ECONOMIC CONVERGENCE OF INCOME DISTRIBUTION WORLDWIDE FROM 1986 TO 2000.[1]

Jhon James MORA RODRIGUEZ
Department of Economics
ICESI University, Calle 18 No. 122-135.
e-mail: jjmora@icesi.edu.co (Cali, Colombia).

AND

José Javier NÚÑEZ VELÁZQUEZ
Department of Statistics, Economic Structure and I.E.O.
University of Alcalá de Henares (Madrid, España).
e-mail: josej.nunez@uah.es

Forthcoming in Journal of Economic Studies (2009)

Abstract:
This paper discusses the role of Markovian transitions related to the economic convergence among countries. Thus, an overview of several classical approaches is developed, including an analysis of fallacies exposed through the literature. The paper continues with a study of convergence clubs through the number of modes found in the distribution involved and in the use of Markov chains. In that way, this study is based upon RGDP data, measured in 100 countries in the period 1986-2000, which makes it possible for us to estimate the number of modes found across the 1986-1991 and 1992-2000 sub-periods. Next, different transition matrices are estimated across the analyzed period 1986-2000, using Markov chains methodology, and then ergodic distributions are obtained to study the existence of *Twin Peaks* along income distribution dynamics. The last results across the period 1986-2000 show the evolution of countries into *convergence clubs*, instead of the existence of economic convergence.

Key Words: Economic Growth, Economic Convergence, Markov Chains.
JEL Classification: C10, O40.

[1] The authors are thankful for the valuable suggestions of Isaías Salgado-Ugarte at the Autonomous University of México, Rubén Garrido Yserte and Tomás Mancha Navarro at the University of Alcalá, and those who attended the seminars at which preliminary versions of this work were presented and the anonymous referees. The authors also appreciate the assistance provided by Juana Domínguez Domínguez at the University of Alcalá. It is clear that they take full responsibility for any mistakes in this paper.

1. Introduction

In the 1990s one of the most interesting discussions in the field of economic growth took place with regard to the potential convergence of the GDP in countries over time. On the one hand, in reaction to the works of authors in the 1990s showing the inexistence of absolute convergence, traditional academics proposed the idea of conditional convergence (Barro and Sala-i-Martín, 1992b) claiming that poor countries would converge into rich countries at an annual growth rate of 2%. The reaction, suggested by Quah's papers (1993, 1996a, 1996b, 1997), was overwhelming and irrefutable in that he proved that there was no convergence into a single distribution, but on the contrary rich countries would be increasingly richer while poor countries would be poorer. He further claimed that over time "Convergence Clubs" would be established [Baumol (1986), Durlauf (1993), and Durlauf and Johnson (1995)].

By the way, several empirical works showed that the series from which convergence characteristics were inferred were not stationary, but on the contrary had at least a unit root[2]. This implied that behind an alleged stable state (summarized in β) there were much more complex dynamic processes than anyone had ever thought which were not incorporated in the equations of convergence. Quah (1996), for instance, showed that behind an annual convergence rate of 2% there might be a unit root process.

We discuss two important issues on the convergence hypothesis. Firstly, the discretization process really matters. If we use quartiles or quintiles, the ergodic distribution doesn't show Twin Peaks: the process shows an equiprobabilistic ergodic (stationary) distribution in the long term. Secondly, the Twin Peaks results need a Markov (time) homogeneous chain. That is, if the dynamic processes of income evolution follow a Markov chain, then Chapman-Kolmogorov´s equation must be satisfied.

The first section of this paper presents the idea of convergence clubs and provides a review of the number of modes in the distribution of the RGDPL (real GDP per worker using the Laspeyres index) for 100 countries in the period from 1986 to 2000. The second section discusses the results obtained from the relevant transition matrices and analyzes the existence of *Twin Peaks* in the distribution of income. The third section discusses the Markovian and homogeneity hypothesis about the stochastic process that underlies income distribution. The most relevant conclusions are summarized in the last section.

2. Convergence and Convergence Clubs: How many modes are there in the distribution of income?

Barro and Sala-i-Martin (1992a) conducted a traditional analysis of the hypothesis of convergence. The empirical foundation of convergence across countries is a regression of convergence that results from linearizing the equations of consumption growth (Ramsey-Keynes rule) and capital stock. The convergence equation shows growth based on GDP levels. Following Sala-i-Martin (2000), the convergence equation will be as follows:

$$T^{-1}(Y_i(T) - Y_i(0)) = a - \left(\frac{1-e^{-\beta T}}{T}\right)Y_i(0) + \mu_i(T) \qquad (1)$$

---

[2] With regard to the GDP of the United States, Nelson and Plosser (1982), and Perron (1989) demonstrated the existence of a unit root in this series. In Colombia Posada (1993), and Mora and Salazar (1994), among others, also showed that there is a unit root. Furthermore, Mora (1997) showed the existence of non-linear patterns in the series of the Colombian GDP. These empirical results, which are also found in many series of GDPs in countries across the world, corroborate that it is inadequate to conduct a cross-section regression in an analysis of convergence.

In equation (1) above, Y_i is the logarithm of the GDP in a given country i, 0 the initial year, T is the final year, a is a constant that depends on the technological parameters of the model, and β is the convergence ratio. "Convergence ratio" means the rate at which an economy approaches a stationary state or the speed at which the economies of poor countries grow faster than those of rich countries. Some empirical results about the estimates of this convergence equation are shown below:

Table 1: Estimates of β-type convergence in different countries

	-Cond.	β regional	Time to close the gap
Europe: 90 regions [1950-1990])		0.015	66 years
Japan: 47 prefectures [1955-1990])		0.019	53 years
Spain: 17 communities [1955-1987]		0.023	43 years
EEUU: 48 states [1880-1990]		0.017	54 years
Barro and Sala-i-Martín (1991) (1960-1988)	0.02		50 years
Colombia: 1960-1989 (Cárdenas)	0.0324	0.041	24 years

Source: Prepared using data from Cárdenas (1993), Sala-i-Martín (2000).

As shown in Table 1 above, estimates of convergence are close to 2%, which is a fairly interesting phenomenon. On the other hand, if we use the calculations by Cardenas (1993), Colombia is undoubtedly a more successful case of regional convergence than the United States, Japan, and Spain, considering that the annual convergence rate in Europe was 1.5% against 4% in Colombia.[3]

In addition to the concept of β-type convergence, there is the concept of σ- or sigma-type convergence showing whether dispersion across countries decreases over time (see Quah, 1996, for details). Based on this criterion, the fact that an economy shows convergence to a stationary state is not relevant; what is actually important is the evolution in the distribution of all countries in a region.[4] At any point in time the dispersion of the GDP is calculated based on the typical deviation of the countries included in the region.

In light of the theories of convergence or existence of "twin peaks", there is an increasingly strong possibility that countries could actually come together in groups or conglomerates with similar income levels called "*convergence clubs*". In this respect, both Durlauf (1993) and Durlauf and Johnson (1995) found evidence of the existence of such convergence clubs.[5]

Quah (1993) also discussed the existence of these convergence clubs, thereby emphasizing that: firstly, countries are endogenously selected in groups, which means that they do not work in an isolated fashion; secondly, specialized production allows countries to tap on economies of scale; and thirdly, ideas are an important driver of growth. Therefore, the existence of coalitions or convergence clubs is an endogenous result across all countries, and the dynamics of convergence will depend on the initial distribution of the characteristics of each

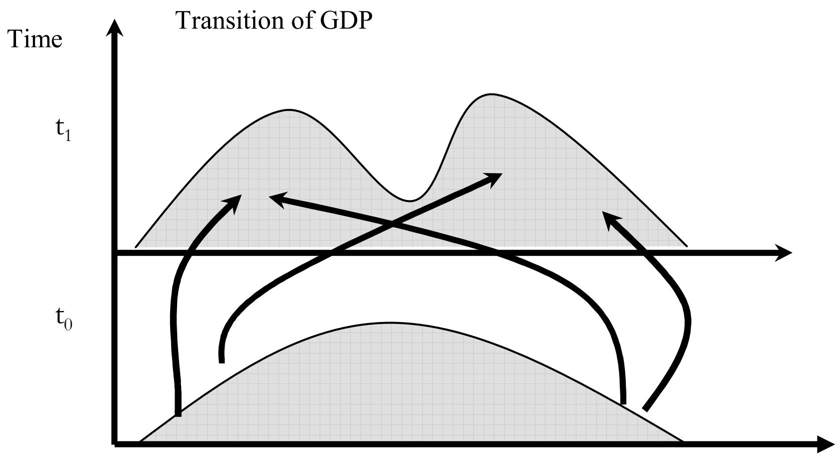
[3] If the conditions remain the same, this means that it would take 19 years to close the gap (disparity) between regions 19 years ($0.052 \times 19.2 \cong 1$). The works of Meisel (1993), Mora and Salazar (1994), Rocha and Vivas (1998), Birchenall and Murcia (1997), Bonet and Meisel (2001) and Mora (2003) showed that there was no convergence of the regions in Colombia during the period of time analyzed by Cardenas (1993).

[4] However, Sala-i-Martín (2000, 196) made it clear that beta-type convergence is a necessary, yet not sufficient, condition for the existence of sigma-type convergence.

[5] Durlauf (1993) incorporated non-linearities by using an interval-based regression method that allows countries to have different growth behaviors. Durlauf and Johnson (1995) use the same method and reject the linear method commonly used in cross-section regressions, which demonstrate the hypothesis of convergence. Baumol (1986) also discussed the existence of convergence clubs.

country. These dynamics include polarization, which makes rich countries richer and poor countries poorer while the middle class fades out in such a way that stratification is consistent with multiple modes in the distribution of the GDP in these countries, thereby providing coverage to diverging economies [Quah (1993,1997)]. Thus, the transition of a country from a period of time t_0 to a period of time t_1 is shown in the following figure:

Figure 1: Transition from a unimodal to a bimodal distribution.



Source: Based on the related figure in Quah (1996).

Figure 1 shows that economies that begin with an initial distribution of the GDP at t_0 could form coalitions or convergence clubs over time in such a way that distribution tends to a bimodal structure at t_1. The exact number of coalitions and their composition depend on the initial distribution of the GDP. Thus, when the GDP is similar in all the countries, there is a trend towards a single coalition which generates convergence to equality and, therefore, the existence of a single distribution mode. On the other hand, if the GDP is uneven, then multiple convergence clubs may be established that match an identical number of relative modes.

In consideration of the above, one of the implications of the theory of economic convergence is that income distribution must have a single mode. As shown in figure 1 above, when countries make a transition to convergence clubs, the number of modes increases as suggested by Quah. Thus, the number of distribution modes might shed a light on the existence or absence of economic convergence.

Bianchi (1997) examined the number of modes in 119 countries in 1970, 1980, and 1989 and found evidence of bimodality in 1980 and 1989, thus corroborating Quah's findings (El-Gamal and Ryu (2003) reached to the same results).

We used the Silverman multimodality test to prove the number of modes in the distribution of the GDP. Consider a random variable x, like say GDP, with realizations x_i i=1,..,n. Let $\hat{F}_n$ be the empirical accumulated distribution of the sample. Thus, the density distribution of the data is

$$\hat{f}(x) = \int_{-\infty}^{+\infty} w_n(x,u)d\hat{F}_n(u) = n^{-1}\sum_{j=1}^{n} w_n(x,X_j), \quad x \in \Re,$$

Where w_n(x,u) satisfies $\int_{-\infty}^{+\infty} w_n(x,u) = 1$. Let $w_n(x,u) = h^{-1} K\left(\frac{x-u}{h}\right), x \in \Re$ where K is Kernel and h>0 is the bandwidth. The Kernel density estimator of f(x) is (see Silverman, 1986; Härdle, 1990)

$$\hat{f}(x) = (nh)^{-1} \sum_{j=1}^{n} K\left(\frac{x - X_j}{h}\right), \quad x \in \Re$$

And there are many discussions about the critical bandwidth, h. The critical bandwidth is defined as the small possible h producing a density. Silverman (1986) and Izenman and Sommer (1988) and Bianchi (1997) showed the procedure for obtaining the optimal bandwidth. We used ASH-WARPing's non-parametric technique [Scott (1992), Häerdle (1991)] implemented by Salgado-Ugarte, Shimizu and Taniuchi (1995)

Following we tested the number of modes in the distribution of the GDP. We used the RGDPL from Penn World, Table 6.1 (Summer and Heston, 1991). The table only includes 100 countries where information for the entire period from 1986 to 2000 was available. The results of the multimodality test are shown in the following table:[678]

Table 2: Estimation of the number of modes

Year	Number	Mode value	Bandwidth
1986	2	(28854.6719, 62388.4805)	4874.1
1987	2	(29401.1758, 63602.5430)	5000.2
1988	2	(29311.0801, 61948.2383)	5197
1989	2	(26176.1641, 59657.3047)	5072.9
1990	2	(23818.5645, 60447.1133)	5003.9
1991	2	(25715.0313, 59929.1016)	5448.1
1992	3	(26961.4805, 62696.1406, 89657.6172)	5349.5
1993	3	(28794.3477, 64954.6914, 98213.2813)	5580.3
1994	3	(29207.3047, 67918.5703, 100371.1328)	5795.1
1995	3	(27092.0156, 68071.5391, 103814.7813)	5691.6
1996	3	(28340.0879, 69064.0781, 106453.9453)	5953.8
1997	3	(30283.8164, 73475.4844, 114433.1094)	6205.7
1998	3	(30638.1680, 73430.3203, 119260.9688)	6330.2
1999	3	(29391.3008, 71267.3984, 122247.0000)	6502.5
2000	3	(28964.4160, 74917.5781, 127554.8281)	6962.6

Source: Author's calculations using data from Penn World (Table 6.1).

[6] Jones (1997, 20) contends that the GDP per capita is generally used in the analysis of the evolution of income. However, considering that a significant share of production in many developing countries is not market-related, using GDP per capita could bias the true income in those countries. In such cases, using the GDP per worker would seem to be the best option [Jones (1997)]. Another reason for using the GDP per worker is that the results will be comparable to those obtained by Jones (1997).

[7] The bandwidth was calculated using Silverman's (1986) equations 3.30 and 3.31, and so its expression is $h = 0.9n^{-1/5}$min (standard deviation, inter-quartile range/1.34), whereas n is the number of observations. The number of modes was calculated using 100 changes of histograms (bar charts) averaged for the Gaussian kernel.

[8] Alternatively one could calculate the number of modes for the logarithm of the RGDP. However, as thoroughly shown by Bianchi (1997), logarithmic transformation simplifies the structure of income distribution because if fails to consider the outliers and makes it difficult to distinguish a bimodal from a unimodal distribution [El-Gamal and Ryu (2003:2)].

Table 2 above shows that the distribution of the RGDPL had two modes in the period from 1986 to 1991. Although the first mode was close to USD 28,000 in 1986, it dropped to USD25,000 in 1991, and the second mode, which started at close to USD62,000 in 1986, decreased to approximately USD$59,000 in 1991. From 1992 to 2000 the number of modes increased giving rise to a third mode of about USD 127,000 in the year 2000.

Therefore, Table 2 clearly shows the existence of three modes in the year 2000. These findings do not provide any supporting evidence of economic convergence at all, but on the contrary reveal that convergence clubs arise.

These findings are different from those reported by Sala–i-Martín (2002)[9] and raise a question as to the historical evolution of income distribution and its future changes on a long-term basis.

3. Transition of the GDP and use of Markov's chains

The use of Markov chains for analyzing convergence started with Quah's work (1993, 1996a, 1996b, 1997). Quah showed that the likelihood of a transition on the ends will be higher than the likelihood in the middle, which, following Baumol (1986), Durlauf (1993), and Durlauf and Johnson (1995) would result in the existence of "twin peaks" or two convergence clubs.

Unlike Quah, Sala-i-Martín (2002) also discussed the existence of "Twin Peaks", but his methodology is different from the one based on the use of a transition matrix because he builds individual density functions for each year and country to calculate the distribution of income worldwide. Sala-i-Martin (2002) believes that any evidence of the existence of two modes in the 1980s disappeared in 1998 [Sala–i-Martín (2002:14)], which means that the second peak ("twin peaks") disappears and a middle class arises worldwide: "*The forecast is that twin peaks are going to disappear in the long term* [Sala-i-Martin (2000: 226)]". Like Sala-i-Martin, Kremer, Onatski and Stock (2001) did not find any statistical evidence of the existence of "twin peaks" using data from a 5-year interval. Their findings show that the transition matrix leads to an ergodic distribution in which most countries are in the category of countries with the highest income. Yet the findings of Kremer, Onatski and Stock (2001) are arguable because, as shown by Quah (2001), if an annual interval is taken, then the methodology proposed by Kremer, Onatski and Stock (2001) does not reject the existence of "twin peaks".

In order to examine the evolution of income, a model will be devised based on a first-order homogeneous Markovian process. If λ_t is a measure of income distribution in the countries at time t, then income distribution will evolve as follows:

$$\lambda_{t+1} = M\lambda_t \qquad (2)$$

As posed by Quah (1996a, b), M represents the intra-distributional dynamics of income in the different countries in equation (2). Therefore, it rigorously contains more information than just consolidated statistics such as means or typical deviations [Quah (1996b: 1370)]. In fact, M incorporates conditional probabilistic information that allows a transition from income

[9] Sala-i-Martin found only one mode in 1998, but his approach to arriving to this result is arguable. First of all, the Gaussian kernel depends on the selection of an optimal bandwidth. If the bandwidth is not optimal, a lower or higher number of distribution modes may occur. The bandwidth depends on the standard deviation, which was selected based on questionable criteria such as the average standard deviation from 1970 to 1998. A less "informal" criterion would be to calculate the standard deviation using a *bootstrap* and taking into account the mean and the number of countries in the sample.

distribution at t to the appropriate (t+1), which is supposed to be invariant during the entire period, thus generating a hypothesis of homogeneity in the related Markov's process. If equation (2) is iterated, then:

$$\begin{aligned}
\lambda_{t+2} &= M\lambda_{t+1} = M(M\lambda_t) = M^2\lambda_t \\
\lambda_{t+3} &= M\lambda_{t+2} = M(M^2\lambda_t) = M^3\lambda_t \\
\lambda_{t+4} &= M\lambda_{t+3} = M(M^3\lambda_t) = M^4\lambda_t \\
\ldots\ldots &= \ldots\ldots\ldots\ldots = \ldots\ldots\ldots\ldots\ldots\ldots = \ldots\ldots\ldots \\
\lambda_{t+s} &= M\lambda_{t+s-1} = M(M^{s-1}\lambda_t) = M^s\lambda_t
\end{aligned} \qquad (3)$$

Now, if one takes the limits when s → ∞, then equation (3) can describe the long-term behavior or ergodic distribution of the GDP in the countries which is the same as the evolution of the GDP.(Quah, 1993b) [10][11]. In order to calculate an estimate (3), Quah breaks down the evolution of income into the five following states: 1/4 of the mean income worldwide, 1/2 the mean income worldwide, the mean income worldwide, twice the mean income worldwide, and more than twice the mean income worldwide. This could be considered a "discretization" system to obtain the states. Consequently, although the space of states determined by income is continuous, it is grouped in a small number of classes which are the new states and determine the nature of the chain shown in the model[12]. Quah's [1993, 1996(a, b), 1997] results are shown in the following table[13]:

Table 3:
RGDPL. Interannual transition matrix with regard to the mean income (1962-1985).

Transitions	¼	1/2	1	2	∞
(456)	0.97	0.03			
(643)	0.05	0.92	0.04		
(639)		0.04	0.92	0.04	
(468)			0.04	0.94	0.02
(508)				0.01	0.99
Ergodic distribution	0.24	0.18	0.16	0.16	0.27

Source: Adapted from Quah (1993b, 1996a, 1997).

The row titled "ergodic distribution" in Table 3 above shows Quah's limit distribution (1993b). These results reveal that there is accumulation at the tails and tapering towards the middle of the ergodic distribution, thus providing indication of the existence of convergence clubs at low and high ends of the RGDPL, while the middle class disappears [Quah (1996:1372)].

---

[10] Following Quah (1993b), if $F_{t+1} = M\ F_t$, , whereas M has the abovementioned characteristics, then this equation "…is like a standard first-order auto-regression, except that its values are distributions (rather scalars or vectors of numbers), and it contains no explicit disturbance or innovation" (page 429).

[11] See a Paap and van Dijk (1998: 1289) for a clearer interpretation of the M matrix.

[12] However, the discretization process is not neutral and should be investigated in more detail because it may incorporate some of the dynamic effect of time on the process. This is what happens, for example, if quartiles are taken into account in the distribution of income when income moves in time (see for example, Hungerford (1993) or Pena and Nuñez (2003)).

[13] As it can be easily checked from Tables 3 to 9, all the transition matrices (M) show its respective models as finite regular -or ergodic- Markov Chains (aperiodical and irreducible ones). Hence, their ergodic distributions always exist and each one of them matches up with the unique solution of the respective equation $\lambda = M'\lambda$. (details can be found for instance in Parzen, 1962, and Iosifescu, 1979).

Nonetheless, this accumulation at the tails is not exclusively the result of using a stochastic kernel as suggested by Quah (1993b). It also occurs when Markov chains are used and the states are grouped. The following are some of the authors' comments with regard to the application of Quah's proposed methodology for the period from 1986 to 2000.

i) The ergodic distribution resulting from the "discretization" by means of quartiles of salary distributions displays an even behavior.
ii) Although the ergodic distribution changes when intermediate information between both periods is not considered, the dynamics of evolution remains unchanged.

With regard to the first of these two comments, this result is also found in Quah's works. Quah (1993a) first discretized the RGDPL using the quartiles of income distribution and achieved the following results:

Table 4: RGDPL. Inter-annual transition matrix by quartiles (1962-1984)

Transitions	0.25	0.50	0.75	1
(667)	0.96	0.04		
(690)	0.04	0.93	0.03	
(667)		0.03	0.95	0.02
(690)			0.02	0.98
Ergodic	0.25	0.25	0.25	0.25

Source: Adapted from Table 3.1. (Quah 1993a).

Table 4 shows a transition matrix for a one-year period. The transition column shows the total observations, and each cell in the matrix shows a recount of annual transitions in the countries over a period of 24 years. The last row above shows an ergodic distribution (not shown by Quah (1993a)) which has been calculated through equation (3), as it was explained before. Although data in Table 4 show an unbalance in the number of transitions, it does not have an impact on the bistochasticity of the transition matrix[14]. On the other hand, the ergodic distribution shows an opposite result to what Quah affirms (1993a). Quah contends that the structure of the likelihood of transition in Tables 3.1 and 3.2 shows the same ergodic distribution as in Quah (1993b), i.e. *"stacking of the mass of likelihood and tapering towards the middle."* As can be shown, in no case is there stacking on the ends, but there is rather an even distribution of likelihood in all states. Quah (1993a) reported the following transition matrix using "quintiles" to discretize:

Table 5: RGDPL. Inter-annual transition matrix by quintiles (1986-2000)

No. of transitions	0.20	0.40	0.60	0.80	1
(529)	0.95	0.05			
(552)	0.05	0.90	0.05		
(530)		0.05	0.90	0.05	
(551)			0.05	0.93	0.02
(552)				0.01	0.99
Ergodic distribution	0.1675	0.1675	0.1675	0.1675	0.33

Source: Adapted from Table 3.2 (Quah 1993a).

In the case presented in Table 5 above, the ergodic distribution (which was also not reported by Quah (1993)) shows different likelihoods between states 5 and the rest of the states. This result is clearly due to an unbalance in the number of transitions which has an effect on the

[14] A matrix with elements in the range from 0 to 1 is called bistochastic or doubly stochastic if not only the sum of its rows, but also that of its columns equal one.

bistochasticity of the transition matrix. This effect is observed after proving deviations with respect to the unit of the sums of the elements in the columns. In this case, the sum of the elements in the last column is apparently what makes the mode move to the last quintile.

The results shown below were obtained from the same process of discretizing data - in quartiles - about the RGDPL with respect to the mean income in the 100 different countries worldwide for the period from 1986 to 2000, thereby following Quah's (1993a) approach.

Table 6: RGDPL, Inter-annual transition matrix by quartiles (1986-2000)

No. of transitions	0.25	0.50	0.75	1
(350)	0.9857	0.0143		
(350)	0.0143	0.9371	0.0486	
(350)		0.0486	0.9371	0.0143
(350)			0.0143	0.9857
Ergodic distribution	0.25	0.25	0.25	0.25

Source: Author's calculations using data from Penn World (Table 6.1).

As shown in Table 6, the likelihood in each of the four states in which the discretization process was carried out equals 25%. When quintiles were used in the distribution, the results were as follows:

Table 7: RGDPL Inter-annual transition matrix by quintiles (1986-2000)

No. of transitions	0.20	0.40	0.60	0.80	1
(280)	0.9571	0.0429			
(280)	0.0429	0.9321	0.0250		
(280)		0.0250	0.9321	0.0429	
(280)			0.0429	0.9500	0.0071
(280)				0.0071	0.9929
Ergodic distribution	0.20	0.20	0.20	0.20	0.20

Source: Author's calculations using data from Penn World (Table 6.1).

Similarly, the results of the bistochastic transition matrix show convergence into an equiprobabilistic vector determined by 20%.

Tables 4, 5, 6, and 7 show that the selection of the states based on a quantilical discretization process is associated with the bistochastic nature of the transition matrix [Formby, Smith and Zheng (2004)]. It is also well-known that the performance of this kind of matrix produces an equalizing effect [e.g. Nuñez (2006)] that leads to an equiprobabilistic ergodic distribution in the long term[15].

In consistence with the second comment above with regard to the selection of states, a contrast of the effect of the dynamic changes that occur in the selected states could consist of comparing the results obtained during the entire period of time and failing to post any resulting interannual changes. This idea is discussed more thoroughly below.

If one applies Quah's (1993b, 1996a and b, 1997), i.e. if discretization of the mean income values for each year is taking into account, the results for the period from 1986 to 2000 are the following:

---

[15] Evidently any deviation with regard to the bistochasticity has a displacing effect on the mode or modes in consistence with the size of the deviation (see Table 5 above).

Table 8: RGDPL, Inter-annual transition matrix with regard to the mean income (1986-2000)

No. of transitions	1/4	1/2	1	2	∞
(345)	0.9884	0.0116			
(219)	0.0274	0.9589	0.0137		
(322)		0.0155	0.9627	0.0217	
(235)			0.0383	0.9447	0.0170
(279)				0.0108	0.9892
Ergodic distribution	0.4278	0.1811	0.1591	0.0901	0.1419

Source: Author's calculations using data from Penn World (Table 6.1).

As shown in Table 8 above, the ergodic distribution reveals accumulation of likelihood on the ends. However, in the long term the likelihood of countries achieving an income level within a range from a quarter to half the mean income worldwide put higher attraction (makes them more attractive) than the likelihood of countries achieving more than twice the mean income worldwide.

Again if one takes the initial and final years of the sample instead of taking the entire period, the following results are obtained:

Table 9: RGDPL, Inter-annual transition matrix with regard to the mean income from 1986 to 2000.

No. of transitions	1/4	1/2	1	2	∞
(24)	0.9167	0.0833			
(16)	0.2500	0.6875	0.0625		
(23)		0.1304	0.7391	0.1304	
(18)			0.2778	0.6111	0.1111
(19)				0.0526	0.9474
Ergodic distribution	0.5793	0.1930	0.0925	0.0434	0.0917

Source: Author's calculations using data from Penn World (Table 6.1).

As shown in Table 9 above, the results are very similar to those listed in Table 8, but the state attraction in the "ranging from a quarter to half the mean income worldwide" is stronger than in the previous case. Yet the evolution of income does not change, which means that there are still "twin peaks" in the distribution of income.

4. -Homogeneity and Markovian nature of income distribution

The various results discussed in the above section regarding the long-term dynamics of income distribution have often been attributed to the "discretization process". This process is needed to obtain the states for the Markov homogeneous chain and its influence has been demonstrated. Thus, the selection of the aforementioned states has a clear impact on the structure of the transition matrix, which means that it determines the configuration of the stationary vector and, therefore, also the characteristics of the dynamic behavior of income distribution.

On the other hand, the most common approach consists of using distribution quantiles as states in the Markov chain (quartiles and quintiles in this case), giving rise to dynamic changes in the states when transition matrices evolve over time. Hence, as it has been demonstrated and expected, the long-term behavior is determined by a uniform probabilistic vector, and therefore all states have the same likelihood.

A possible solution to the problem of selecting the states - which has often been used in literature - is a gradual decrease of the widths leading in the end to the use of stochastic kernels. These allow us estimating the transition density function between consecutive periods - in a non-parametric fashion - which is still assumed to be stable throughout the period (homogeneity of the underlying Markov process). The most relevant characteristic of these stochastic kernels is again determined by the predominant structure of the diagonal elements as demonstrated in the previously reviewed transition matrices.

This idea allows us speculating about the characteristics associated with the mobility of the distribution. Thus, an analysis of mobility associated to the involved distributions was conducted. To this end, the authors used Shorrocks' [16] mobility index, which is defined by the following expression:

$$M(P) = \frac{n - tr(P)}{n - 1} \qquad (4)$$

Since (4) is a standardized ratio, it reflects total immobility when it equals zero, but it indicates total mobility when it equals 1. In the above formula, *n* is the number of states, *P* is the corresponding transition matrix, and *tr(P)*, its trace.

Table 10: Transition matrices using consecutive years (1986-2000).

1986-1987	1/4	½	1	2	∞
¼	100	0	0	0	0
½	0	100	0	0	0
1	0	0	100	0	0
2	0	0	11.11	88.89	0
∞	0	0	0	0	100

1987-1988	**1/4**	**1/2**	**1**	**2**	**∞**
1/4	100	0	0	0	0
1/2	0	100	0	0	0
1	0	4	92	4	0
2	0	0	6.25	93.75	0
∞	0	0	0	0	100

1988-1989	1/4	½	1	2	∞
¼	100	0	0	0	0
½	5.88	94.12	0	0	0
1	0	0	91.67	8.33	0
2	0	0	6.25	93.75	0
∞	0	0	0	5.26	94.74

1989-1990	1/4	1/2	1	2	∞
1/4	92	8	0	0	0
1/2	0	94	6.25	0	0
1	0	0	100	0	0
2	0	0	0	88.89	11.11
∞	0	0	0	0	100

1990-1991	1/4	½	1	2	∞
¼	100	0	0	0	0
½	5.88	94.12	0	0	0
1	0	0	87.5	12.5	0
2	0	0	6.25	87.5	6.25
∞	0	0	0	5	95

1991-1992	1/4	1/2	1	2	∞
1/4	96	4.2	0	0	0
1/2	6.3	94	0	0	0
1	0	4.6	95.5	0	0
2	0	0	0	100	0
∞	0	0	0	0	100

1992-1993	1/4	½	1	2	∞
¼	100	0	0	0	0
½	5.88	94.12	0	0	0
1	0	0	100	0	0
2	0	0	0	94.44	5.56
∞	0	0	0	0	100

1993-1994	1/4	1/2	1	2	∞
1/4	100	0	0	0	0
1/2	6.3	88	6.25	0	0
1	0	0	95.2	4.76	0
2	0	0	5.88	94.12	0
∞	0	0	0	0	100

[16] See detailed characteristics in Prais (1955) and Shorrocks (1978).

1994-1995	1/4	½	1	2	∞
¼	96.15	3.85	0	0	0
½	0	92.86	7.14	0	0
1	0	0	100	0	0
2	0	0	0	100	0
∞	0	0	0	0	100

1995-1996	1/4	1/2	1	2	∞
1/4	100	0	0	0	0
1/2	0	100	0	0	0
1	0	0	100	0	0
2	0	0	0	100	0
∞	0	0	0	0	100

1996-1997	1/4	½	1	2	∞
¼	100	0	0	0	0
½	0	100	0	0	0
1	0	4.35	95.65	0	0
2	0	0	5.88	94.12	0
∞	0	0	0	4.76	95.24

1997-1998	1/4	1/2	1	2	∞
1/4	100	0	0	0	0
1/2	0	100	0	0	0
1	0	4.4	95.7	0	0
2	0	0	11.8	88.24	0
∞	0	0	0	0	100

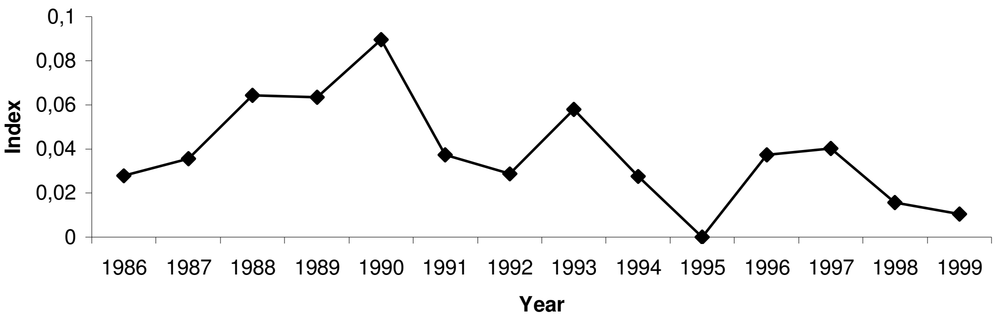
1998-1999	1/4	½	1	2	∞
¼	100	0	0	0	0
½	6.25	93.75	0	0	0
1	0	0	100	0	0
2	0	0	0	100	0
∞	0	0	0	0	100

1999-2000	1/4	1/2	1	2	∞
1/4	100	0	0	0	0
1/2	0	100	0	0	0
1	0	4.2	95.8	0	0
2	0	0	0	100	0
∞	0	0	0	0	100

Source: *Author's calculations using data from Penn World* (Table 6.1).

Table 10 shows all interannual transition matrices for the period 1986-2000. Given the above stated arguments with regard to the selection of quantiles as states of the Markov chains involved, we decided to make a selection of the states based on each year's mean income worldwide as explained above. Figure (3) below shows the values of Shorrocks' mobility indexes for all the resulting matrices.

Figure 3: Shorrocks' mobility indexes from 1986 to 2000



Source: Author's calculations using data from Penn World (Table 6.1).

Figure 3 clearly shows that interannual transitions were extremely rigid because the mobility index reached a peak of 0.09 for the transition matrix from 1990 to 1991. Although 1991 was not the breakpoint year to build the sub-periods analyzed in the previous section, it was certainly the last year in which there were two modes in the distribution of income worldwide. It was a 3-mode distribution from 1992 to 2000.

This fact strongly supports the assumption that the homogeneity hypothesis of involved Markov chains could fail[17]. Thus, it would then account for the lack of consistence between the number of distribution modes and the conclusions drawn following Quah's proposed technique.

Moreover, if one acknowledges a Markov chain model as the stochastic process underlying the income dynamics, then Chapman-Kolmogorov equation - which reflects the behavior of transition matrices - necessarily has to hold true.

$$P(1986,2000) = \prod_{t=1986}^{1999} P(t,t+1) \qquad (5)$$

Thus, if we use the transition matrices in Table 10, then the second element of equation (5) above is as shown in Table 11 below.

Table 11. Inter-annual transition matrices (1986-2000).

States	1/4	½	1	2	∞
1/4	0.8764	0.1124	0.0115	0	0
1/2	0.2764	0.5917	0.1268	0.0074	0.0004
1	0.0203	0.1327	0.6593	0.1738	0.0280
2	0.0042	0.0432	0.3055	0.4951	0.1576
∞	0.0001	0.0020	0.0244	0.1039	0.8700

Source: Author's calculations using data from Penn World (Table 6.1).

A comparison of the matrices in Tables 9 and 11 provides a tool to demonstrate the existence of very significant differences, especially concentrated on the diagonal of both matrices because of the rigidness of income distribution. It is worth noting that the Markovian hypothesis evidently does not hold true with respect to income distribution by countries worldwide. Therefore, we could also contend that the hypothesis of homogeneity is not meaningful and thus, it can explain why the methodology based on the Markovian hypothesis is unable to show the three modes presented herein.

5. Conclusions

Danny Quah's work about convergence proved to be transcendental because he showed that beta-type convergence regressions only provide proof of unit roots, and a β value of 0.02 verifies that the GDP series is not stationary. This means that the long discussed 2% global convergence would be nothing but the evidence that the series is a unit root process.

When Quah discusses sigma-type convergence, his results are similarly illustrative in that he concludes the following: "Stability of convergence could occur in different potential scenarios or worlds; some would have gaps between intervals while others would have GDP convergence clubs or polarization."

In his articles published in 1993, 1996b y 1997, in calculating the first-order matrix, Quah uses the stochastic kernel as a non-parametric estimator of the conditional density function when he refers to M. This could essentially be considered a Markov chain, in which the classification of income in different stages (discretization process) has been progressively polished [18] to deal with the inconveniences apparently associated with the selection of states, whilst there is actually a trend to a Markovian process in discrete time.

[17] Shorrocks (1976) presented a similar argument with regard to the analysis of income mobility (1976).
[18] Quah and Durlauf (1998:62) reported this in similar terms.

Many authors have made attempts to minimize Quah's argument. Jones (1997), for example, is one of those outstanding authors [emphatically quoted by Sala-i-Martin (2000)] who, having calculated the M matrix, was able to show that the percentage of rich countries is going to increase in the long term, while the number of poor countries will decrease under the assumption that the economies of these countries will grow at a faster rate than they decrease.

The idea of convergence clubs is irrefutable when one reviews the distributions in 1980, 1985 and 1990 [Quah (1993, 1996a, 1996b, 1997), Bianchi (1997), El-Gamal and Ryu (2003)]. Our results show that the distribution of the RGDPL in the period from 1986 to 2000 is not unimodal. And our results corroborate the occurrence of "twin peaks" in the distribution of income. However, after conducting analysis of the number of modes in global distribution of income, there are only 2 modes until 1991, but thereafter 3 modes are found. This leads us to consider the existence of convergence clubs rather than the existence of convergence to a bimodal distribution, which has not yet occurred.

Lastly, it has been shown that the above mentioned inconsistencies are present because the Markovian hypothesis fails to hold true. In fact, the high rigidness of the distribution leads to the evidence that Chapman-Kolmogorov condition, which is necessary for these kinds of processes, is not met. This fact also renders the condition of homogeneity meaningless which would otherwise be at least arguable.

The solutions to the above problems are complex in each case, but one could argue that, if countries whose income distribution is clearly stagnant are classified in a different group, then the Markovian condition could hold true for the rest. This would be along the lines of the proposal made by Bartholomew (1973) or Shorrocks (1976) within the framework of occupational mobility. This is undoubtedly a worthy research topic to be addressed in detail in the future.

References

Barro R.J. and Sala-I-Martin, X. (1992a). *Economic Growth*. McGraw-Hill.
-------------. (1992b). “Convergence". *Journal of Political Economy*, 100(2): 223-251.

Bartholomew, D.J. (1973). *Stochastic models for social processes.* 2ª ed. John Wiley.

Baumol, W.J. (1986). “Productivity Growth, Convergence, and Welfare: What the Long-Run Data Show”. *American Economic Review*, 76(5):1072-1085.

Bianchi, M. (1997). “Testing for convergence: evidence from non-parametric multimodality test”. *Journal of Applied Econometrics*, 12(4): 393-409.

Birchenall, J and Murcia G. (1997). “Convergencia regional: una revisión del caso Colombiano”. *Desarrollo y Sociedad*, 40: 273-308.

Bonet, J and Meisel, A. (2001). “La convergencia regional en Colombia: una visión de largo plazo, 1926-1995”. In *Regiones, ciudades y crecimiento económico en Colombia*, A. Meisel (ed.), Colección de Economía Regional, Banco de la República.

Cárdenas, M. (1993). “Crecimiento y convergencia en Colombia: 1950-1990”. *Revista Planeación y Desarrollo*, 24: 53-80.

Cohen, D. (1995). "Test of the Convergence Hypothesis: Some Further Results". *CEPR Working Paper* 1163.

Durlauf, S. (1993). "Nonergodic economic growth". *Review of Economic Studies*, 60:349-366.

Durlauf, S. and Johnson, P.A. (1995). "Multiple Regimes and Cross Country Growth Behavior". *Journal of Applied Econometrics*, 10(4):365-384.

El-Gamal, M.A. and Ryu, D. (2003). "´Convergence´ Hypotheses are Ill-Posed: Non-stationarity of Cross-Country Income Distribution Dynamics". Available at: http://www.ruf.rice.edu/~dhryu/growth.pdf.

Friedman, M. (1992). "Do Old Fallacies Ever Die". *Journal of Economic Literature*, 30 (2): 2129-2132

Formby, J.P.; Smith, W.J. and Zheng, B. (2004). "Mobility measurement, transition matrices and statistical inference". *Journal of Econometrics,* 120: 181-205.

Härdle, W. (1991). *Smoothing Techniques, With Implementations in S*. Springer. New York.

Hungerford, T.L. (1993). "U.S. income mobility in the seventies and eighties". *Review of Income and Wealth*, 39(4): 403-417.

Iosifescu, M. (1979). *Finite Markov processes and their applications.* John Wiley & sons.

Jones, Ch. I. (1997). "On the Evolution of the World Income Distribution". *Journal of Economics Perspectives*, 11(3): 19-36.

Kremer, M., Onatski, A. and J. Stock.(2001). "Searching for Prosperity". *NBER Working Paper* No. 8250.

Mankiw, G.; Romer, D. and Weil, N. (1992). "A Contribution to the Empirics of Economic Growth". *Quarterly Journal of Economics*, 107: 407-413.

McCloskey, D. (1994). "1780-1860: A Survey". In *The Economic History of Britain Since 1700*, R. Floud, D. McCloskey, (eds.), vol.1, Cambridge, pp. 242-71.

Meisel, A. (1993). "Polarización o convergencia? a propósito de Cárdenas, Pontón y Trujillo". *Coyuntura Económica*, 23(2):153-160.

Mora, J.J. (2003). "Crecimiento y convergencia: A propósito de Quah". *Estudios Gerenciales*, 89:57-72.

Mora, J.J. (1997). "No-linealidades y ciclos asimétricos en el PIB Colombiano". *Estudios Económicos*, 12(2):183-195.

Mora, J.J. and Salazar, B. (1994). "Fábula y trama en el relato de la convergencia". *Boletin Socioeconomico*, 27: 99-116.

Murcia, G. (1997). "Convergencia departamental en Colombia", *DNP-UMACRO*.

Nelson, Ch.R and Plosser, Ch.I. (1982). "Trends and Random Walks in Macroeconomics Time Series: Some evidence and Implications". *Journal of Monetary Economics*, 10: 139-162.

Núñez, J.J. (2006). "Inequality measures, Lorenz curves and generating functions". In *Distribution Models Theory* (Herrerías, R.; Callejón, J.; Herrerías, J.M., eds.). World Scientific Publishing Co. Singapore, 189-220.

Paap, R and Van Dijk, H. (1998). "Distribution and mobility of wealth of nations". *European Economic Review*, Vol 42(7): 1269-1293.

Parzen, E. (1962). *Stochastic Processes.* Holden Day.

Pena, J.B. and Núñez, J.J. (2003). "La movilidad en la distribución personal de la renta en España". *Special Issue of the Inter-American Statistical Institute (IASI) Journal of Statistics*, (C. Dagum, ed.). Ch. 2, Sec. I, Vol. 55: 164-165.

Perron, P. (1989). "The Great Crash, the Oil Shock and the Unit Root Hypothesis". *Econometrica*, 57: 1361-1401.

Prais, S.J. (1955). "Measuring social mobility". *Journal of the Royal Statistical Society,* A, Part I, 118: 56-66.

Quah, D. (1993a). "Galton´s Fallacy and Tests of the Convergence Hypothesis". *W.P. Department of Economics L.S.E* (1993); *Scandinavian Journal of Economics*, December, 1993 (Reprinted in T.M. Andersen and K.O. Moene (eds.), *Endogenous Growth*, Blackwell, 1993).
--------------(1993b) "Empirical cross-section dynamics in economic growth". *European Economic Review,* 37: 426–434.
------------- (1996a). "Empirics for Economic Growth and Convergence". *European Economic Review*, 40(6): 1353-1375.
----------------(1996b). "Twin Peaks: Growth and Convergence in Models of Distribution Dynamics". *The Economic Journal*, 106: 1045-1055
-------------- (1997). "Empirics for Growth and Convergence: Stratification, Polarization, and Convergence Clubs". *Journal of Economic Growth*, 2(1): 27-60.
------------- (2001). "Discussion of 'Searching for Prosperity' by Michael Kremer, Alexei Onatski, and James Stock". Available at http://econ.lse.ac.uk/~dquah/.

Quah, D. and Durlauf, S.N. (1998). "The new empirics of economic growth". *CEPR Working Paper,* nº 384.

Rocha, R. and Vivas, A. (1998). "Crecimiento regional en Colombia: ¿persiste la desigualdad?". *Revista de Economía del Rosario*, 1(1): 67-108.

Sala-I-Martin, X. (2000). *Apuntes de crecimiento económico.* Antoni Bosch (editor), segunda edición.
---------------(2002). "The World Distribution of income (estimated from individual country distributions". *NBER, WP.* No. 8933.

Salgado-Ugarte, I.H., M. Shimizu, and T. Taniuchi. (1995). "ASH-WARPing, and kernel density estimation for univariate data". *Stata Technical Bulletin*, 26: 23-31.

Scott, D.W. (1992). *Multivariate Density Estimation: Theory, Practice, and Visualization.* John Wiley.

Shorrocks, A.F. (1976). "Income mobility and the Markov assumption". *Economic Journal*, 86: 566-578.
---------------------(1978). "The measurement of mobility". *Econometrica*, 46(5): 1013-1025.

Silverman, B.W. (1986). *Density Estimation for Statistics and Data Analysis*. Chapman and Hall.

Summers, A and Heston, R (1991). "The Penn World Table (Mark 5): An expanded set of international comparations, 1950-1988". *Quarterly Journal of Economics*, vol. 106 (2): 327, 368. Available at http://www.chass.utoronto.ca/