ECONOMIC CONVERGENCE OF INCOME DISTRIBUTION WORLDWIDE FROM 1986 TO 2000.

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, volume 36 issue 6 (November 2009)

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

- 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.

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

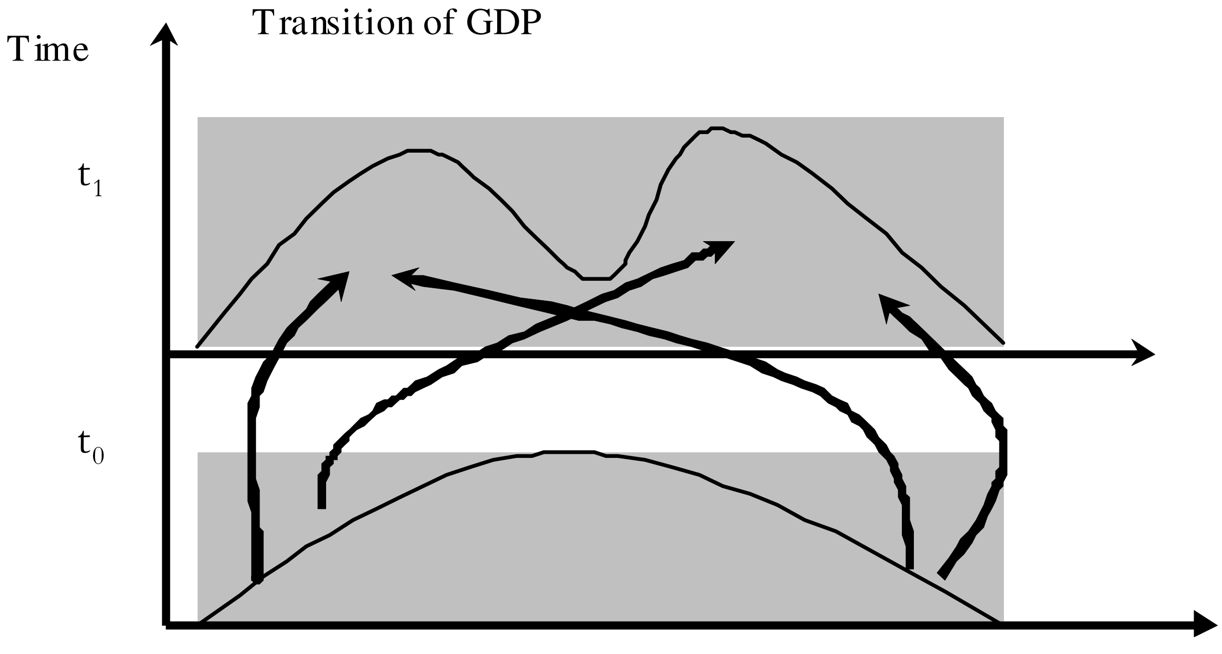
Table 1: Estimates of β-type convergence in different countries

	-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).

2.- Convergence clubs?

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



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

Multimodality test

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 cumulative distribution of the sample. Thus, the probability density function of the data is

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

where w_n(x,u) satisfies $\int_{-\infty}^{+\infty} w_n(x,u).du = 1$. Let $w_n(x,u) = h^{-1} K\left(\frac{x-u}{h}\right), x \in \Re$ where K is some Kernel and h>0 is the bandwidth. The Kernel density estimator of f(x) is (see Silverman, 1986; Härdle, 1991):

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

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).

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.

The bandwidth was calculated using Silverman's (1986) equations 3.30 and 3.31, and so its expression is h = 0.9n-1/5min (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.

3.- Transition of the GDP and use of Markov chains (2)

If λ_t is a measure of income distribution in the countries at time t, then income distribution will evolve as follows,

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

• M represents the intra-distributional dynamics of income in the different countries.

• M contains more information than just consolidated statistics such as means or typical deviations [Quah (1996b: 1370)].

• M incorporates conditional probabilistic information that allows a transition from income 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 process.

If equation (2) is iterated, then

$$\lambda_{t+2} = M\lambda_{t+1} = M\left(M\lambda_t\right) = M^2\lambda_t$$

$$\lambda_{t+3} = M\lambda_{t+2} = M\left(M^2\lambda_t\right) = M^3\lambda_t$$

$$\lambda_{t+4} = M\lambda_{t+3} = M\left(M^3\lambda_t\right) = M^4\lambda_t$$

$$\ldots\ldots\ldots = \ldots\ldots\ldots\ldots\ldots = \ldots\ldots\ldots\ldots\ldots\ldots\ldots = \ldots\ldots\ldots$$

$$\lambda_{t+s} = M\lambda_{t+s-1} = M\left(M^{s-1}\lambda_t\right) = M^s\lambda_t \qquad (3)$$

Now, if one takes the limits when s $\rightarrow \infty$, 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 breaks down the evolution of income into the five following states: 1/4 of the mean income worldwide, 1/2 of the mean income worldwide, the mean income worldwide, twice the mean income worldwide, and more than twice the mean income worldwide.

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 modes

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).

N.T.T= 2714

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.

1. The ergodic distribution resulting from the "discretization" by means of quartiles of salary distributions displays an even behavior.

2. Although the ergodic distribution changes when intermediate information between both periods is not considered, the dynamics of evolution remains unchanged.

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).

N.T.T= 2714

Table 5: RGDPL. Inter-annual transition matrix by quintiles (1962-1984)

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

N.T.T= 2714

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

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:

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).

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).

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

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).

Again if one takes the initial and final years of the sample instead of taking the entire period, then

Table 9: RGDPL, 24-year transition (1986 – 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).

4. -Homogeneity and Markovian nature of income distribution

The selection of the states has a clear impact on the structure of the transition matrix - 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.

But, 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' mobility index, which is defined by

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

n is the number of states, *P* is the corresponding transition matrix, and *tr(P)*, its trace.

M (P) = 0 total immobility
M (P) = 1 total mobility

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

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

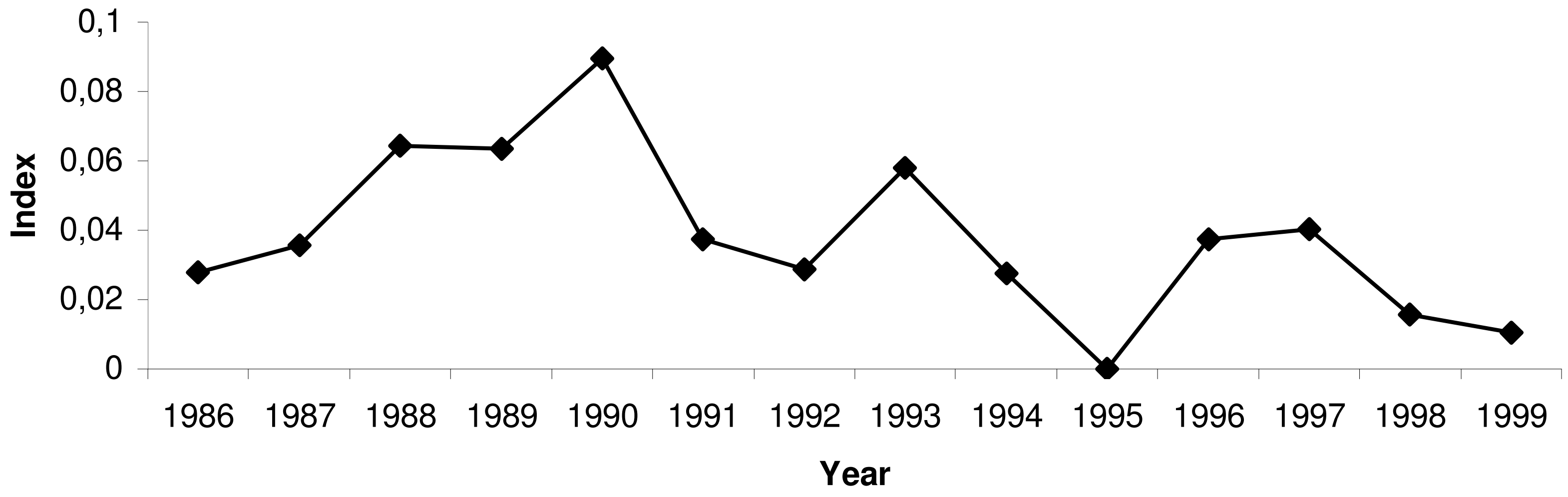
1997-1998	1/4	1/2	1	2	∞
1/4	100	0	0	0	0
1/2	0	100	0	0	0
1	0	4.4	95.7	0	0
2	0	0	11.8	88.24	0
∞	0	0	0	0	100

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).

Figure 2: Shorrocks' mobility indexes from 1986 to 2000



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

This fact strongly supports the assumption that the homogeneity hypothesis of involved Markov chains could fail. 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,

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

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 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.

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.