

# What Have We Not Learned from the Convergence Debate?

Leone Leonida <sup>1,2</sup> 

<sup>1</sup> King's Business School, King's College London, 30 Aldwych, London WC2B 4BG, UK; leone.leonida@kcl.ac.uk; Tel.: +44-(0)207-848-4152

<sup>2</sup> DESMAS, University of Messina, Piazza Pugliatti 1, 98100 Messina, Italy; leonida@unime.it

**Abstract:** We argue that the alternative hypotheses about cross-country convergence dynamics, namely the Conditional Convergence Hypothesis, Absolute Convergence Hypothesis, and Club Convergence hypothesis, build upon different modelling choices and answer different empirical questions. Hence, results favoring one hypothesis are not necessarily evidence against the other hypotheses. We apply several modelling approaches to a sample of world economies to support our argument, and present empirical evidence that yields controversial conclusions if the hypotheses about convergence are taken as competing. However, the controversy disappears as we note that there are neither theoretical nor empirical reasons to take evidence in favor of the Absolute Convergence Hypothesis as necessarily being against the Club Convergence Hypothesis, and vice versa. We present results for the world economies where the two processes co-exist. We conclude by arguing that when analysis is conducted at the regional level, the two processes are more likely to co-exist because regions share the same institutions, culture, natural resources, and other fundamental causes of growth. Consequently, a test for convergence that studies whether clusters converge or diverge once eventually emerged in the distribution of incomes is still needed. Finally, we argue that to study convergence dynamics, it is necessary to model relationships between economies.

**Keywords:** absolute convergence; club convergence; conditional convergence; regional analysis

**MSC:** 91-10; 91-11

**JEL Classification:** C14; O40



**Citation:** Leonida, L. What Have We Not Learned from the Convergence Debate? *Mathematics* **2023**, *11*, 2119. <https://doi.org/10.3390/math11092119>

Academic Editor: Manuel Alberto M. Ferreira

Received: 14 March 2023

Revised: 11 April 2023

Accepted: 25 April 2023

Published: 29 April 2023



**Copyright:** © 2023 by the author. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

## 1. Introduction

Economists and social scientists have long been involved in studying growth processes and whether economies tend towards convergence with one another in the long run—the so-called convergence debate. The debate originated in the seminal research by Baumol (1986) [1] and Abramovitz (1986) [2], and was substantially fuelled by the challenging evidence of Barro and Sala-i-Martin (1991) [3], who claimed that economies converge to each other at a constant speed. Building on this initial body of research, new definitions of convergence have been proposed, and many empirical tools have been employed to study whether economies display a tendency to converge (among many others, see Barro and Sala-i-Martin, 1992 [4] and 1997 [5]; Bernard and Durlauf, 1995 [6]; Ben-David, 1994 [7]; Canova and Marcet, 1995 [8]; Islam, 1995 [9]; Quah, 1997 [10]; Bianchi, 1997 [11]; Caggiano and Leonida 2007 [12], 2008 [13]).

Several reasons justify the attention that researchers have been paying to this question. Whether poor economies are converging to rich economies, how long the convergence process will eventually take to complete, and what determines persistence in the differences between poor and rich economies are questions “of paramount importance for human welfare” (Islam, 2003:309) [14], so much that it is “hard to think of a more fundamental question to answer” (Temple, 1999:112) [15]. Policy implications about the possible solutions to under development and poverty may be drawn from this research. In a sense, government

policy should be driven by answers that economists provide to these questions [15]. Researchers interested in the area believe that “in terms of medium and long-term welfare, the trend is more important than the cycle—provided the volatility of incomes remains as low as it has been during the last few decades” (De la Fuente 2000:25) [16]. Research on convergence is also useful from a theoretical perspective, given there exists an apparent “conflict between growth models that predict convergence, and those models, which do not” (Lamo 2000:681) [17] and that researchers believe that the validity of alternative growth models can be tested by building on their predictions about convergence dynamics [14]. The resulting large amount of literature has led “to many different interpretations of convergence and to a wide array of empirical results, so much so that a feeling of exasperation” (Islam 2003:309–310) [14] exists. Some have expressed the view that the new growth literature in general, and the convergence literature in particular, have not produced anything new or substantive (Parente, 2001) [18]. Some others dispute this argument, noting that not all the evidence on convergence may be interpreted in a neoclassical framework, or by highlighting that the neoclassical model has been changed and adapted to account for evidence coming from the convergence literature (Durlauf and Quah, 1999 [19]; Temple, 1999 [15]; Islam, 2003 [14]).

Islam (2003) [14] asks what we have learnt from the huge efforts that researchers have undertaken in this regard. He suggests that there is some agreement among the results, and that the evidence did help modelling growth theory and the development of new growth models by shedding light on the large technological and institutional differences across economies. Dowrick and DeLong (2003) [20] discuss convergence in a more historical perspective and note that the club convergence among rich economies has been occurring with some economies, which initially take off but lose the opportunity to fully catch up with rich economies. Grinin and Korotayev (2015) [21] frame the discussion about convergence in the more general context of the economic and political relationships between the West and the East. Johnson and Papageorgiu (2020) [22] ask what remains of cross-country convergence and discuss evidence showing that the gap between economies is not closing (see also Caselli et al. 1996) [23].

This paper does not aim at providing a new and comprehensive review of the debate on convergence, as excellent survey papers on convergence and its relationships with the growth empirics already exist (Sala-i-Martin, 1996 [24]; de la Fuente, 1997 [16]; Durlauf and Quah, 1999 [19]; Temple, 1999 [15]; Islam, 2003 [14]; Johnson and Papageorgiu, 2020 [22]). What we aim at instead, is a discussion of some aspects of the relationships between the hypothesis of convergence that have been left unaddressed—or only marginally addressed—by the relevant studies. In turn, our discussion will allow scope for further research by the interested reader.

The above is a sensible aim for at least two reasons. First, even if the debate about cross-country convergence has considerably slowed down, some poor countries have been catching up with the rich economies and, according to Kremer et al. (2021) [25] and Patel et al. (2021) [26], recent evidence is in favour of absolute unconditional convergence. Hence, it is key to clarify what the different empirical tools are testing for, and what the relationships are among the alternative hypotheses about convergence. In this respect, we note that authors involved in the analysis of convergence take as a given the relationships between the different hypotheses about convergence dynamics, namely the Conditional Convergence Hypothesis, the Absolute Convergence Hypothesis (ACH hereinafter), and the Club Convergence Hypothesis (CCH hereinafter). Second, and possibly more importantly, the set of hypotheses, theoretical frameworks and empirical tools that researchers have developed in the cross-country convergence debate have been recently applied to studying convergence at the regional level. Regional convergence has been studied, for example, for the cases of Brazil (Lima 2010) [27], Spain (Gonzalez-Paramo and Martinez-Lopez 2003 [28], Puente 2017 [29]), the EU (Goech and Huter 2016 [30]; Eichengreen, 2019 [31]; Savoia, 2020 [32]), the US (Ram, 2021 [33]), and Russia (Kholodilin, Oshchepkov and Siliverstovs, 2012 [34]; Lehmann, Oshchepkov and

Silvagni, 2020 [35]), among many others. However, none of the authors has asked—to the best of our knowledge—whether and in which direction the empirical frameworks should be adapted when regional studies are of interest. The discussion will highlight some interesting points when convergence is studied at the regional level.

In detail, we contribute to the literature in three respects. The first contribution we make is that we show that the evidence about the different convergence hypotheses offered by the literature should be taken as controversial, despite [14] arguing that the evidence from different empirical frameworks can somewhat be reconciled. Upon closer examination, indeed, the evidence yields different conclusions according to the framework adopted for the empirical analysis. Hence, in Section 2, we describe the data that we adopt to present our arguments, and the rationale for choosing the data span under analysis. In Section 3, we illustrate our argument by adopting some of the empirical tools proposed in the literature on convergence to the sample of world economies described in Section 2—namely, the  $\beta$ -convergence approach, the  $\sigma$ -convergence, and the approach building on the distributional dynamics. In doing so, we briefly discuss the advantages and shortcomings of all these frameworks, paying particular attention to the hypothesis of convergence they are testing for. We note that results arising from alternative empirical approaches are controversial. Some papers explicitly conclude in favour of convergence (Barro and Sala-i-Martin, 1991 [3]; Barro, 1999 [36] and, more recently, Kremer et al. 2021 [25] and Patel et al. 2021 [26]). Some others, however, lead to the opposite conclusion (Ben David, 1994 [7]; Bernard and Durlauf, 1995 [6]; Quah, 1997 [10]; Bianchi, 1997 [11], Caggiano and Leonida, 2008 [19]). This is essentially because of modelling issues behind empirical frameworks that make it difficult to reconcile the conflicting evidence under the same convergence hypothesis. In performing our analysis, we note that the different frameworks test for different hypotheses about convergence, and that this is the main reason for the evidence appearing controversial at first sight. We note that once the set of regressors in a growth determinant equation is augmented to account for steady state determinants, not only can the empirical framework inform only about conditional convergence processes, but also that without the growth determinants in the set of regressors, the empirical model is likely to be mis-specified, and the conclusions about absolute convergence are at risk.

Our second contribution to the literature is that, quite surprisingly, a test for the CCH does not necessarily provide information about the ACH, and that a test for the ACH does not necessarily provide evidence about the CCH. This implies that evidence in favour of the CCH should not be taken as being against the ACH, and vice versa. This argument is presented in Section 4, where our reasoning shows that empirical results are less controversial than they appear—not because the evidence they provide can be somewhat reconciled, but because the alternative empirical approaches underlie different testable hypotheses. Indeed, in empirical analyses, club convergence dynamics and absolute convergence processes are taken as competing hypothesis. However, we show that this is not necessarily the case. Consequently, a cautious approach must be taken when results arising from different approaches are compared, and conclusions on convergence processes are drawn.

Once the relationship between the ACH and the CCH is clarified in Section 5, we discuss some theoretical arguments for the co-existence of CCH and ACH processes, and the potential consequences of testing for convergence when regions, instead of countries, are of interest. It is likely that similar economies, or economies that are geographically closer, converge to each other before eventually converging with those that are at more different stages of economic development or geographically more distant from each other. In turn, this means that studies about regional convergence are more likely to conclude in favour of absolute convergence than those studying the world economies, as the former share institutions, culture, and government policies, among others. This also means that the speed of convergence is likely to be faster for the case of regional studies and overshadow potential absolute convergence dynamics. However, because the relationships among the different hypotheses of convergence are not clear, care must be taken when the set

of empirical tools is adapted to studying samples of economies rather than the world economies. Further, it is necessary to adopt a theory-free empirical framework to test for CCH and ACH. In turn, this makes it difficult to adopt economic theory to explain mobility dynamics and convergence processes. We also note that results on convergence may not help in discriminating between existing growth models because both theoretical results using the neo-classical framework and those from alternative growth models are able to conform to empirical evidence. This is our third contribution to the literature.

Section 6 concludes by noting that a test is needed to study the behaviour over time of the clusters that eventually emerge in the distribution of incomes, and that in testing for convergence, economies are typically taken as independent observations. In turn, this means that interactions among economies are often excluded. This is quite surprising given that interactions among economies, through international trade and migration processes, for example, define winners and losers in growth—and, therefore, whether poor converge to rich countries and, ultimately, whether economies converge to each other. From this perspective, both the theoretical and, especially, the empirical debates largely understate the role that the relationship between economies should play in explaining convergence and growth processes.

## 2. Data

In order to illustrate the arguments, we present results from empirical analyses that are similar in spirit to those offered by the relevant literature. We adopt some of the techniques that have been proposed and test for the different hypotheses on convergence for a sample of 53 countries from 1964 to 1999 (Table 1). The sample period covers a similar period adopted by the bulk of the literature where the debate about convergence has originated. This is because we are not interested in studying whether world economies are currently converging, but in examining results from the different methodologies at work.

**Table 1.** Sample of World Economies.

Num	Code	Country	Num	Code	Country	Num	Code	Country
1	ARG	Argentina	19	GTM	Guatemala	37	NOR	Norway
2	AUS	Austria	20	HKG	Hong Kong	38	NZL	New Zealand
3	AUT	Australia	21	HND	Honduras	39	PAN	Panama
4	BEL	Belgium	22	IND	Indonesia	40	PER	Perù
5	BOL	Bolivia	23	IRL	Ireland	41	PHL	Philippines
6	BWA	Botswana	24	IRN	Iran	42	PRT	Portugal
7	CAN	Canada	25	ISL	Iceland	43	PRY	Paraguay
8	CHE	Chile	26	ISR	Israel	44	SLE	Sierra Leone
9	CHL	Switzerland	27	ITA	Italy	45	SWE	Sweden
10	COL	Colombia	28	JAM	Jamaica	46	SYR	Syria
11	DNK	Denmark	29	JPN	Japan	47	THA	Thailand
12	DOM	Dominican Rep.	30	KEN	Kenya	48	TUR	Turkey
13	ECU	Ecuador	31	KOR	Korea	49	TWN	Taiwan
14	ESP	Spain	32	LKA	Sri Lanka	50	USA	U.S.A.
15	FIN	Finland	33	MEX	Mexico	51	VEN	Venezuela
16	FRA	France	34	MUS	Mauritius	52	ZMB	Zambia
17	GBR	Great Britain	35	MWI	Malawi	53	ZWE	Zimbabwe
18	GRC	Greece	36	NLD	Netherland			

The sample selection criteria have been mainly guided by data availability in the Penn World Table (Aten, Summers and Heston 2001 [37]) and Barro and Lee (2000) [38]. The advantage of the former is that it denominates variables in a common set of prices and in a common currency so that real international comparisons are possible both over time and across countries. The sample results are almost equally distributed across continents. When we observe countries over such a long period of time, the results are likely to be biased towards rich economies, where statistics are likely to be recorded on a more systematic basis

(this is an issue that is known in the PWT). Appendix A reports the correlation analysis, showing that correlations are in line with expectations.

We use the chain index of the per capita GDP, employment and population size, and a measure of the degree of trade openness, defined as the sum of exports plus imports over GDP. We add this variable to the set of regressors in an attempt to control for technological transfers among economies. The human capital stock is measured as the average number of years of schooling, and is available in the Barro and Lee (2000) [38] dataset. Because investment is available in PWT 6.1, the producer capital stock per worker is obtained—up to 1992—from PWT 5.6 by Summer and Heston (1996) [39]. In order to lengthen the series until 1998, we apply a permanent inventory equation using the investment series in PWT 6.1. Growth rates are calculated from stocks.

### 3. Conflicting Views

#### 3.1. The Analysis of the Mean: $\beta$ -Convergence

The analysis of convergence can be framed into the neoclassical optimal growth model where, once the factor accumulation is taken into account, countries converge to their long-run equilibrium, i.e., the so-called steady state. In this class of models, the representative economy is typically assumed to have non-increasing returns to scale; in turn, this assumption ensures the existence of the steady state. According to Barro and Sala-i-Martin (1991) [3] and (1992) [4], the theoretical model implies that, other things being equal, poor countries grow faster than rich economies. Bairam and Mc Rae (1999) [40] propose a more flexible test of this assumption, which allows the empirical model for non-increasing returns to scale. Let:

$$Y_{it} = A_{it}K_{it}^{\alpha}H_{it}^{\beta}L_{it}^{\lambda}, \tag{1}$$

be the production function where in country  $i$  at time  $t$ ,  $Y$ ,  $A$ ,  $K$ ,  $H$ , and  $L$  are aggregate output, technology, physical capital stock, human capital stock, and labour, respectively; and  $\alpha$ ,  $\beta$ , and  $\lambda$  are factors elasticity. Dividing both sides of (1) by  $L$ , taking natural logs and differentiating it with respect to time yields:

$$g_{YLi} = g_{ai} + \alpha g_{KLi} + \beta g_{HLi} + \varphi g_{Li}, \tag{2}$$

where  $g$  is the growth rate of the relevant variable. In Equation (2), the parameter  $\varphi = (\alpha + \beta + \lambda - 1)$  helps to test for the degree of returns to scale, where  $\varphi <, >$  or  $=0$  implies decreasing, increasing, or constant returns to scale, respectively. The term  $g_{ai}$  is the rate of technological change. Ref. [40] suggest testing for two hypotheses:

$$g_{ai} = \gamma_0 + \gamma_1(Y_{it_0}/L_{it_0}), \tag{3}$$

and

$$g_{ai} = \gamma_0 + \gamma_1(Y_{it_0}/L_{it_0}) + \gamma_2(Y_{it_0}/L_{it_0})^2, \tag{4}$$

where  $Y_{it_0}/L_{it_0}$  is the labour productivity level in the initial year. Equations (3) and (4) assume that growth rates differ among countries because of their technological gap; these parameters are usually employed to test for linear and non-linear convergence. For the sake of our discussion, it is also useful to test for the hypothesis that the more an economy interacts in terms of flow of goods with other economies, the higher the growth rate because technology is embedded into exchanged goods. Hence, we also assume:

$$g_{ai} = \gamma_0 + \gamma_1(Y_{it_0}/L_{it_0}) + \gamma_2(Y_{it_0}/L_{it_0})^2 + \gamma_3(OPEN/Y_i), \tag{5}$$

where Equation (5) includes a measure of openness to foreign trade as an explaining factor of technological change as a fraction of the total output, i.e.,  $OPEN$  is the sum of export and import.

The discussion above leads to five estimating models:

$$g_{YLi} = \gamma_0 + \gamma_1(Y_{it_0}/L_{it_0}) + \varepsilon_i, \tag{6}$$

$$g_{YLi} = \gamma_0 + \gamma_1(Y_{it_0}/L_{it_0}) + \gamma_2(Y_{it_0}/L_{it_0})^2 + \varepsilon_i, \tag{7}$$

$$g_{YLi} = \gamma_0 + \gamma_1(Y_{it_0}/L_{it_0}) + \gamma_2(Y_{it_0}/L_{it_0})^2 + \gamma_1(OPEN/Y_i) + \varepsilon_i, \tag{8}$$

$$g_{YLi} = \gamma_0 + \gamma_1(Y_{it_0}/L_{it_0}) + \gamma_2(Y_{it_0}/L_{it_0})^2 + \gamma_1(OPEN/Y_i) + \alpha g_{KLi} + \beta g_{HLi} + \varepsilon_i, \tag{9}$$

$$g_{YLi} = \gamma_0 + \gamma_1(Y_{it_0}/L_{it_0}) + \gamma_2(Y_{it_0}/L_{it_0})^2 + \gamma_1(OPEN/Y_i) + \alpha g_{KLi} + \beta g_{HLi} + \varphi g_{Li} + \varepsilon_i, \tag{10}$$

The empirical strategy consists of allowing for the maximum flexibility of the empirical framework to test for: the absolute linear and non-linear convergence using Equations (6) and (7); linear and non-linear convergence conditional to the openness variable using Equation (8); linear and non-linear conditional convergence under constant returns to scale using Equation (9); and, finally, the regime of returns to scale with Equation (10). Initially, the studies on growth determinants were conducted in the context of an OLS framework, using cross-sectional averages for time-series data (Mankiw, D. Romer and Weil, 1992 [41], Barro and Sala-i-Martin, 1997 [5], Barro, 1999 [36]).

However, according to [5], if country and/or time effects that control the estimate for unobserved fixed and time effects do exist, the error term should be written as:

$$\varepsilon_{it} = \mu_i + \eta_t + v_{it}, \tag{11}$$

where  $\mu_i$  denotes the country individual effect, which is constant through time and varies with the cross-sectional units;  $\eta_t$  is the time effect, which is constant across units and varies with different time periods; and  $v_{it}$  is the random error term. Because the fixed effects  $\mu_i$  and  $\eta_t$  are proxies of the country- and time-specific technological change, they may affect the factor accumulation process as well as the output growth process. If we account for this decomposition, Equation (10), for example, would be written as:

$$g_{YLi} = \gamma_0 + \gamma_1(Y_{it_0}/L_{it_0}) + \gamma_2(Y_{it_0}/L_{it_0})^2 + \gamma_1(OPEN/Y_{it}) + \alpha g_{KLit} + \beta g_{HLit} + \varphi g_{Lit} + \mu_i + \eta_t + v_{it} \tag{12}$$

where  $t$  are typically averages of 5 years and  $t_0$  is now the initial observation of each of these five-year intervals. The term  $\mu_i$  controls the estimated parameters for unknown country heterogeneity, such as different production functions, and  $\eta_t$  controls for unknown exogenous shocks that affect all economies equally. In this case, the estimation framework would be the within-groups estimator (WG).

The resulting framework addresses a number of interesting questions, especially in light of the discussion we want to undertake. First, because the technological progress is modelled as a function of the initial level of per worker GDP, we test the Absolute Convergence Hypothesis, which is assumed to be both a linear and a non-linear process. Second, we can test the degree of the returns of scale directly, without imposing any restriction on the estimates and by controlling for unobserved time and country heterogeneity. Third, we can investigate the hypothesis that international trade provides access to imported inputs, which embody new technology and increase the effective size of the market facing producers, which raises the returns to innovation.

Table 2 reports empirical results, building on seven 5-year intervals, as is typical in such regression approaches. We estimate models (6)–(10) by both the OLS and the WG framework. The results reported in Panel A show that, when estimated by the OLS,  $\gamma_1$  has a negative sign and is statistically significant at the 10% s.l. Hence, the results about the initial level of GDP suggest that economies with a lower initial GDP grow, on average, faster than economies with higher initial GDP. The evidence in Column (a), therefore, supports

the ACH hypothesis. We note that the coefficient is smaller than the typical 2% speed; we believe that this is due to the particular sample we are observing, which is biased towards rich economies that are closer with each other and produces a lower average convergence speed. The evidence in Column (b) rejects the hypothesis that convergence takes place at a decreasing rate as  $\gamma_1$  and  $\gamma_2$  have the correct sign but are not statistically significant.

When the set of regressors is expanded to include the variable OPEN, the corresponding parameter is statistically significant, with the expected sign and an estimated elasticity coefficient of about 1.4%. This result supports the hypothesis that international trade helps disseminate technological progress across world economies. Physical capital accumulation accounts for about 55% of the total variation in labour productivity; the human capital estimated coefficient has the expected sign but is not robustly correlated to growth. Evidence in favour of the constant returns to scale hypothesis emerges, as  $\varphi$  is not statistically significant.

In some of the models estimated by the OLS, the constant term is not statistically significant. This suggests that there is some unexplained heterogeneity that the single pooled intercept does not capture. This suggests, in turn, that economies may not share the same production function, and calls for the full WG approach. The results obtained with the WG estimator are reported in Panel (B). The results reported in Columns (f) and (g) suggest that  $\gamma_1$  is negative and  $\gamma_2$  is positive. The evidence is again in favour of the ACH, as it is in favour of absolute convergence at a decelerating rate.

The estimates of the parameters regarding OPEN are always positive and statistically significant. The impact of this variable is much higher when estimated via the WG than via the OLS. In this case, convergence to the steady state is found to be non-linear and dependent on how much the average economy trades with the rest of the world. The elasticities of output growth to inputs are sensitive to change compared with the OLS estimates: physical capital elasticity reduces to about 20%, and labour input also exhibits a lower elasticity (23%). Human capital is positively and robustly correlated to growth; its elasticity is estimated to be about 16%. Hence, the evidence suggests that returns to scale are decreasing: if all factors of production increase by  $\lambda$ , output grows on average by  $0.59\lambda$ .

Are the results in favour of the ACH? In model (6), when estimated via OLS, the coefficient associated with the gap variable displays the expected sign. When estimated by panel data, the results are statistically significant, and a negative relationship between the rate of growth and the initial level of productivity appears. If anything, the evidence reported should be interpreted as support for the ACH. Once we allow economies to have different intercepts, we allow a certain degree of heterogeneity of the production functions. In this case, the question arises whether the WG estimator is a framework that allows for an analysis of the Absolute Convergence Hypothesis at all. This is because, if the set of regressors is expanded to include growth determinants, then the results should not be taken as evidence in favour of or against the ACH, as empirical results would be conditional on the growth determinants. In this case, the evidence is in favour of the Conditional Convergence Hypothesis. However, if these growth determinants are taken out of the estimating model, then although the results assess the ACH, the empirical model is likely mis-specified because of omitted variable bias.

To summarise, the results are (mildly) in favour of the ACH and also in favour of the Conditional Convergence Hypothesis. This framework, as it is typically adopted, cannot test for the CCH unless further adjustments are made to the set of regressors.

**Table 2.** Empirical results from estimating growth equations.

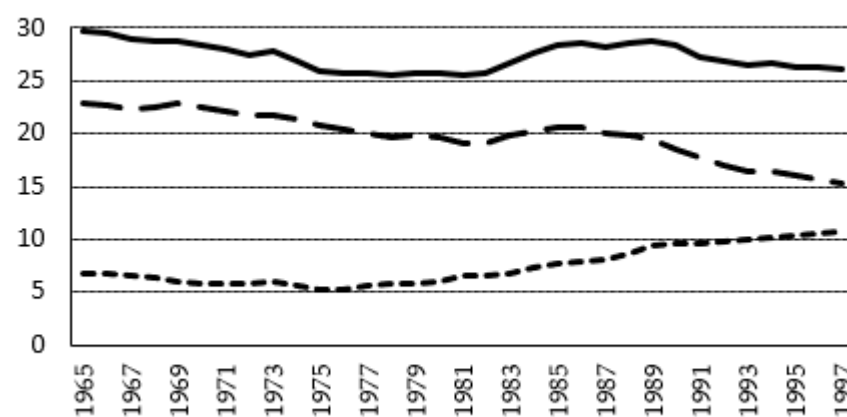
	Panel A: OLS					Panel B: Within Groups				
	(a)	(b)	(c)	(d)	(e)	(f)	(g)	(h)	(i)	(j)
Const.	0.025 (5.16) ***	0.023 (2.98) ***	0.013 (1.82) *	−0.008 (0.83)	0.003 (0.27)					
YL <sub>to</sub>	−0.0004 (−1.75) *	0.0001 (0.15)	0.0001 (0.11)	0.0003 (0.44)	−0.0001 (−0.07)	−0.001 (3.86) ***	−0.003 (3.09) ***	−0.004 (4.11) ***	−0.003 (3.04) ***	−0.004 (3.70) ***
(YL <sub>to</sub> ) <sup>2</sup>		−1.4 × 10 <sup>−8</sup> (0.66)	−1.2 × 10 <sup>−8</sup> (0.59)	−1.2 × 10 <sup>−8</sup> (0.76)	−0.6 × 10 <sup>−8</sup> (0.32)		3.4 × 10 <sup>−8</sup> (2.05) **	4.1 × 10 <sup>−8</sup> (2.55) ***	2.8 × 10 <sup>−8</sup> (1.79) **	3.4 × 10 <sup>−8</sup> (2.30) ***
OPEN/Y			0.016 (2.77) ***	0.014 (3.60) ***	0.013 (3.05) ***			0.046 (3.05) ***	0.020 (1.75) **	0.024 (1.90) **
g <sub>KL</sub>				0.55 (6.64) ***	0.49 (5.06) ***				0.260 (3.92) ***	0.20 (3.22) ***
g <sub>HL</sub>				0.07 (0.34)	0.12 (0.60)				0.24 (3.09) ***	0.16 (2.09) **
g <sub>L</sub>					−0.33 (1.57)					−0.41 (3.46) ***
R <sub>tS</sub>				1.00	0.67				1	0.59
R <sup>2</sup>	0.05	0.05	0.13	0.64	0.66	0.34	0.35	0.38	0.44	0.45

Note: The dependent variable is the average growth rate of the labour productivity—g<sub>YL</sub>. All regression models build upon 53 countries observed over 7 5-year intervals, for a total of 371 observations. Figures in brackets are *t*-statistics. Standard errors are robust to heteroskedasticity and autocorrelation, and clustered at the country level. \*\*\* (\*\*) [\*] means that the estimated coefficient is significant at 1% (5%) [10%]. *R<sub>tS</sub>* is the sum of the estimated elasticities of the factors of production.



### 3.2. The Analysis of Variance: $\sigma$ -Convergence

Researchers recognise that the study of the mean of a sample may not say much about the dynamics of the individual countries in terms of the distribution of incomes. New definitions of convergence have been proposed, and many mathematical and statistical tools have been employed for studying whether economies display a tendency to converge. The underlying idea of all these approaches is that whether countries converge is an empirical issue. The  $\beta$ -convergence framework has been integrated by means of a  $\sigma$ -convergence analysis. Barro and Sala-i-Martin (1997) [5] propose to determine, after having found that conditional convergence holds, whether the dispersion of the real income per capita across groups of economies tends to fall over time. This concept of convergence involves the estimation of the cross-sectional dispersion: a decreasing (increasing) variance is evidence in favour (against) of the Absolute Convergence Hypothesis. Figure 1 reports an estimate of the variance for the sample under analysis.



**Figure 1.**  $\sigma$ -Convergence Analysis. The full line represents the evolution of the variance of GDP per capita among a sample of 53 economies. The dashed and the dotted lines are the variance between and within groups. Groups are defined on the basis of their GDP per capita in 1965 (lower or higher than the mean).

Following this definition of convergence, we find convergence until 1977, and divergence until 1989. The ANOVA analysis in Table 3, that holds having decided a priori how to group economies as rich and poor based on whether in 1965, country  $i$  has a per capita GDP higher or lower than the mean, is reported in Table 4. The distance between the two groups decreases over time. The  $F$  test rejects the null of equality of the means of the groups. Is the evidence in favour of the Absolute Convergence Hypothesis? Again, it seems to be the case for the period under analysis. However, the results crucially depend on how we chose the winners in 1965; to put it differently, the analysis depends on how the groups are defined. Should they be the rich economies in 1965, as in our example, or those that are rich in 1998? This decision may change the conclusion about whether the ACH holds (Ben David 1994 [7]).

Furthermore, the  $\sigma$ -convergence approach may not provide sufficient information to test the ACH. Even if the variance is unchanged through time—showing no convergence or divergence—the economies underlying the cross-section could still be changing their position within the invariant distribution (Quah 1996 [42]).

Panel A of Figure 2 illustrates the difficulty when testing for the ACH by adopting an approach based upon the variance. The countries are represented on the vertical axis, and a group of them follows the trend pattern marked as “1”; a second group follows, instead, the pattern marked “2”. A third group stays in the same position in this distribution over time. Some observations increase their position, while others lose position in the distribution of incomes. In this case, the variance of the distribution is constant, but we would not see the intra-distributional movements of the two groups of countries. Some of

them are catching up with the rich economies, and some of them are moving towards the poor tail of the distribution, but the variance does not tell us much about these dynamics. Second, and more importantly, we would not uncover the difference from this scenario and that represented in Panel B, in which nothing changes, and all the economies stay in the same position.

**Table 3.** ANOVA (selected years).

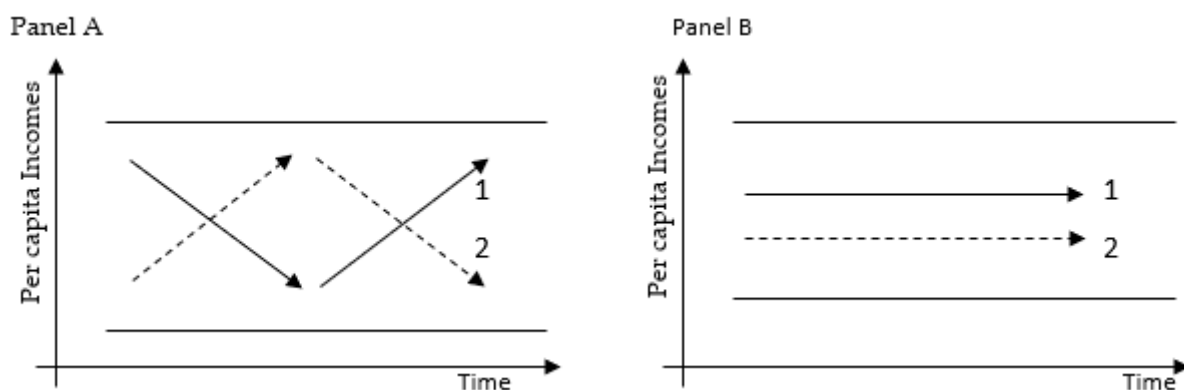
Year	Var	Bet.	Whi.	F-Stat	G <sub>1</sub>	G <sub>2</sub>	Var G <sub>1</sub>	Var G <sub>2</sub>	Dist.
1965	29.61	22.94	6.67	175.28	1.81	0.47	0.24	0.06	1.35
1974	26.93	21.26	5.67	191.20	1.78	0.49	0.17	0.08	1.29
1975	25.95	20.80	5.15	205.92	1.77	0.49	0.14	0.08	1.28
1976	25.69	20.38	5.31	195.69	1.77	0.50	0.14	0.08	1.27
1998	26.25	15.77	10.48	76.73	1.67	0.56	0.17	0.23	1.12

Note: Var. is the variance across the sample. Bet. and Whi. are the variance between and within the two groups, respectively. F-stat tests the difference in mean between the two groups, respectively G<sub>1</sub> and G<sub>2</sub> are the mean incomes of the two groups, and Var G<sub>1</sub> and Var G<sub>2</sub> are the variances inside each group. Dist. represents the distance between the means of the two groups. For groups definition, see Figure 1.

**Table 4.** The Transitional Matrix approach.

Panel A						Panel B					
# Obs.	Upper Endpoint					# Obs.	Upper Endpoint				
	0.25	0.50	1.00	2.00	>2		0.25	0.50	1.00	2.00	>2
212	0.97	0.03	0.00	0.00	0	7	0.57	0.29	0.14	0.00	0
445	0.02	0.96	0.02	0.00	0	14	0.29	0.57	0.00	0.14	0
324	0	0.03	0.93	0.03	0	11	0	0.27	0.36	0.36	0
609	0	0.00	0.01	0.97	0	14	0	0.00	0.14	0.79	0.07
159	0	0.00	0.00	0.11	0.89	7	0	0.00	0.00	0.86	0.14

Note: Panel A and B report transitions of one year (t to t + 1) and 34-year transitions (t to t + 33), respectively, for the sample of world economies described in Table 1.



**Figure 2.**  $\sigma$ -Convergence and club convergence. Panel A reports hypothetical dynamics inside the distribution with a constant variance. Panel B reports hypothetical dynamics where all economies are in their steady state position.

### 3.3. The Analysis of the Distribution of Incomes

Hence, the  $\sigma$ -convergence approach fails to explain the intra-distribution dynamics, and this is may be the most important feature if we want to analyse persistence of income disparities over time. To put it differently, the  $\sigma$ -convergence analysis alone is not sufficient to study convergence unless more information is gained on how units move within the distribution. A law of motion for the cross-section distribution as the realisation of a random element in the space of the distribution is needed. To apply a transitional matrix

framework to the data, we let  $F_t$  denote the distribution of incomes across countries at time  $t$ . We describe  $\{F_t: \text{integer } t\}$ 's evolution using the following law of motion:

$$F_{t+1} = M * F_t, \quad (13)$$

The matrix  $M$  maps the distribution at  $F_t$  into another distribution  $F_{t+1}$ , where in  $F_{t+1}$ , observations at  $F_t$  finish with a lag of 1. In other words,  $M$  encodes information on whether economies such as Korea and the Philippines, for example, which were close together at the beginning of the period, transit subsequently to widely different income levels. This is a standard vector autoregression equation, apart from the fact that its values are a distribution, and the innovation term is absorbed in the  $M$  operator. If, for simplicity, we assume that it is a first order law of motion, by iterating:

$$F_{t+s} = M^s * F_t, \quad (14)$$

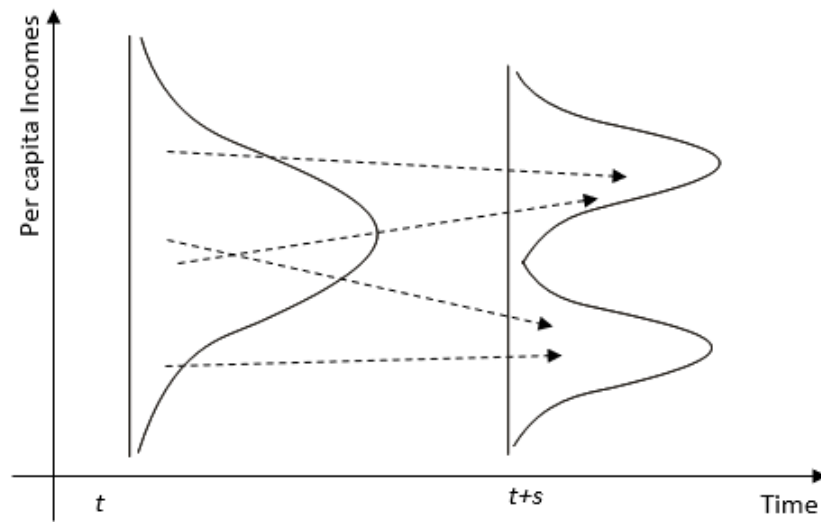
and, taking the limit for  $s$  going to infinity, we obtain a characterisation of the long-run distribution across countries. Convergence might manifest if  $F_{t+s}$  tends to a point mass; convergence would not manifest if  $F_{t+s}$  tends to a bimodal distribution of incomes, with poor and rich economies.

What are the long-run tendencies of incomes across countries? To answer this question, Quah suggests to discretise the set of possible values into the intervals of the distribution. Table 4 reports two transitional matrices. Panel (A) in Table 4 reports the estimates for a one-year transition.

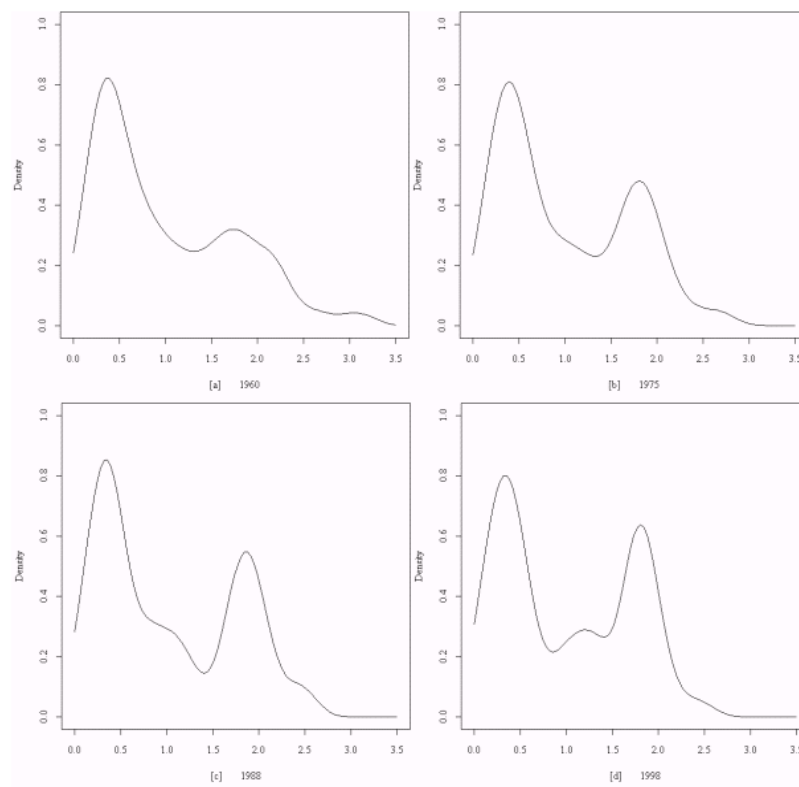
All the relevant properties of  $M$  are described by a  $5 \times 5$  Markov chain transition matrix whose  $(j,k)$  entry is the probability that an economy in state  $j$  transits to state  $k$  (Quah 1993) [43]. The one-step annual transition matrix is estimated by averaging the observed one-year transitions over every year from 1965–66 to 1997–98. The first column provides the total number of transitions with starting points in that income space. For example, the second row shows that 445 observations fell in state 2 across the entire sample (53 countries and 34 years). Of these, 96% remained at the same state. In this one-year period, the predominant feature is, somewhat obviously, high persistence. All the estimated probabilities are higher than 92% along the main diagonal. From such a state, an economy has the same probability of going ahead or behind in the distribution. Notice, however, that the middle class has a higher probability of moving. The second panel describes a 34-year transition. Although persistence is less pronounced, it is, nonetheless, still the main feature of the estimate. Here, clearly, the middle class tends to vanish, and the observations tend to accumulate at the tail of the distribution. Moreover, notice that a representative economy of state 2 is marginally more likely to fall behind than to take off. The contrary is true for an economy starting in state 4. The economies are converging to two different points, and creating two clusters, with a vanishing middle class.

Quah (1993) [43] further suggests estimating the shape of the distribution directly so that results do not depend on groupings. By estimating the shape of the involved densities, it would be straightforward to assess the underlying dynamic and whether a conditional or a club conditional dynamic exists. Figure 3 represents an income distribution at time  $t$  and another (possible) distribution at time  $t + s$ .

If the distribution collapses from a unimodal to a bimodal one (“emerging twin peaks”), intra-convergence is found inside groups of GDP per-capita (clustering), but there is divergence from the other group. However, if the distribution collapses from a bimodal to a unimodal shape, economies are converging over time. In other words, analysing the overall distribution of the income per-capita across countries, regions, provinces, and so on enables the study of both shape and mobility dynamics simultaneously. Formulating the problem of economic growth in the form of Figure 4 draws an equivalence between the analysis of growth and of distributions. It is not that higher growth can cause or, alternatively, be driven by greater inequality, but rather that the two are considered simultaneously [10].



**Figure 3.** The Distributional Approach to Convergence. The vanishing middle class suggests that economies are approaching different long-run equilibria.



**Figure 4.** Kernel Density Estimates and Convergence Testing. Density estimates of GDP per capita. The estimates refers to 1960, 1975, 1988 and 1990, respectively. Densities are obtained by means of a Gaussian Kernel and a Least squares cross-validation bandwidth. The sample size is 53.

Estimates of the distribution may be obtained via a non-parametric approach. There exist various methods in the non-parametric context to uncover how many modes (and so how many clusters) the distribution shows. The kernel density estimate is as follows:

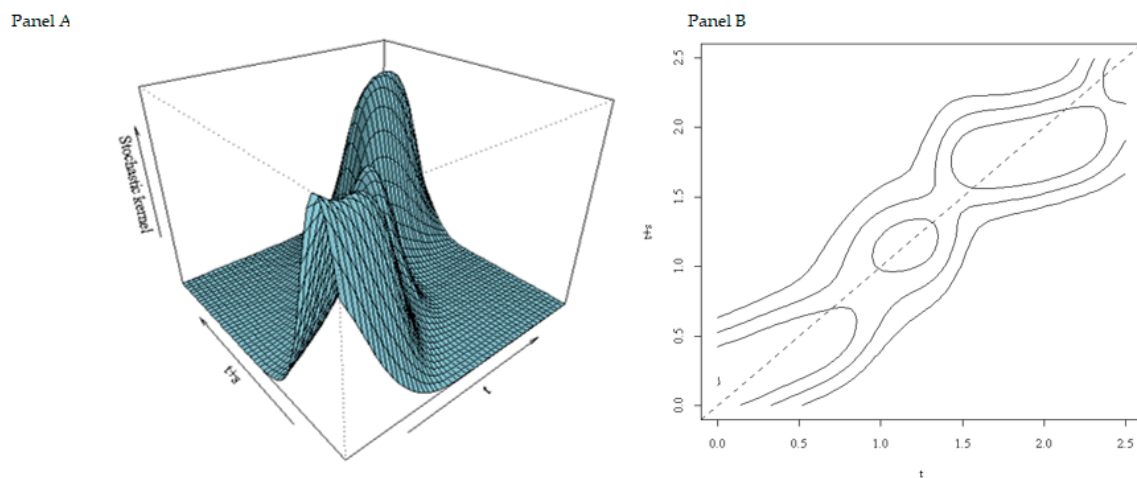
$$\hat{f}_{n,h} = \frac{1}{nh} \sum_{i=1}^N K\left(\frac{x - x_i}{h}\right), \tag{15}$$

where  $h > 0$  is the bandwidth and  $K(\cdot)$  is the kernel, which is an estimate of the density function of the random variables  $x$ , estimated using  $N$  cross-sectional realisations of GDP per capita. The density for  $x$  is estimated at a point  $x = x_0$  as a weighted sum of all the observations by  $1/n$ ; the weight for the  $i$ th observation  $x_i$  is given by the height of the function  $K(\cdot)$  evaluated at  $x_i - x_0$ . This weighting function is the kernel. The magnitude of  $h$  determines which observation we are looking at: if the chosen value of  $h$  is too small, the kernel assigns non-negligible weight only to the observation very close to  $x_0$ , with the result that the density function is insufficiently smoothed and uninformative. However, if the chosen value of  $h$  is too large, the kernel assigns a non-negligible weight even to observations far from  $x_0$ , leading to over-smoothing of the estimated density function and the loss of crucial information about the true shape of the distribution. For this reason, the bandwidth magnitude is the crucial issue for the effective estimation of the density: under the Gaussian kernel, the number of modes is a decreasing function of the bandwidth magnitude. Figure 4 is obtained following exactly this approach.

In the first panel, the second mode is beginning to be visible. In the second panel, the mode is more pronounced. Such a tendency seems to be monotone: the data show no reversals in the dynamic we are describing. The data show that there is no evidence in favour of absolute convergence among countries. The bimodal distribution is taken as proof of a non-convergence dynamic and evidence against the ACH. The second step consists of estimating stochastic kernels to study the mobility dynamics in the distribution. The stochastic kernel estimates the ex-post probability, conditional to the initial position, of the observations changing or staying in the same position. This allows us to examine whether rich countries at  $t$  are still the rich at  $t + s$  (persistence), i.e., if some poor countries at  $t + s$  began as rich (churning or mobility) and/or if some groups of these economies that were originally close together in the middle class separated because of a process of divergence (separability). To observe such dynamic using Bayes law, a stochastic kernel can be defined as:

$$\hat{f}(y_{t+s}|y_t) = \frac{\hat{f}(y_{t+s}, y_t)}{\hat{f}(y_t)} \tag{16}$$

The stochastic kernel (the joint distribution normalised by the implied marginal) also represents a continuous transition matrix and informs about both the modes and the dynamic of the clusters with respect to each other over the years, conditional to the GDP level from which the observations begin. Vectors  $t$  and  $t + s$  are obtained as in the transitional matrix approach. This estimate is reported in Figure 5 for a 15-year transition.

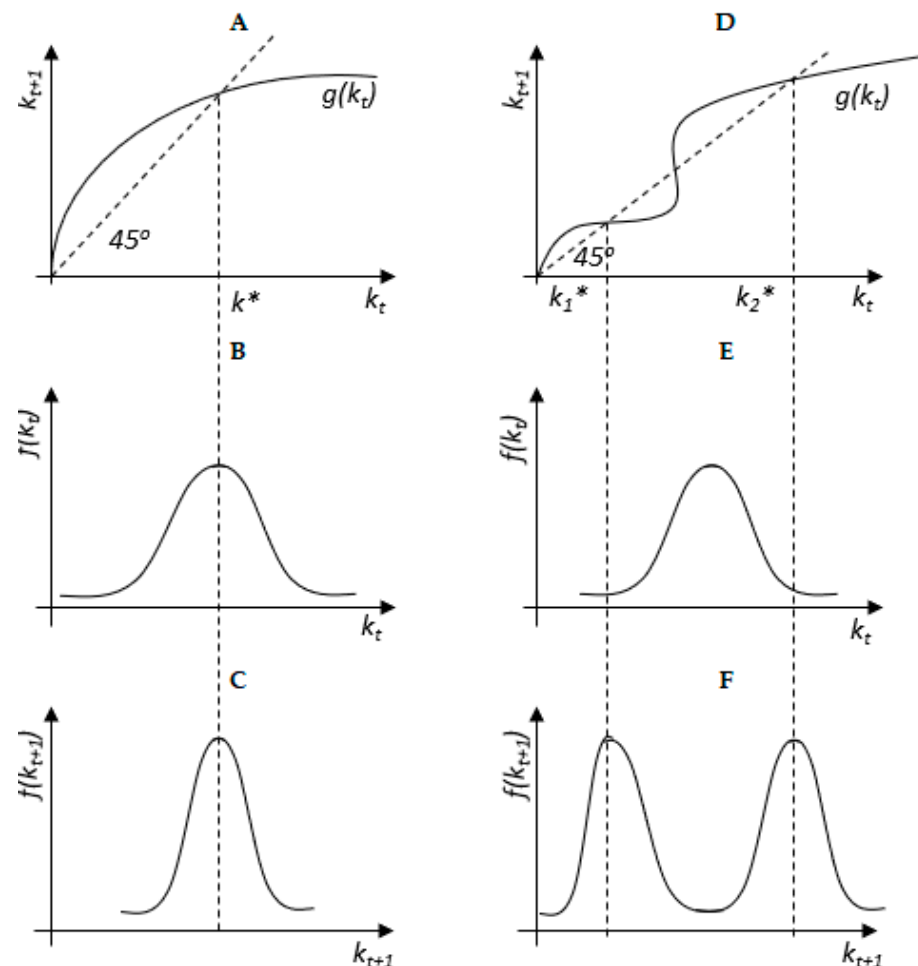


**Figure 5.** Stochastic Kernel and Convergence. Panel A shows the stochastic kernel estimated between  $t$  and  $t + 20$  for the sample of economies, by means of a Gaussian kernel and the Least-Squares Cross-Validation bandwidth. The sample size is 742. Panel B shows the contour plot of the estimate at 0.4, 0.6 and 0.8.

#### 4. Conflicting Views? Testing for Convergence and the Shape of the Distribution of Incomes

Ref. [10] proposes studying the shape of the distribution over time to test for convergence dynamics. If the distribution collapses to a unimodal shape, we have evidence in favour of the ACH and against the CCH. Conversely, if more than one peak emerges in the distribution, countries are catching up with one another, but only within particular subgroups; this evidence would be taken in favour of the CCH, and against the ACH.

The testing approach can be framed in the context of set of assumptions that underlie growth models. Panel A of Figure 5 reports the capital accumulation path of the economy under the assumption that the production function  $g(k)$ , where  $k$  is the capital per employee ratio, has diminishing returns to capital and that returns to scale are constant. Under the additional condition that all the economies share this concave technology because, for example, international trade disseminates it, there is one equilibrium at  $k^*$ . This equilibrium is also stable as the capital accumulation path crosses the 45-degree line from above. In this case, both if the distribution of incomes is unimodal (as in Panel B) or bimodal, then economies will converge in the long run to the unique equilibrium. Because the equilibrium is also stable, we must observe a reduction in the variance as time passes and technological progress spills over across countries, as plotted in Panel C of Figure 6.



**Figure 6.** Transition to a single and to multiple equilibria. Panel A shows the accumulation pattern in the presence of constant returns to scale and diminishing returns to capital, that leads to the reduction in the variance described from Panel B to Panel C. Panel D shows the accumulation pattern in the presence of non-convexities in the accumulation function, that leads that leads to the clustering and the emergence of the twin peaks dynamics described from Panel E to Panel F.

If instead, the production function for the representative economy has some non-convexities, as reported in the example in Panel D, then there are three equilibria, as the accumulation function crosses the 45-degree line three times. In this case, the equilibrium at the centre of the distribution is unstable, and the equilibria located at the tails are stable instead. Hence, at time  $t$ , even though the economies are clustered about the same equilibrium, as in Panel E, in the long run, the economies tend to catch up with each other only within subgroups, and more than one peak emerges, as from Panel E to F. The presence of non-convexities in this case leads to a multiple equilibria configuration (Caggiano and Leonida 2013 [44]).

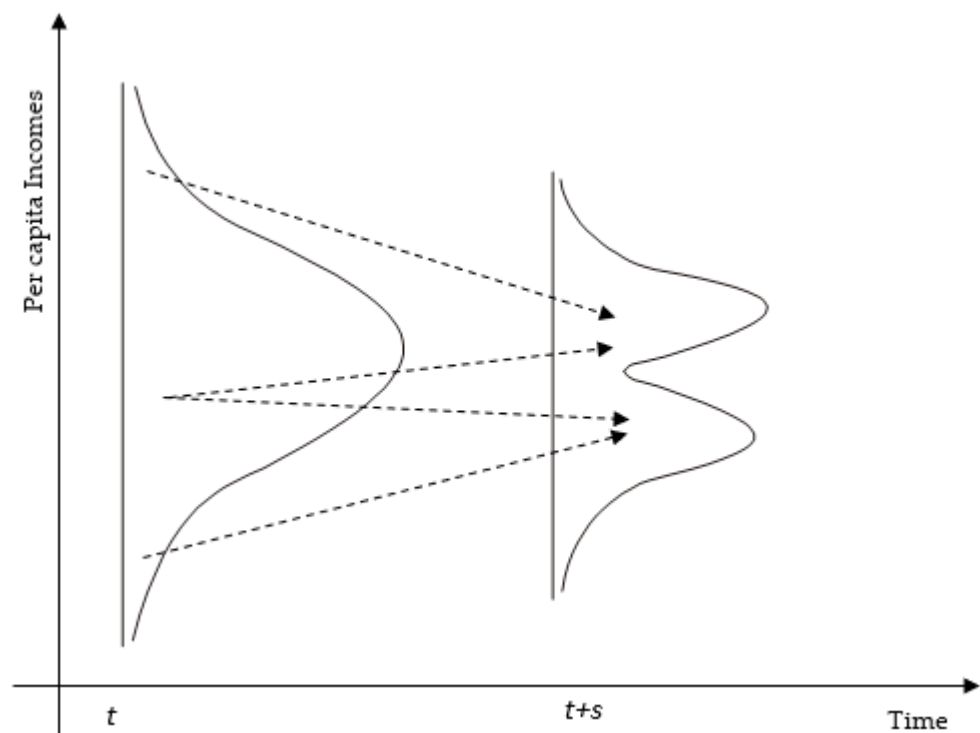
The question to answer is, therefore, whether evidence in favour of a bimodal distribution of incomes would necessarily exclude the absolute convergence dynamics. Quite surprisingly, under closer examination, the answer is no. We argue that the presence of a certain number of groups within the distribution of the per capita GDP does not per se provide evidence against the ACH. Bimodality informs on the number of groups present in the distribution of incomes, but groups may converge over time. It is actually reasonable to expect that similar economies, or economies that are geographically closer, converge more rapidly with each other than with those that are at different levels of development or are geographically more distant. In this latter case, we can observe both clustering processes and convergence dynamics.

A club convergence pattern may well be consistent with both absolute divergence and convergence dynamics (Caggiano and Leonida, 2013 [44]). To illustrate this potential dynamic, Figure 7 shows a hypothetical succession of distributions of incomes. The underlying dynamics are obviously a club convergence process. Indeed, the middle class vanishes and the economies are approaching different equilibria. Two groups of economies arise in the distribution of incomes. However, observations located in the upper part of the distribution converge to the centre of the distribution, and part of the vanishing middle class grows and converges to the rich group of economies. Moreover, observations located in the lower tail of the distribution are growing and approaching the fraction of the middle class that loses position. At the end of this process, the tails of the distribution are less fat. We observe an emerging twin peaks dynamic, with a falling variance also. If the observations located on the tails converge to the centre of the distribution, and this process occurs while the middle class vanishes, then the club convergence combines with absolute convergence. The two processes co-exist. The evidence does not necessarily lead to rejection of the ACH. In this case, we observe a process that may be called clustering convergence. Therefore, a vanishing middle class is a sufficient condition for polarization; however, it is also a sufficient condition for rejection of the ACH, provided clustering is not offset by a mean-reverting movement of economies located on the tails of the distribution.

Another point is that the probability mass shifting from one mode to another may reduce the dispersion of each group, providing evidence in favour of clustering, even if some poor economies actually catch up with observations for the rich ones. If we assume that over time, a set of observations coming from the poor part of the poor cluster of economies catches up with the rich cluster of economies and that these observations move to about the average of the richer cluster, the rich cluster now comprises a higher number of observations with a per capita income which is lower than the mode representing this group. For this reason, the mode representing this cluster is more pronounced. At the same time, the mode representing the poor cluster is less pronounced and it shifts to the right. The dynamics of such a group of countries, therefore, lead to the rejection of unimodality with less strength than earlier. This is the correct inference to make: because a catch-up process occurred, some poor economies catch up with the richer economies.

If we assume instead that they converge with the richer part of the rich cluster, for example, following what it is called a “growth miracle”, a larger number of observations with a per capita income higher than the mode of this cluster compose the rich cluster after the catching-up process. The resulting mode is again more pronounced, and it shifts to the right. The mode representing the poor group shifts to the left. The two modes appear

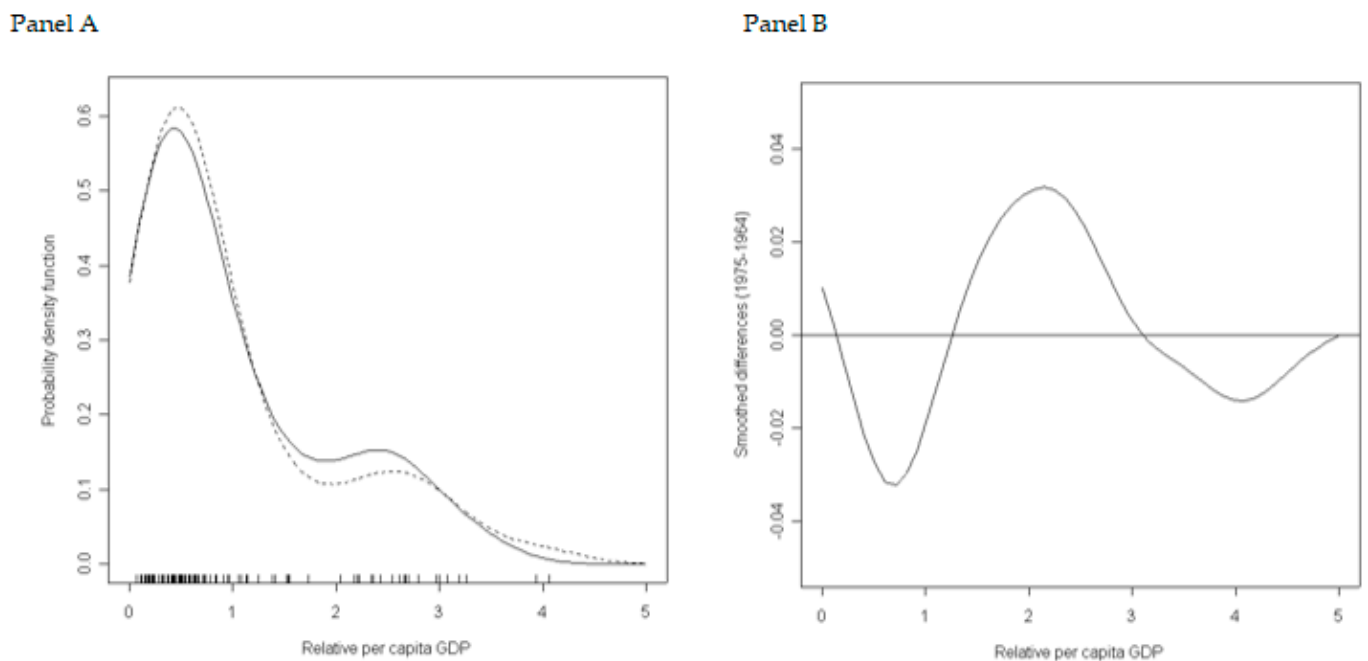
more distant than earlier, even if a catch-up process occurred. Unimodality would be more strongly rejected. However, given a convergence process has occurred, this conclusion would be wrong. As an example, Figure 8 reports estimates of the distributions relative to the per capita GDP across economies for the years 1964 and 1975. These estimates were obtained using an average bandwidth as it is otherwise impossible to compare distributions estimated for different years [44]. Note that between these years, our sample of countries experienced an absolute convergence process. There is a mass shift from the modes to the centre of the distribution. We highlight this in panel B, by showing the smoothed differences between the two densities. Notice that the distributions display a bimodal shape. Hence, bimodality does not necessarily indicate evidence against the ACH.



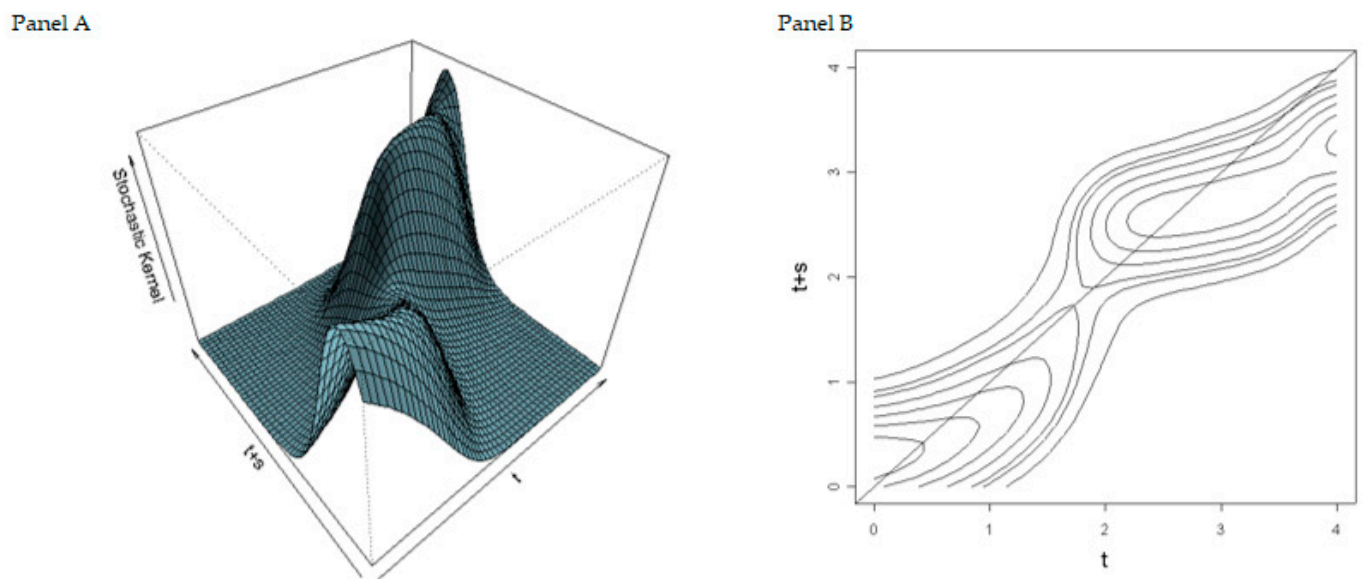
**Figure 7.** Club Convergence and Absolute Convergence. If observations located on the tails of the distribution converge to the centre and the middle class vanishes, absolute convergence and club convergence processes may co-exist.

The issue we are raising is not solved by employing a more refined analysis that builds on the stochastic kernel. Indeed, this analytical tool overshadows the dynamics we have described. The stochastic kernel measures the mobility in the sample, and especially among clusters of economies. However, the two clusters may still catch up with each other, and the stochastic kernel would be silent about this dynamic. To illustrate this, in Figure 9, we report the stochastic kernel estimated for the period 1967 to 1975. It is clear that the distribution changes its shape over time, as suggested by the fact that the distribution has a much less evident mode in 1967 and the second mode is more evident in 1975. Notice also that the range of the estimate decreases because the variance reduces. While it is true that the dispersion cannot inform about all the dynamics within the distribution, this does not mean that it says nothing about convergence dynamics. The evidence confirms, in this case, the issue we are discussing. There is both convergence and polarization in the period from 1964 to 1975. Hence, the evidence supports both the ACH and the CCH. This is the case because both the variance decreases and the observations cluster in groups.





**Figure 8.** Convergence between 1964 and 1975. Full lines represent estimates for 1975; dashed lines represent estimates for 1964. Panel A shows estimates obtained under the hypothesis that the data are normally distributed; panel B shows the smoothed differences between 1975 and 1964.



**Figure 9.** Stochastic Kernel Analysis (1964–1975). Panel A shows the stochastic kernel estimated for transitions of 20 years (i.e.  $s = 20$ ) between 1964 and 1975 for world countries. Vectors  $t$  and  $t + s$  consist of 1248 observations of relative per capita GDP. The estimate is performed by means of the Gaussian Kernel, under the hypothesis that the data are normally distributed. Panel B shows the contour plot of the estimate for fixed levels of probability ( $p = 0.3, 0.4, 0.5, 0.6, 0.7, 0.8, 0.9$ ).

The finding that the modes are about the 45-degree line is interpreted as evidence in favour of the presence of some basins of attraction for the sample, in the sense that they represent the long-run equilibrium positions for the economies. However, persistence may not be the only conclusion in such a case. Suppose we analyse a 40-year time span, during which economies experience convergence in the first 20 years and divergence thereafter. By applying the stochastic kernel to the entire period, we conclude in favour of persistence and

against the ACH. Actually, the observations exhibit interesting and qualitatively different dynamics over time; however, from a historical perspective, we are unable to determine in what year the convergence process begins or ends, or whether there has been only persistence over time. To the extent that this information would help the researcher to assess the reasons why these changes occurred, being able to discriminate across periods is clearly important. However, the stochastic kernel approach is silent in this respect:  $s$  has to be decided a priori.

### 5. Conflicting Views? The Theoretical Relationship between the CCH and the ACH

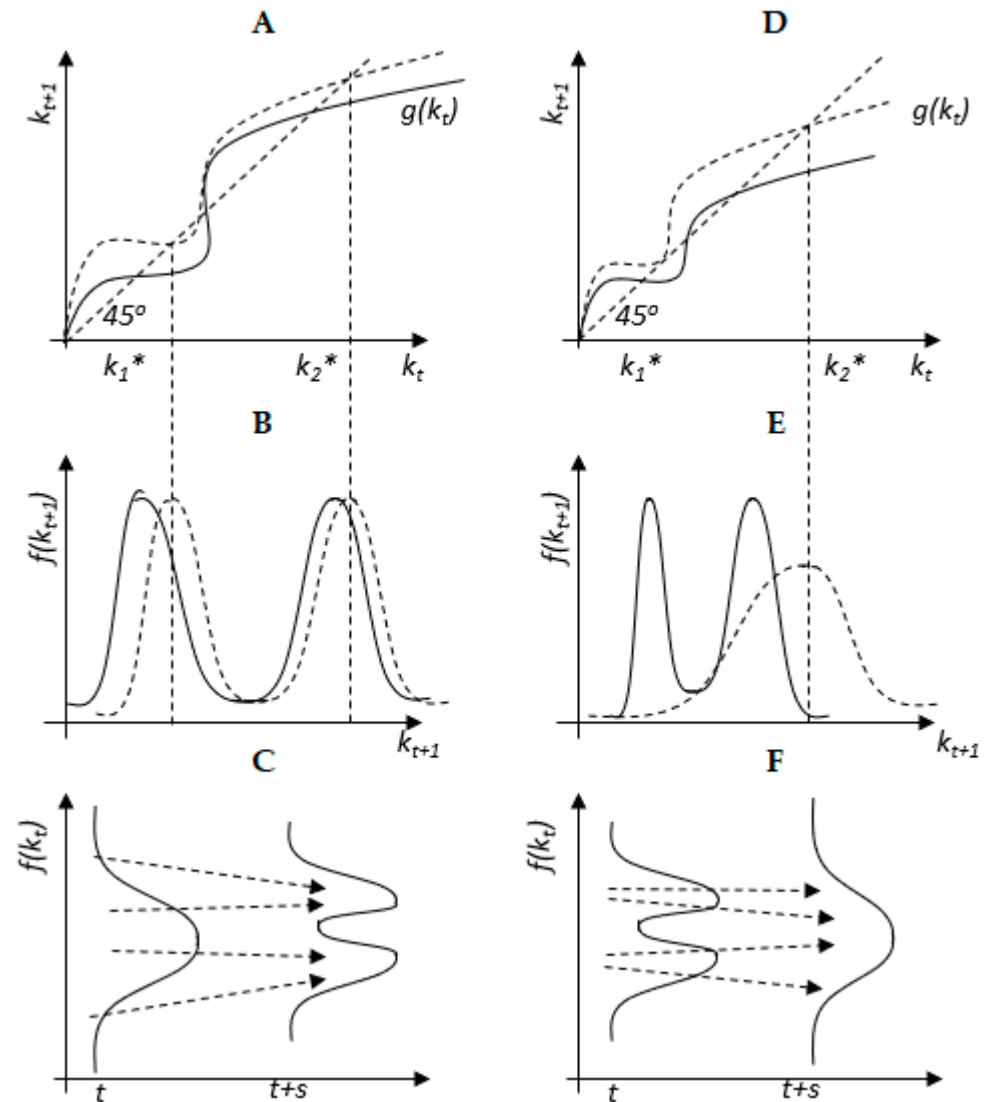
The approach according to which the presence of clubs in the distribution of incomes is necessarily evidence against the ACH does not find theoretical support. On the one hand, absolute convergence dynamics involve all the economies under analysis; on the other hand, club convergence involves only a subgroup of countries (or regions) within the population or the sample under examination. Hence, it is reasonable to expect that, if both club and absolute convergence are in place simultaneously, they occur at different speeds. Under closer examination, if they occur jointly, it is also reasonable to expect that the club convergence dynamics are faster than the absolute convergence process. This is because it is likely that economies or regions that are geographically closer to each other, or maybe share the same institutions, or are under the same economic policy, converge with each other first—if they converge—and that this club convergence dynamic is faster than the process of convergence with the eventual remaining regions or economies.

In all the examples above, club convergence would hide the absolute convergence dynamics. In offering a theoretical argument in favour of this hypothesis, Galor (1996) [45] contends that multiple equilibria can be a temporary phenomenon when technological progress spills over among a given group of economies first and then to the remaining units in the distribution, possibly because of a trade agreement, or foreign direct investment, or because of geographic proximity. In this case, the emergence of clubs is a likely intermediate process towards a more general absolute convergence pattern. Clearly, it is difficult to give a definition of the time span for the two different processes to eventually emerge; therefore, we cannot rule out the potential case that two long-run equilibria coexist with a variance that reduces over time. Figure 10 provides a graphical illustration of the argument.

In Panel A, we report the accumulation path of capital per worker,  $g(k)$ , in the case where, as described, there are non-convexities in the production function. If we assume that the economies share the same technology, as described above, depending on their initial positions, the economies will converge to the high-capital stable equilibrium, or to the low-capital stable equilibrium. Therefore, there will be two long-run stable equilibria in the distribution of incomes per capita, as in Panel B (full line). Assuming now that there is technological progress, the position of the accumulation path will change. More specifically, the path of capital accumulation rotates counter-clockwise. The new path of the accumulation of capital is represented with a dashed line, again in Panel A. Next, we assume that economies clustered at the lower stable equilibrium gain a greater advantage from the technological progress. This assumption is far from being new and it relates to the possibility that poorer economies gain advantage from the technological progress as they can copy it and have lower costs [46]. Therefore, because of the technological progress, all the economies will grow, but poor countries will grow more than rich economies. The new distribution of income is now represented in Panel B with a dashed line. The distribution is still bimodal; however, the low cluster shifts to the right more than the high cluster. We have two clusters of economies, but an absolute convergence process did take place. The variance in the distribution decreases.

In Panel C we report a hypothetical succession of densities, where the middle class tends to vanish. The observations approach to two equilibria. However, over time the tails are less fat and the variance reduces. In its extreme synthesis, we have a twin-peaks dynamic, but also a falling variance. If we look at the distribution only, the emergence of a bimodal distribution supports the CCH and rejects the ACH. Should the ACH be

necessarily rejected? We argue that it may not: because the dispersion decreases, the economies tend to converge to the centre of the distribution of incomes. Therefore, even if two clusters emerge and the middle class vanishes, club convergence may coexist with absolute convergence. Indeed, the two equilibria may or may not converge thereafter.



**Figure 10.** Transition to equilibrium and technological progress. Panel A shows the convergence dynamics in presence of non-constant returns to scale. Panel B shows the convergence patterns of a sample of economies in presence of technological progress.

Quite surprisingly, the tendency to unimodality does not necessarily support the ACH. In Panel D, we report the accumulation path for all the economies in the sample (full line). There exist two equilibria so that the implied distribution of income is bimodal (Panel E, full line). If the technological progress causes the accumulation path to rotate enough, there are no more equilibria on the non-convex part of the accumulation path (Panel D, dashed line). In this case, the distribution is unimodal (Panel E, dashed lines). In the long run, the economies may or may not converge to a uniquely defined equilibrium. However, over time we see a distribution with only one mode, and with a larger variance than the distribution with two modes at  $t$ . Therefore, the ACH should be rejected even if the distribution transforms from bimodality to unimodality (Panel E).

## 6. Conclusions and Directions for Further Research

We did not aim at providing a new and comprehensive review of the debate on convergence. This is the main limitation of this article. Instead, we contribute to the literature in three respects, and suggest directions for further research. The first contribution we offer is the conclusion is that the different hypotheses about convergence, namely, the Conditional Convergence Hypothesis, the ACH, and the CCH, should not be taken as competing hypotheses when performing empirical analysis.

In turn, the above suggests that evidence in favour of one of them is not necessarily against the other, as we might expect. Indeed, the Conditional Convergence Hypothesis is based on testing whether each economy converges to its steady-state position. The steady-state position of the various economies may differ because of different habits or behaviours of the agents with respect to their saving choices, for example, but also due to other fundamental variables such as institutions, natural resource endowment, culture, and so on. Therefore, evidence in favour of the Conditional Convergence Hypothesis (or against it) should not be taken as being in favour or against the ACH or the CCH. Of course, it is still possible that these fundamental variables cluster in groups or diverge in the sample, and therefore, we may also observe polarization and or divergence. However, as we do not know whether this is the case, conclusions about the CCH and the ACH building on evidence against or in favour of the Conditional Convergence Hypothesis are essentially weakened.

The second contribution we offer is that, more interestingly, evidence in favour of the CCH is not necessarily against the ACH, and evidence in favour of the ACH is not necessarily against the CCH. Polarization in the distribution of incomes is typically taken as evidence in favour of the CCH, and against the ACH. However, there are neither empirical nor theoretical reasons to expect this to be the case. As an example, we have reported the case of economies between 1964 and 1975. During this 11-year time span, the mode representing the rich cluster becomes increasingly apparent, so the evidence is in favour of the CCH. However, this mode is closer to the mode representing the poor cluster, and the variance reduces. Hence, the evidence is in favour of the ACH also. This would be the correct inference to make, as when the two processes are occurring at the same time, they are likely to occur at different speeds. Clustering involves observations that are likely similar to each other in some respect; hence, the polarization process is likely to occur at a higher speed than the absolute convergence process and, therefore, to overshadow it. As with the case above, when testing for the CCH, it is risky to conclude in favour or against the ACH based on this evidence. In turn, this implies that an explicit test of the behaviour of clusters eventually emerging in the distribution of incomes is still needed. This is our third contribution to the literature.

One final point of discussion relates to the econometric modelling of convergence dynamics and, more generally, of growth processes. In the majority of empirical works, economies are taken as independent observations. This should be regarded a strong hypothesis because with the marked globalization process, interactions among economies, through international trade and migration processes, for example, define the winners and losers in growth, and whether observations converge to each other. Studies about convergence, especially at the regional level, should instead explicitly model such interactions.

**Funding:** This research received no external funding.

**Data Availability Statement:** Not applicable.

**Conflicts of Interest:** The author declares no conflict of interest. The funders had no role in the design of the study; in the collection, analyses, or interpretation of data; in the writing of the manuscript; or in the decision to publish the results.

## Appendix A

Table A1. Correlations between Variables.

	gYN	gYL	YL <sub>to</sub>	OPEN/Y	gHL	gKL	gL
gYN	1						
gYL	0.97 **	1					
YL <sub>to</sub>	−0.11	−0.21	1				
OPEN/Y	0.28 *	0.30 *	−0.04	1			
gHL	0.56 **	0.64 **	−0.42 **	0.25	1		
gKL	0.72 **	0.74 **	−0.12 *	0.05	0.45 **	1	
gL	−0.30 *	−0.33 *	−0.46 **	−0.14	0.04	−0.32 **	1

Note: Correlations are calculated on the basis of the entire period. \*\* (\*) stands for significance at the 1% (5%) level.

## References

- Baumol, W.J. Productivity Growth, Convergence and Welfare. What Long-Run Data Show. *Am. Econ. Rev.* **1986**, *76*, 1072–1085.
- Abramowitz, M. Catching up, Forging Ahead and Falling Behind. *J. Econ. Hist.* **1986**, *46*, 385–406. [[CrossRef](#)]
- Barro, R.J.; Sala-i-Martin, X. Convergence across States and Regions. *Brook. Pap. Econ. Act.* **1991**, *1*, 1107–1182. [[CrossRef](#)]
- Barro, R.J.; Sala-i-Martin, X. Convergence. *J. Political Econ.* **1992**, *100*, 223–251. [[CrossRef](#)]
- Barro, R.J.; Sala-i-Martin, X. Technological Diffusion, Convergence, and Growth. *J. Econ. Growth* **1997**, *2*, 1–26. [[CrossRef](#)]
- Bernard, A.; Durlauf, S.N. Convergence in International Output. *J. Appl. Economet.* **1995**, *10*, 97–108. [[CrossRef](#)]
- Ben-David, D. *Convergence Club and Diverging Economies*; Working Paper 922; CEPR: Washington, DC, USA, 1994.
- Canova, F.; Marcet, A. *The Poor Stay Poor: Non-Convergence across Countries and Regions*; Discussion Paper No. 1265; Centre for Economic Policy Research (CEPR): London, UK, 1995.
- Islam, N. Growth Empirics: A Panel Data Approach. *Q. J. Econ* **1995**, *110*, 1127–1170. [[CrossRef](#)]
- Quah, D. Empirics for Growth and Distribution: Stratification, Polarization and Convergence Clubs. *J. Econ. Growth* **1997**, *2*, 27–59. [[CrossRef](#)]
- Bianchi, M. Testing for convergence: Evidence from nonparametric multimodality test. *J. Appl. Economet.* **1997**, *12*, 393–409. [[CrossRef](#)]
- Caggiano, G.; Leonida, L. A note on the empirics of the neoclassical growth model. *Econ. Lett.* **2007**, *94*, 170–176. [[CrossRef](#)]
- Caggiano, G.; Leonida, L. International output convergence: Evidence from an autocorrelation function approach. *J. Appl. Economet.* **2008**, *24*, 139–162. [[CrossRef](#)]
- Islam, N. What Have We Learnt from the Convergence Debate? *J. Econ. Surv.* **2003**, *17*, 309–362. [[CrossRef](#)]
- Temple, J. The New Growth Evidence. *J. Econ. Lit.* **1999**, *37*, 112–156. [[CrossRef](#)]
- De la Fuente, A. Convergence Across Countries and Regions: Theory and Empirics. *EIB Pap.* **2000**, *2*, 25–43.
- Lamo, A. On Convergence Empirics: Some Evidence for Spanish Regions. *Investig. Econ.* **2000**, *24*, 681–707.
- Parente, S. The Failure of Endogenous Growth. *Knowl. Technol. Policy* **2001**, *13*, 49–58. [[CrossRef](#)]
- Durlauf, S.N.; Quah, D.T. The New Empirics of Economic Growth. In *Handbook of Macroeconomics*; Taylor, J., Woodford, M., Eds.; Elsevier: Amsterdam, The Netherlands, 1999; pp. 235–308.
- Dowrick, S.; DeLong, B.J. Globalization and Convergence. In *Globalisation in Historical Perspective*; Bordo, M.D., Taylor, A.M., Williamson, J.G., Eds.; University of Chicago Press: Chicago, IL, USA, 2003; pp. 191–225.
- Grinin, L.; Korotayev, A. *Great Divergence and Great Convergence: A Global Perspective*; Springer: Berlin/Heidelberg, Germany, 2015; pp. 1–102.
- Johnson, P.; Papageorgiou, C. What Remains of Cross Country Convergence? *J. Econ. Lit.* **2020**, *8*, 129–175. [[CrossRef](#)]
- Caselli, F.; Esquivel, G.; Lefort, F. Reopening the convergence debate: A new look at cross-country growth empirics. *J. Econ. Growth* **1996**, *1*, 363–389. [[CrossRef](#)]
- Sala-i-Martin, X. The Classical Approach to Convergence Analysis. *Econ. J.* **1996**, *106*, 1019–1036. [[CrossRef](#)]
- Kremer, M.; Willis, J.; Yu, Y. Converging to Convergence. *BFI WP* **2021**, 1–50. [[CrossRef](#)]
- Patel, D.; Sandefur, J.; Subramanian, A. The new era of unconditional convergence. *J. Dev. Econ.* **2021**, *152*, 102687. [[CrossRef](#)]
- Lima, L.R.; Hostalacio, H.; Reis Gomes, F.A. Empirical Evidence on Convergence Across Brazilian States. *Rev. Braz. Econ.* **2010**, *64*, 135–160.
- Gonzalez-Paramo, J.M.; Martinez-Lopez, D. Convergence Across Spanish Regions: New Evidence on the Effect of Public Investment. *Rev. Reg. Stud.* **2003**, *33*, 184–205. [[CrossRef](#)]
- Puente, S. *Regional Convergence in Spain: 1980–2015*; Banco de Espana: Madrid, Spain, 2017; pp. 1–7.
- Goeche, H.; Huter, M. Regional Convergence in Europe. *Intereconomics* **2016**, *51*, 165–171. [[CrossRef](#)]
- Eichengreen, B. Convergence and Divergence in the EU: Lessons from Italy. *Intereconomics* **2019**, *54*, 31–34. [[CrossRef](#)]
- Savoia, F. Income Inequality Convergence Across EU Regions. *LIS WP* **2020**, *760*, 1–40.

33. Ram, R. Income convergence across the U.S. states: Further evidence from new recent data. *J. Econ. Finan.* **2021**, *45*, 372–380. [[CrossRef](#)]
34. Kholodilin, K.A.; Oshchepkov, A.; Siliverstovs, B. The Russian Regional Convergence. *East. Eur. Econ.* **2012**, *50*, 5–26. [[CrossRef](#)]
35. Lehmann, H.; Oshchepkov, A.Y.; Silvagni, M.G. *Regional Convergence in Russia: Estimating a Neoclassical Growth Model*; IZA—Institute of Labor Economics: Bonn, Germany, 2020; p. 13039.
36. Barro, R.J. Notes on Growth Accounting. *J. Econ. Growth* **1999**, *24*, 119–137. [[CrossRef](#)]
37. Aten, B.; Heston, A.; Summers, R. *Penn World Table*, Version 6.1; CIC UPENN: Philadelphia, PA, USA, 2001.
38. Barro, R.; Lee, J.W. International Measures of Schooling Years and Schooling Quality. *AER Pap. Proc.* **2000**, *86*, 218–223.
39. Summers, R.; Heston, A. The Penn World Table (Mark 5.6): An Expanded Set of International Comparisons, 1950–1988. *Q. J. Econ.* **1996**, *106*, 327–368. [[CrossRef](#)]
40. Bairam, E.I.; McRae, S.D. Testing the Convergence Hypothesis: A New Approach. *Econ. Lett.* **1999**, *64*, 351–355. [[CrossRef](#)]
41. Mankiw, N.G.; Romer, D.; Weil, D. A Contribution to the Empirics of Economic Growth. *Q. J. Econ.* **1992**, *107*, 407–437. [[CrossRef](#)]
42. Quah, D. Twin Peaks: Growth and Convergence in Models of Distribution Dynamics. *Econ. J.* **1996**, *106*, 1045–1055. [[CrossRef](#)]
43. Quah, D. Empirical Cross-Section Dynamics in Economic Growth. *Eur. Econ. R.* **1993**, *37*, 426–434. [[CrossRef](#)]
44. Caggiano, G.; Leonida, L. Multimodality in the distribution of GDP and the absolute convergence hypothesis. *Empir. Econ.* **2013**, *44*, 1203–1215. [[CrossRef](#)]
45. Galor, O. Convergence? Inference from Theoretical Models. *Econ. J.* **1996**, *106*, 1056–1069. [[CrossRef](#)]
46. Jones, C.I. Time Series Tests of Endogenous Growth Models. *Q. J. Econ.* **1995**, *110*, 495–525. [[CrossRef](#)]

**Disclaimer/Publisher’s Note:** The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.