incomes, as in Tsionas (2000), Pittau (2005), and Pittau and Zelli (2006).

As far as we know, mixtures of growth regression have been previously used by Bloom et al. (2003), and by Battisti and Di Vaio (2008) in a regional context. The perspective adopted here, however, differs from the work by Bloom et al. (2003) because we do not explicitly test the club convergence hypothesis. This latter phenomenon would imply that, for each country, the probability of falling in a regime should depend on some specific variables. On the contrary, we consider the probability to belong to a club as a parameter to be estimated in the model.[3] Thus our model can be seen as a test of multiple regimes, independently on the initial conditions of countries, and it aims to provide a correct assessment of which countries fall in the specific regime.

Let's start assuming that for each country i, the average growth rate of per capita income, $g_i = \left[\log(y_{i,T}) - (y_{i,t})\right]/T$, between time t and T, is given by

$$g_i = \alpha_j + \beta_j \log(y_{i,0}) + \varepsilon_{i,j}, \text{ with probability } \phi_j, \tag{1}$$

where $y_{i,0}$ denotes the income per capita level at the beginning of the period, α_j is a constant representing the steady–state determinants of the economy, β_j is a convergence parameter approximating the speed at which the economy reaches the steady–state, $\varepsilon_{i,j} \sim N(0, \sigma_j^2)$ is a random shock affecting the growth rate of the economy, and $j = 1, ..., k$ is the regime which the country belongs to. Expression (1) is usually named 'beta–convergence' equation, after the famous study by Barro and Sala–i–Martin (1992). If β_j is estimated with a negative (positive) sign, the evidence supports that poor countries tend to converge (diverge) to the rich ones.[4]

Assume g_i be distributed as a finite mixture of conditional univariate normal densities:

$$g_i \sim \sum_{j=1}^{k} \phi_j f_{i,j}(g_i \mid \log(y_{i,0}), \beta_j, \sigma_j^2), \tag{2}$$

where

$$f_{i,j}(g_i \mid \log(y_{i,0}), \beta_j, \sigma_j^2) = \frac{1}{\sqrt{2\pi\sigma_j^2}} \exp\left[\frac{-\left(g_i - \alpha_j - \beta_j \log(y_{i,0})\right)^2}{2\sigma_j^2}\right]. \tag{3}$$

The mixing proportions ϕ_j, i.e. the probabilities to belong to a regime, are unknown and should be estimated along with the other parameters of the model.[5] Higher the probability, more precise the identification of the regime is. This aspect makes clear

[3] Such parameter does not depend on the initial level of income, nor it is subject to any threshold in the factor accumulation, as in multiple equilibria models by Azariadis and Drazen (1990), and Galor (1996).

[4] As is well known, the estimated convergence parameter is usually biased if the steady–state determinants vary across the economies. In this framework, we treat this problem allowing for different intercepts across regimes.

[5] The mixing proportions sum to one.

that an *ad hoc* choice of the club may be conducive of misleading results, due to an arbitrary imposition of the probabilities. Particularly, we might erroneously asses that, for instance, a country obeys to a converging pattern, while, on the contrary, it follows a diverging one.

If the set of observations g_i is independently and identically distributed, the joint density or likelihood of the model, L, can be written as

$$L = \prod_{i=1}^{n} \left[\sum_{j=1}^{k} \phi_j \frac{1}{\sqrt{2\pi\sigma_j^2}} \exp\left[\frac{-\left(g_i - \alpha_j - \beta_j \log(y_{i,0})\right)^2}{2\sigma_j^2} \right] \right], \tag{4}$$

or, in its logarithmic form,

$$\log L = \sum_{i=1}^{n} \log \left[\sum_{j=1}^{k} \phi_j \frac{1}{\sqrt{2\pi\sigma_j^2}} \exp\left[\frac{-\left(g_i - \alpha_j - \beta_j \log(y_{i,0})\right)^2}{2\sigma_j^2} \right] \right]. \tag{5}$$

Estimation of the parameters of interest, α_j, β_j, σ_j^2, and ϕ_j, can be conducted maximizing equation (5), subject to the constraint $\sum_{j=1}^{k} \phi_j = 1$. The condition $\sigma_j^2 > 0$ is required to avoid the unboundedness of the likelihood function. Once estimates are obtained, i.e. $\widehat{\alpha}_j$, $\widehat{\beta}_j$, $\widehat{\sigma}_j^2$, and $\widehat{\phi}_j$, each country i is assigned to regime j looking at the posterior probabilities $\widehat{\psi}_{i,j}$, calculated by means of Bayes rule as

$$\widehat{\psi}_{i,j} = \frac{\widehat{\phi}_j f_{i,j}(g_i \mid \log(y_{i,0}), \widehat{\alpha}_j, \widehat{\beta}_j, \widehat{\sigma}_j^2)}{\sum_{j=1}^{k} \widehat{\phi}_j f_{i,j}(g_i \mid \log(y_{i,0}), \widehat{\alpha}_j, \widehat{\beta}_j, \widehat{\sigma}_j^2)}. \tag{6}$$

Basically,

$$\text{country } i \ \in \ \text{regime } j \text{ if } \widehat{\psi}_{i,j} > \widehat{\psi}_{i,m} \ \forall m \neq j = 1, ..., k. \tag{7}$$

The stationary equations of the maximum loglikelihood expressed in (5) are derived by DeSarbo and Cron (1988). Estimation can be straightforwardly dealt with the application of the Expectations–Maximization (EM) algorithm (see Dempster et al. 1977). The EM algorithm works as follows: in the E–step, estimates of ϕ_j and $\psi_{i,j}$ are obtained maximizing the expected loglikelihood, while in the M–step α_j, β_j, σ_j^2 are estimated performing k weighted least squares regressions with weights given by the posterior probabilities. This latter step has been proved to be equivalent to maximum likelihood estimation (see DeSarbo and Cron 1988, for technical details). After the starting values of the parameters are assigned,[6] the algorithm iterates until a specified convergence criterion is achieved.[7] While the procedure provides monotone increasing values of the objective function, convergence to a global optimum is not ensured, due

[6]In absence of specific priors, as in the present case, they are generated randomly.
[7]We set a threshold equal to 0.000001.

to non–convexity of the loglikelihood function. To check the robustness of the results, several trials are carried out.[8]

Making inference, as well as to calculate confidence intervals, requires the variance–covariance matrix of the parameters, which are asymptotically normal, being estimated by maximum likelihood. Louis (1982) shows how to derive the Fisher information matrix in EM environments. The inverse of this matrix provides the estimated covariance matrix (see Turner 2000, for computational aspects).

An open issue relates to the choice of the k components, i.e. regimes, of the mixture. In principle, there is no need of a mixture whenever a one–component model fits the data well. On the contrary, if a mixture model is chosen, the preference of two instead of three components, for instance, has to be proved better. A decision criterion should be adopted, even though no universal rule exists in the literature. We found our decision choice upon two main rules. First, following Turner (2000) we calculate a sequential likelihood ratio (LR) test of k versus $k+1$ components. The test is based on parametric bootstrap, since the likelihood ratio statistic is not regularly distributed.[9] Second, according to Hawkins et al. (2001), we look at the Bayesian information criterion (BIC),

$$BIC = -2 \log L + n_p \log n, \tag{8}$$

where n_p is the number of free parameters, which equals the dimension of the parameter vector minus one. The rationale of this criterion relies on assigning a penalty function to the less parsimonious model, because the loglikelihood can be an increasing function of the components number. The BIC is the recommended criterion for choosing between one and two components in the case of mixtures of linear regressions (see Hawkins et al., 2001). Finally, the model is selected according to the results of the two rules.

4 Description of the data

To estimate the model described in Section 2, we only need data on per capita GDP.

These are taken for 64 countries, over the period 1870–2003, from the database *Historical Statistics for the World Economy: 1-2003 AD*, developed by Angus Maddison, which is downloadable from the Internet page http://www.ggdc.net/maddison (last update: August 2007. For the list of countries see Table 4). GDP per capita is expressed in 1990 International Geary–Khamis dollars.[10]

Some doubts have been raised about the reliability of such data, since they are extrapolated from present–day PPP adjusted GDP levels, on the basis of volume indices of real product. This approach implies that the basket of goods and services used to construct the end–year PPP converter is supposed to be stable over time, something

[8] We run 100 trials, choosing the estimates from the model with the highest loglikelihood value. The results, however, are very stable.

[9] The test is conducted launching 1000 replications.

[10] For detailed notes see Maddison 1995; 2001; 2003.

which can be not very realistic in the long–run (Prados de la Escosura 2000). Unfortunately, alternative estimates usually include too few observations, not allowing to make inference on a larger set of countries. This way we prefer to use the Maddison's data for the present analysis.

5 Discussion of the results

As for the full period 1870–2003, the model clearly identifies two distinct regimes, according to both the adopted selection criteria. The sequential LR test shown in Table 1 strongly rejects the null hypothesis of one versus two components of the mixture (the P–value is 1%), while it is not able to reject the null of two versus three components at any accepted significance level.[11]

[TABLE 1 AROUND HERE]

[FIGURE 1 AROUND HERE]

The values of the BIC reported in Table 2 also suggest the selection of two components. This means that a single one–component growth regression is not the best model to fit the data and produces misleading results, due to the assignment of the same growth pattern to all the countries in the sample.

[TABLE 2 AROUND HERE]

The results from the estimated mixture model are shown in Table 3. Over the period 1870–2003, one regime has been significantly converging (with a beta coefficient of -0.007) and the other one significantly diverging (beta is 0.003). This result indicates that convergence of income per capita towards a common steady–state is not a general feature between the countries of the world in the long–run. The majority of the countries fall in the diverging regime, but the model identifies a convergence club consisting of 19 countries. Those countries can be identified from Table 4 in which the posterior probabilities of a country belonging to each regime are shown. Regime 1 here refers to the convergence club whereas regime 2 stands for divergence. As seen from the table, many of the OECD countries show large probabilities of being assigned to regime 1.

[TABLE 3 AROUND HERE]

[11] The empirical distribution of the test is shown in Figure 1.

A few exceptions emerge: Germany, Greece, Portugal and Spain fall in the diverging regime due to relatively low average growth rates for the full period. Iberia's failure to converge to the OECD is well–documented in economic history and has been explained by a relative failure to industrialize during the late 19th century. Tortella (1994) argues that the Iberian retardation can be assigned to agricultural backwardness and low levels of investment in human capital, as evidenced by low enrolment and literacy rates. It is also the case that the revision of the Maddison's GDP per worker–hour data for Italy and the Iberian countries has cast some doubt on the unconditional convergence hypothesis that was supported by the early studies (see O'Rourke and Williamson 1997, p. 161 for a discussion). Germany is particularly penalized by the slow growth of the inter–war years and thus therefore assigned to the diverging regime. The relative decline of German industrial productivity during the inter–war years has been documented by early scholars, although more recent work has emphasized that Germany's modest economic performance was rather due to large peasant agriculture and backwardness in the service sector (Broadberry 1997).

In addition, USA and New Zealand fall in the diverging regime, since they are both countries having high GDP in 1870 but growing richer over time.

[TABLE 4 AROUND HERE]

Although we find evidence of a long–run OECD club we do not find any converging regime during the First Global Wave 1870–1913, as it can be seen from Table 3. In this case, however, the selection criteria provide a discordant information, since the LR test does not reject the null of one versus two components, while the BIC chooses the two–component model.[12] We prefer the parsimonious specification, given by the one–component model. Anyway, if we would be willing to accept the model with two components, results did not differ in qualitative terms, since two diverging regimes were estimated instead of one (see Table 3).[13] This clearly contradicts the Williamson (1996) notion of convergence between trading nations during the First Global Wave.

The scatter plots in Figure 2 show the estimated models' fit during the different epochs. The full period is displayed in the upper left panel where the converging regime stands out as a range of countries positioned along a straight line with a clear negative slope and a small confidence band. The diverging regime shows up in the slightly positive slope of the fitted line, but the confidence band is much larger. As opposed to the scatter plot from the full period, the plot in the upper right panel does not indicate any convergence club during the first epoch of globalization. The slope of the fitted line, produced by the one–component specification, rather shows divergence and no distinct growth pattern is found among the countries that had the highest logged GDP per capita in 1870.

[12] Shifting from one to two components produces only a small decrease in the value of the BIC (see Table 2).

[13] What is more, the convergence parameters of regime 1 and 2 are almost identical (0.002 and 0.004, respectively).

[FIGURE 2 AROUND HERE]

From convergence theory we would expect the countries to face modest growth rates due to decreasing marginal returns to capital. Instead, looking at the period 1870–1913, these countries are positioned in two clusters on each side of the fitted line. Countries like Australia, Belgium, Ireland, Netherlands, United Kingdom and Uruguay are all below the fitted line and thus exhibit some tendency for slower growth than many countries in the sample. Simultaneously, however, there is a set of initially rich countries that are showing relatively high growth rates and diverging tendencies during the period. These countries cluster above the fitted line and are Austria, Canada, Germany, Denmark, France and Switzerland. Many of these nations have earlier been assigned to a converging regime in wages and labor productivity. Taylor (1999), for example, specifically identifies the labor productivity growth patterns of Germany, France, Denmark, Sweden, UK, USA and Australia as providing evidence that 1870–1914 was an era of convergence, with a speed of about 1 percent per annum (Taylor 1999, p. 1623).

The late 19th century growth of the Scandinavian countries has often been described as a catch–up phenomenon and taken as evidence for the strong forces of convergence during the First Global Wave. O'Rourke and Williamson document a spectacular catch–up in factor prices, but smaller effects in GDP per worker–hour and even less so in GDP per capita (1997, pp. 158–59). A lion's part of the estimated factor price convergence is assigned to mass migration from Scandinavia to the New World. However, the scatter plot in Figure 2 does not suggest that the growth pattern of the Scandinavian countries contributed to a general picture of convergence among the 64 countries during the First Global Wave. On the contrary, Sweden, Norway and Denmark cluster in line with the confidence interval of the fitted line's positive slope. This is because these three countries were initially quite rich compared with the rest of the sample, and showed relatively high GDP per capita growth rates of 1.4–1.6 per cent annually. Thus, although the Scandinavian countries have been singled out as backward and fast–growing in accordance with the convergence hypothesis in the OECD–context, those countries cannot be considered poor in 1870 if compared with the rest of our sample.

So what about capital movements? According to theory we would expect capital to flow from the rich industrial core to the poorer periphery and contribute to convergence. This pattern is for example confirmed in the Scandinavian case. The Scandinavian countries were net importers of capital during these years, and the combination of capital inflow and outward migration has been suggested as a main source behind their growth rates. However, it is also interesting to note that the exporters of capital, such as the initially rich countries Germany and France also showed high growth rates during the same period and that we cannot find any clear pattern of fast–growing capital importers and slow–growing capital exporters. In addition, enormous amounts of capital were placed in the New World, although countries like USA, New Zealand and Australia belonged to the richest countries in the sample in 1870. The tendency for capital accumulation to have a diverging effect on the income distribution during this period has earlier been documented by Taylor and Williamson (1997) among others.

Still, one country that did adhere to the expected convergence pattern was Britain, which was one of the wealthiest country in 1870, who did export large amounts of capital to the New World and also did experience modest growth rates of only 1 percent annually until 1913.

Even though the open market forces of migration and capital did create growth in several parts of the world during the First Global Wave, the data do not unambiguously support the claim of an early converging OECD–club. This becomes especially clear in the larger country sample that we provide. For example, the Scandinavian countries did experience growth rates that were about as high as many of the poorer countries in the sample and their growth experiences do not stand out as remarkable convergence regimes given their initially rather strong positions. It also appears that capital was flowing to countries that were already wealthy in 1870 and therefore acted as a countervailing force to convergence.

As for the period 1913–1950 we identify only one regime. Also in this case, the selection criteria do not provide a clear indication. The LR test selects the two–component specification, while the BIC chooses the model with one component. Anyway, the rejection of the null of one versus two components, produced by the LR test, is not particularly strong, since the P–value is on the significance threshold of 5% (see Table 1). Results do not substantially differ between the two specifications, since both of them support divergence–or at least persistence–of per capita income, as it can be seen from Table 3. This finding underlines the pre–existing historical notion that the inter–war period was characterized by a closing of markets that suppressed the alleged convergence forces from the First Global Wave. The scatter plot in the lower left panel in Figure 2 also shows that the period was characterized by diverging tendencies and modest growth rates. The only exception is Venezuela, an initially poor country showing growth rates of remarkably 5 percent annually due to the discovery of oil in the region.

The post–war period stands out as a period in which a group of 20 countries, mainly members of the OECD, show strong and significant convergence, while the rest of the sample exhibits no clear patterns, e.g. persistence of per capita income. Looking at the selection criteria, the choice of the two–component mixture is clearly supported by the LR test, while the BIC seems to suggest a three–component specification. The identification of the convergence regime, however, is robust to the choice of the components number, since the estimation of a three–component mixture produces the division of the large persistent regime in two smaller ones.[14]

Table 3 displays that the point estimate of the beta–coefficient of regime 1 (-0.016) is roughly twice as large as the estimate from the full period. From the posterior probabilities in Table 4 we note that all but three of the converging countries in regime 1 (Taiwan, Hong Kong and Singapore) belong to the OECD. On the other hand, 44 countries in our sample are assigned to regime 2, which exhibits no significant pattern. This suggests that, excluding the Asian tigers, large parts of the poor world

[14]We do not show the results to save space.

has not experienced the predicted convergence to the OECD. We do however also find a few OECD–countries, like Norway, United States (rich countries that were getting richer), Greece and Portugal (countries with disappointing growth rates given their initial GDP), in regime 2.

In addition, Table 4 indicates that the countries belonging to the long–run convergence club are as good as identical to those singled out for the post–war period.[15] The long–run convergence pattern, that we estimated by means of the mixture of growth regression, thus appears to be completely determined by the dynamics of the post–war period.

6 Concluding remarks

Although forces of globalization have been well–documented for the period 1870–1913, this article shows that growth patterns have been diverse since 1870 and that it was not until after World War II that globalization seems to have been accompanied by convergence for a subset of nations belonging mainly to the OECD. The results highlight that using a restricted sample of nations that have converged ex post, may lead to misleading results about the cross–country growth patterns in the long–run. In fact, this latter were segmented in two worldwide regimes, the first one being characterized by high growth rates and convergence of per capita income, and the other one denoted by slow growth and divergence.

The identification of the two growth regimes seems to be consistent with the predictions of some recent economic theories (Galor and Mountford 2006; 2008), which stress the role of trade for the evolution of the long–term development patterns. In fact, the increase in world trade due to globalization might have affected the growth rate of per capita income asymmetrically according to the comparative advantages of the nations. On the one side, the resource–abundant countries tended to specialize in the production of primary goods, something that depressed the demand for skilled labor, reducing the incentive to invest in human capital and delaying the demographic transition. On the other one, the resource–scarce countries specialized in manufacturing, which activated the need of skilled workers, raising the investment in human capital, spurring the demographic transition and the shift into a sustained stage of growth.

However, looking at the converging 'club', when three historical epochs of global (1870–1913, 1950–2003) and anti–global (1913–1950) waves are analyzed separately, results show that the dynamics which come to dominate over the whole period emerged only during the second globalization stage. Such outcome is in line with the one obtained by Epstein et al. (2003), who reached similar conclusions for a sample of industrialized countries, even using a completely different methodology.

[15] Norway was assigned to the converging regime for the full period, but its remarkable growth after the discovery of oil in the 1970s puts it in the diverging regime when the post–war period is analyzed separately. Spain, on the contrary, was assigned to the diverging regime for the full period, but due to rapid catch–up during the last decades it is assigned to the converging regime after 1950.

Thus, our results indicate that the two trade–booms were not as similar in terms of convergence as what has been previously argued. During the First Global Wave capital did not seek out its highest marginal returns since much of it went to the resource rich New World and although migration acted as a converging force, the net result appears to have been divergence in per capita GDP, even among the industrialized nations. It therefore seems that First Global Wave exhibited a complex inter–play between migration, capital and trade that made it less similar to the postwar period, in terms of convergence, than what previously has been recognized. According to our findings, history only ambiguously supports the positive relationship between globalization and convergence.

Future research should investigate why globalization has brought about convergence in some countries but not in others and why these forces appear to be limited to the last decades only. Especially the trade flows and their composition need to be further investigated.

References

[1] Abramovitz, M. (1986) Catching up, forging ahead, and falling behind. *Journal of Economic History*, *46(2)*, 385–406.

[2] Azariadis, C., Drazen, A. (1990) Threshold externalities in economic development. *Quarterly Journal of Economics*, *105(2)*, 501–526.

[3] Barro, R.J., Sala–i–Martin, X. (1992) Convergence. *Journal of Political Economy*, *100(2)*, 223–251.

[4] Barro R.J., & Sala–i–Martin, X. (2004) *Economic growth*. Cambridge, MA: MIT Press.

[5] Baumol, W.J. (1986) Productivity growth, convergence, and welfare: What the long-run data show. *American Economic Review*, *76(5)*, 1072–1085.

[6] Battisti, M., Di Vaio, G. (2008) A spatially filtered mixture of β–convergence regressions for EU regions, 1980–2002. *Empirical Economics*, *34(1)*, 105–121.

[7] Bloom, D.E., Canning, D., Sevilla, J. (2003) Geography and poverty traps. *Journal of Economic Growth*, *8(4)*, 355–378.

[8] Broadberry, S.N. (1993) Manufacturing and the convergence hypothesis: What the long–run data show. *Journal of Economic History*, *53(4)*, 772–795.

[9] Broadberry, S.N. (1997) *The productivity race: British manufacturing in international perspective 1850–1990*. Cambridge: Cambridge University Press.

[10] Dempster, A.P., Laird, N.M., Rubin, D.B. (1977) Maximum likelihood from incomplete data via the EM algorithm. *Journal of the Royal Statistical Society B*, *39(1)*, 1–38.

[11] De Long, J.B. (1988) Productivity growth, convergence, and welfare: Comment. *American Economic Review*, *78(5)*, 1138–1154.

[12] De Sarbo, W.S., Cron, W.L. (1988) A maximum likelihood methodology for clusterwise linear regression. *Journal of Classification*, *5(2)*, 249–282.

[13] Epstein, P., Howlett, P., Schulze, M.–S. (2003) Distribution dynamics: Stratification, polarization, and convergence among OECD economies, 1870–1992. *Explorations in Economic History*, *40(1)*, 78–97.

[14] Estevadeordal, A., Frantz, B., Taylor, A.M. (2003) The rise and fall of world trade, 1870–1939. *Quarterly Journal of Economics*, *118(2)*, 359–407.

[15] Feinstein, C. (1988) Economic growth since 1870: Britain's performance in international perspective. *Oxford Review of Economic Policy*, *4(1)*, 1–13.

[16] Hawkins, D.S., Allen, D.M., Stromberg, A.J. (2001) Determining the number of components in mixture of linear models. *Computational Statistics and Data Analysis*, *38(1)*, 15–48.

[17] Galor, O. (1996) Convergence? Inferences from theoretical models. *Economic Journal*, *106(437)*, 1056–1069.

[18] Galor, O., Mountford, A. (2006) Trade and the great divergence: The family connection. *American Economic Review*, *96(2)*, 299–303.

[19] Galor, O., Mountford, A. (2008) Trading population for productivity: Theory and evidence. *Review of Economic Studies*, *75(4)*, 1143–1179.

[20] Louis, T.A. (1982) Finding the observed information matrix when using the EM algorithm. *Journal of the Royal Statistical Society B*, *44(2)*, 226–233.

[21] Maddison, A. (1982) *Phases of capitalist development*. Oxford and New York: Oxford University Press.

[22] Maddison, A. (1987) Growth and slowdown in advanced capitalist economies: Techniques of quantitative assessment. *Journal of Economic Literature*, *25(2)*, 649–698.

[23] Maddison, A. (1995) *Monitoring the World Economy 1820–1992*. Paris: OECD.

[24] Maddison, A. (2001) *The World Economy: A Millennial Perspective*. Paris: OECD Development Centre.

[25] Maddison, A. (2003) *The World Economy: Historical Statistics*. Paris: OECD Development Centre.

[26] 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(1)*, 68–91.

[27] O'Rourke, K.H., Williamson, J.G. (1997) Around the European periphery 1870–1913: Globalization, schooling and growth. *European Review of Economic History*, *1(2)*, 153–190.

[28] Paap, R., van Dijk, H.K. (1998) Distribution and mobility of wealth of nations. *European Economic Review*, *42(7)*, 1269–1293.

[29] Pittau, M.G. (2005) Fitting regional income distributions in the European Union. *Oxford Bulletin of Economics and Statistics*, *67(2)*, 135–161.

[30] Pittau, M.G., Zelli, R. (2006) Empirical evidence of income dynamics across EU regions. *Journal of Applied Econometrics*, *21(5)*, 605–628.

[31] Prados de la Escosura, L. (2000) International comparisons of real product, 1820–1990: An alternative data set. *Explorations in Economic History*, *37(1)*, 1–41.

[32] Quah, D.T. (1993) Empirical cross–section dynamics in economic growth. *European Economic Review*, *37(2/3)*, 426–434.

[33] Quah, D.T. (1996) Twin peaks: Growth and convergence in models of distribution dynamics. *Economic Journal*, *106(437)*, 1045–1055.

[34] Taylor, A. (1999) Source of convergence in the late nineteenth century. *European Economic Review*, *43(9)*, 1621–1645.

[35] Taylor, A., Williamson, J. (1997) Convergence in the age of mass migration. *European Review of Economic History*, *1(1)*, 27–63.

[36] Toniolo, G. (1998) Europe's golden age, 1950–1973: Speculations from a long–run perspective. *Economic History Review*, *51(2)*, 252–267.

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

[38] Turner, R.T. (2000) Estimating the propagation rate of a viral infection of potato plants via mixtures of regressions. *Applied Statistics*, *49(3)*, 371–384.

[39] Tsionas, E.G. (2000) Regional growth and convergence: Evidence from the United States. *Regional Studies*, *34(3)*, 231–238.

[40] Williamson, J.G. (1996) Globalization, convergence, and history. *Journal of Economic History*, *56(2)*, 277–306.

Tables and figures

Table 1. Sequential LR test of k versus $k+1$ components*

Period	1 vs 2	2 vs 3
1870–2003	26.6 (.01)	3.97 (.81)
1870–1913	3.85 (.76)	-
1913–1950	17.2 (.05)	5.99 (.71)
1950–2003	26.3 (.00)	23.3 (.17)

* P-values between parentheses. Maximum number of components: $k = 3$.

Table 2. Bayesian information criterion (BIC)

	1 Comp.	2 Comp.	3 Comp.	Selected
Period 1870–2003	-491.4	-497.2	-484.5	2 Comp.
Period 1870–1913	-510.2	-512.5	-510.9	2 Comp.
Period 1913–1950	-407.2	-403.6	-400.0	1 Comp.
Period 1950–2003	-378.8	-385.1	-391.8	3 Comp.

Table 3. Cross–country growth regimes: estimation results

	1 Component (OLS)	2 Components (ML)	
		Regime 1	Regime 2
	Period 1870–2003		
Constant	.005	.070***	-.003
Log of p.c. GDP 1870	.002	-.007***	.003**
Weight (%)	-	23	77
R-squared	.03	-	-
Log-likelihood	249.9	263.1	-
	Period 1870–1913		
Constant	-.010	-.004***	-.013
Log of p.c. GDP 1870	.003***	.002***	.004***
Weight (%)	-	18	82
R-squared	.15	-	-
Log-likelihood	259.2	270.8	-
	Period 1913–1950		
Constant	-.015	-.038***	-.008
Log of p.c. GDP 1870	.003*	.006***	.002
Weight (%)	-	19	81
R-squared	.05	-	-
Log-likelihood	207.7	216.4	-
	Period 1950–2003		
Constant	.045***	.168***	.044***
Log of p.c. GDP 1870	-.003	-.016***	-.003
Weight (%)	-	25	75
R-squared	.03	-	-
Log-likelihood	193.6	207.1	-

***, **, * denote statistical significance at 1%, 5%, and 10%, respectively. In grey: mixture results. Dependent variable: average growth rate of per capita GDP (various periods). Observations: 64.

Table 4. Posterior probabilities*

	Period 1870-2003		Period 1950-2003	
	Regime1	Regime 2	Regime 1	Regime 2
Austria	0.60	0.40	0.84	0.16
Belgium	0.56	0.44	0.70	0.30
Denmark	0.56	0.44	0.69	0.31
Finland	0.86	0.14	0.77	0.23
France	0.60	0.40	0.74	0.26
Germany	0.46	0.54	0.68	0.32
Italy	0.58	0.42	0.75	0.25
Netherlands	0.57	0.43	0.68	0.32
Norway	0.56	0.44	0.49	0.51
Sweden	0.66	0.34	0.64	0.36
Switzerland	0.57	0.43	0.60	0.40
United Kingdom	0.61	0.39	0.61	0.39
Ireland	0.51	0.49	0.69	0.31
Greece	0.03	0.97	0.01	0.99
Portugal	0.02	0.98	0.01	0.99
Spain	0.42	0.58	0.77	0.23
Australia	0.62	0.38	0.68	0.32
New Zealand	0.21	0.79	0.02	0.98
Canada	0.62	0.38	0.68	0.32
United States	0.10	0.90	0.20	0.80
Albania	0.00	1.00	0.00	1.00
Bulgaria	0.00	1.00	0.00	1.00
Czechoslovakia	0.00	1.00	0.00	1.00
Hungary	0.00	1.00	0.00	1.00
Poland	0.00	1.00	0.00	1.00
Romania	0.00	1.00	0.00	1.00
Russia (USSR)	0.00	1.00	0.00	1.00
Yugoslavia	0.00	1.00	0.00	1.00
Argentina	0.00	1.00	0.00	1.00
Brazil	0.00	1.00	0.00	1.00
Chile	0.00	1.00	0.00	1.00
Mexico	0.00	1.00	0.00	1.00
Uruguay	0.00	1.00	0.00	1.00
Venezuela	0.00	1.00	0.00	1.00
Jamaica	0.00	1.00	0.00	1.00
China	0.00	1.00	0.00	1.00
India	0.00	1.00	0.00	1.00
Indonesia	0.00	1.00	0.00	1.00

* In grey: regime 1.

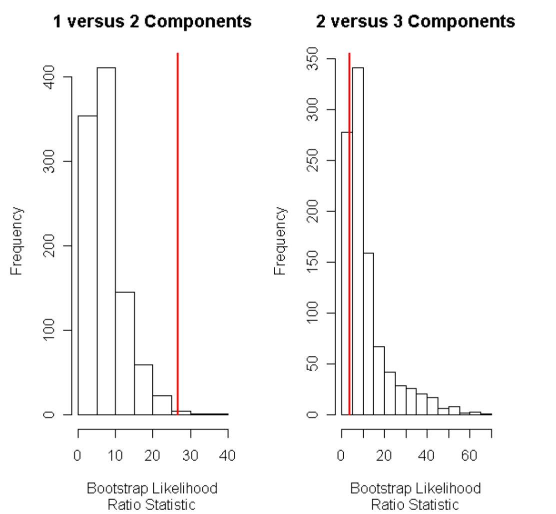
Table 4. Continued*

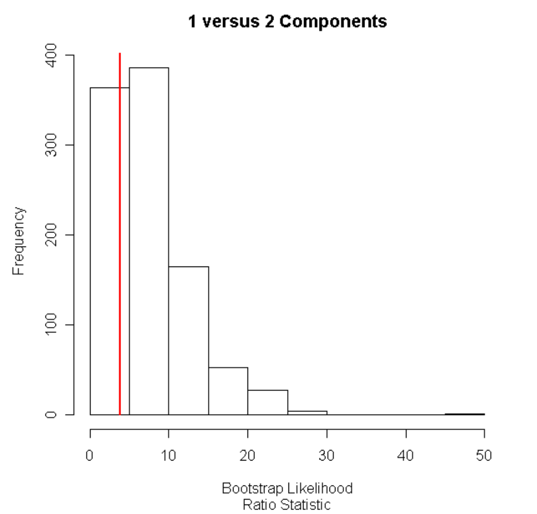
Japan	0.99	0.01	0.96	0.04
Philippines	0.00	1.00	0.00	1.00
South Korea	0.95	0.05	0.99	0.01
Thailand	0.00	1.00	0.00	1.00
Taiwan	1.00	0.00	1.00	0.00
Burma	0.00	1.00	0.00	1.00
Hong Kong	0.99	0.01	0.83	0.17
Malaysia	0.00	1.00	0.00	1.00
Nepal	0.00	1.00	0.00	1.00
Singapore	0.99	0.01	0.95	0.05
Sri Lanka	0.00	1.00	0.00	1.00
North Korea	0.00	1.00	0.00	1.00
Vietnam	0.00	1.00	0.00	1.00
Iran	0.00	1.00	0.00	1.00
Iraq	0.00	1.00	0.00	1.00
Jordan	0.00	1.00	0.00	1.00
Lebanon	0.00	1.00	0.00	1.00
Syria	0.00	1.00	0.00	1.00
Turkey	0.00	1.00	0.00	1.00
West Bank and Gaza	0.00	1.00	0.00	1.00
Algeria	0.00	1.00	0.00	1.00
Egypt	0.00	1.00	0.00	1.00
Ghana	0.00	1.00	0.00	1.00
Morocco	0.00	1.00	0.00	1.00
South Africa	0.00	1.00	0.00	1.00
Tunisia	0.00	1.00	0.00	1.00

* In grey: regime 1.

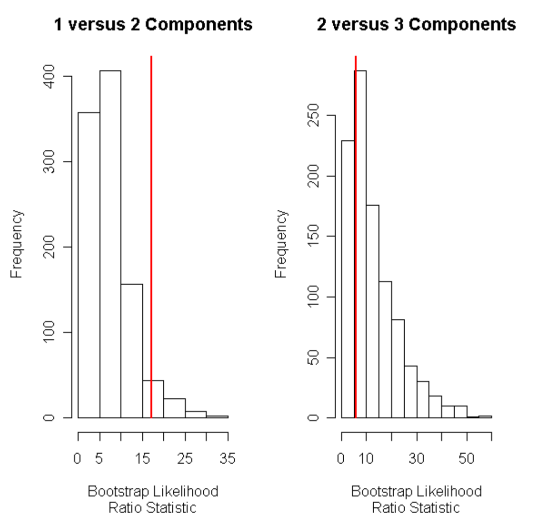
Figure 1. Empirical distribution of the LR test*

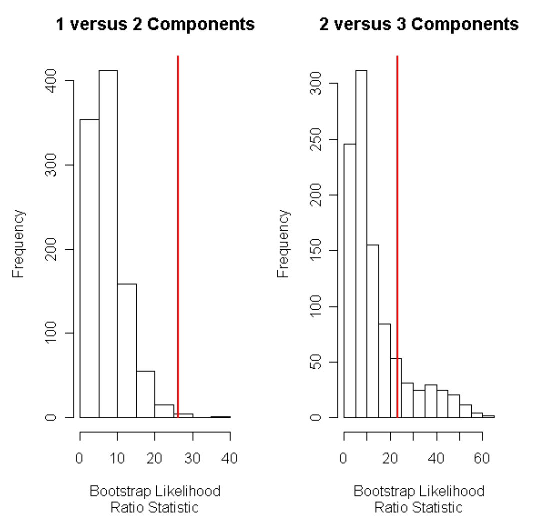
Period 1870–2003	*Period 1870–1913*





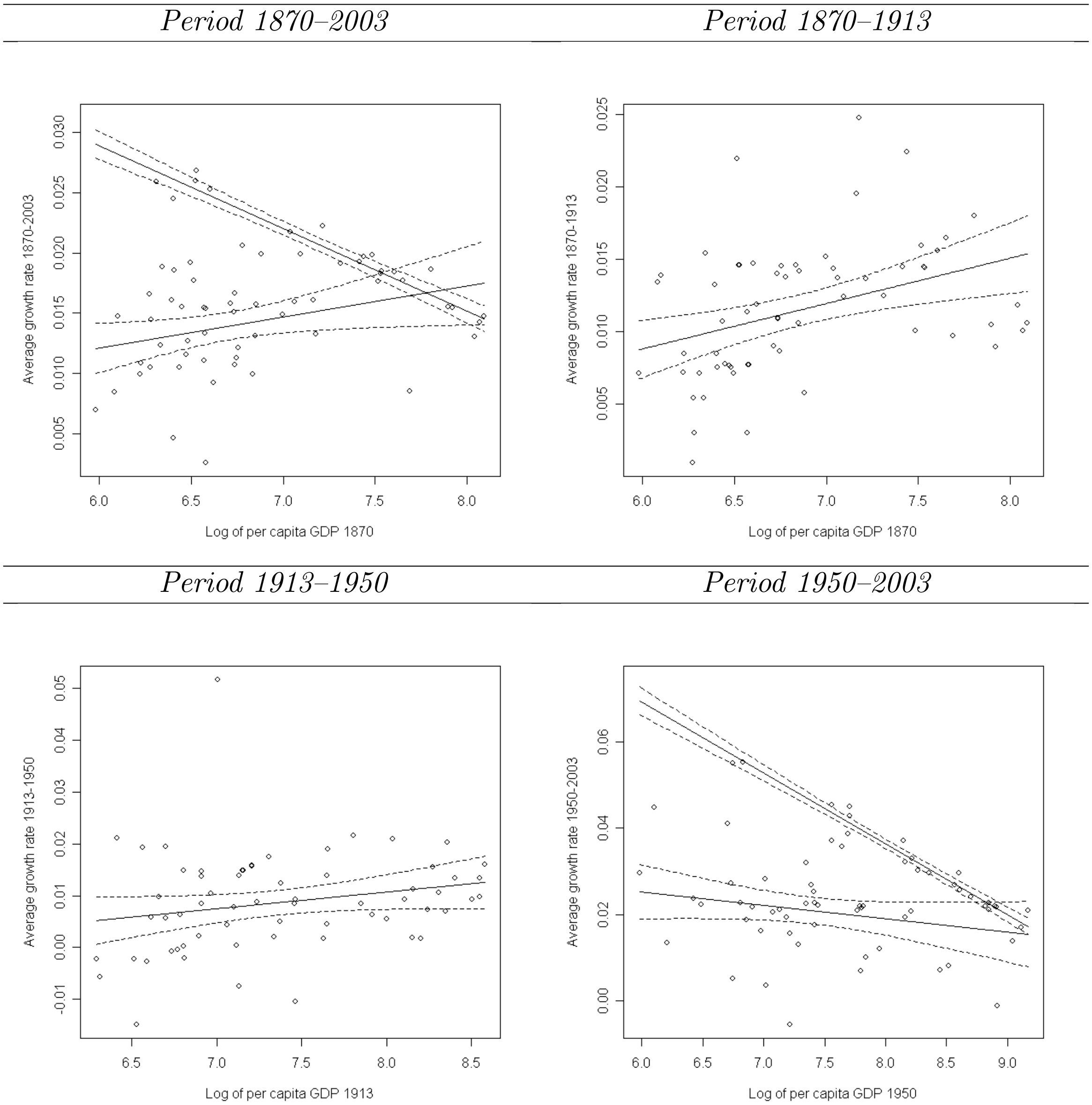
Period 1913–1950	*Period 1950–2003*





*1000 bootstrap replications.

Figure 2. Cross–country growth regimes: model's fit*



*Solid line: regression fit; dotted line: confidence band at 95% level.