

Article

Lack of Global Convergence and the Formation of Multiple Welfare Clubs across Countries: An Unsupervised Machine Learning Approach

Carlos Mendez 

Graduate School of International Development, Nagoya University, Aichi 464-8601, Japan;
carlos@gsid.nagoya-u.ac.jp

Received: 6 June 2019; Accepted: 8 July 2019; Published: 17 July 2019



Abstract: The cross-country convergence hypothesis is one of the central topics of long-run macroeconomics. This paper revisits this hypothesis in a context beyond GDP. It uses a novel welfare index that incorporates measures of consumption, leisure, life expectancy, and inequality. Based on a sample of 128 countries over the 1980–2007 period, the lack of global sigma and beta convergence is first documented. Next, the paper incorporates some recent developments from the unsupervised machine learning literature to evaluate the existence of local convergence. In particular, the application of a distribution-based clustering algorithm suggests the formation of three local convergence clubs. Under this classification, beta convergence is recovered for each club. However, only the core members of the richest club appear to be reducing their welfare differences in a way that is consistent with the strong notion of sigma convergence. Overall, these results re-emphasize the finding that beta convergence is necessary, but not sufficient for sigma convergence, even within convergence clubs and in a context beyond GDP.

Keywords: convergence; welfare; machine learning; non-parametric distribution; distribution dynamics; distribution-based clustering

1. Introduction

In the study of long-run macroeconomics, a central topic is the empirical testing of the convergence hypothesis across countries (Johnson and Papageorgiou 2018). Differences across countries are conceptually defined in terms of living standards or national welfare. However, from an empirical and operational standpoint, GDP per-capita has been used in most of the literature as the key proxy variable for measuring living standards. Although the use of GDP per-capita is sometimes useful and informative,¹ economists are well aware that it is an incomplete measure of national welfare (Fleurbaey 2009; Fleurbaey and Blanchet 2013; Stiglitz et al. 2010, 2019). In this context, a series of different alternatives has been proposed by economists and non-economists alike (Becker et al. 2005; Cordoba and Verdier 2008; Fleurbaey and Gaulier 2009). One of the most recent attempts to go beyond GDP is the work of Jones and Klenow (2016). Using a rigorous expected utility framework, these authors combined measures of national consumption (private and public), leisure, life expectancy, and inequality to construct a theoretically-appealing new welfare index. Given the comprehensive coverage of this index, it can be used as an alternative proxy variable for the study of the convergence hypothesis.

Machine learning is a fast-growing field of research that focuses on the development of prediction, classification, and clustering algorithms (Athey 2018; Athey and Imbens 2019). It is commonly divided

¹ GDP per capita is a useful variable in the sense that it correlates with other human development variables such as educational attainment, life expectancy, and even subjective happiness.

into two sub-fields: supervised and unsupervised machine learning. The latter focuses on finding clusters of observations using covariates, and it is the main approach used in the current paper. Traditional clustering frameworks such K-means and hierarchical clustering derive their results based on some criteria of distance. In contrast, distribution-based clustering frameworks derive their results by defining clusters as regions of high-probability density separated by low-probability regions (Hartigan 1975; Kaufman and Rousseeuw 2010).

This paper revisits the cross-country convergence hypothesis in the context beyond GDP suggested by Jones and Klenow (2016). From a methodological standpoint, this paper first applies the classical sigma and beta convergence tests of Barro and Sala-i Martin (1992). Then, its fundamental contribution relies on the implementation of the distributional convergence approach of Quah (1993, 1996b, 1997). Furthermore, by taking advantage of some recent developments in the unsupervised machine learning literature, it extends the distributional convergence framework by integrating the distribution-based clustering algorithm of Azzalini and Menardi (2014a). This extension allows for a more detailed identification and characterization of convergence clubs.

Based on a sample of 128 countries over the 1980–2007 period, this paper finds that welfare differences across countries are characterized by a lack of both sigma and beta convergence. Moreover, limited country mobility suggests the possible existence of convergence clubs. The distributional convergence approach builds on these findings and suggests that the cross-country welfare distribution appears to be characterized by three convergence clubs or clusters. Interestingly, under this classification, the beta convergence is recovered for each club. However, only the core members of the richest club appear to be converging in a way that is consistent with the strong notion of sigma convergence. Overall, these results re-emphasize a central finding in the long-run macroeconomics literature: beta convergence is necessary, but not sufficient for sigma convergence, even within convergence clubs and in a context beyond GDP.

The rest of the paper is organized as follows: Section 2 describes the data and the different convergence frameworks. Section 3 introduces some stylized convergence facts about welfare differences across countries. Section 4 presents the convergence clubs' results. Section 5 discusses the results and outlines some suggestions for further research. Finally, Section 6 offers some concluding remarks.

2. Data and Methods

2.1. Beyond GDP Data

The work in Jones and Klenow (2016) proposed a new summary statistic that aims to quantify the level of welfare of people in a country. This statistic incorporates measures of consumption (private and public), leisure, life expectancy, and inequality. Internationally-comparable measures of consumption are taken from the Penn World Table 8.0. This database also provides the necessary inputs to compute a measure of leisure time from the available series of hours worked, employment, and population data. Life expectancy data are from the World Bank's HNPStatsdatabase. Finally, inequality data are from the UNI-WIDERWorld Income Inequality database 3.0.²

To aggregate these measures, in a theoretically-consistent way, the work in Jones and Klenow (2016) calibrated an expected utility function and evaluated the consumption-equivalent level of the four variables. Using this statistic, they found that, on average, welfare is highly correlated with GDP per capita. However, they also reported that there are often large deviations from GDP, in particular in developing countries. For instance, in terms of the relative welfare differences across regions, Western Europe appears closer to the United States; fast-growing Asia has not caught up as much; and most countries in Latin America and Africa are lagging further behind. The work in Jones and Klenow

² For further details on the measurement and calculations of the variables, see the Appendix of Jones and Klenow (2016).

(2016) constructed this welfare statistic for 152 countries in the year 2007. They also calculated the average growth rate for the 1980–2007 period using a sub-sample of 128 countries³. Given the welfare level data for the final year (2007) and the average growth rate (1980–2007 period), it is possible to recover cross-sectional data for the initial year (1980). Thus, based on the availability of the data, this paper uses a sample of 128 countries for two time periods: 1980 and 2007.⁴

2.2. Sigma and Beta Convergence

In the empirical literature of economic growth, two concepts of convergence are typically discussed. On the one hand, the concept of sigma (σ) convergence describes the evolution of the cross-sectional dispersion of a variable. From this perspective, convergence occurs when the cross-sectional dispersion declines over time, so the level of the variable under study becomes increasingly more similar across countries (Barro and Sala-i Martin 1992; Baumol 1986; Dowrick and Nguyen 1989). Typically, sigma convergence is measured by the coefficient of variation or by the standard deviation of the logarithm of the variable under study. In the context of the current paper, the latter indicator is adopted, and the measurement of sigma convergence is implemented as follows:

$$\sigma_t = \sqrt{\frac{1}{N-1} \sum_{i=1}^N \left(\log(y_i) - \overline{\log(y)} \right)^2}, \quad (1)$$

where σ_t is the cross-country dispersion in national welfare, N is the number of countries, $\log(y_i)$ is the natural logarithm of the welfare level of country i , and $\overline{\log(y)}$ is the sample average of the logarithm of welfare.

On the other hand, the concept of unconditional beta (β) convergence describes the inverse relationship between the initial level of a variable and its average growth rate. From this perspective, if such an inverse relationship exists, it means that, on average, poor countries tend to grow faster than the rich ones, so over time, poor countries tend to catch up with the level of the rich ones (Barro and Sala-i Martin 1992; Baumol 1986; Sala-i Martin 1996). Typically, beta convergence is measured by the estimation of an Ordinary Least Squares (OLS) regression, in which the growth rate of a variable inversely depends on its initial level. In the context of the current paper, the measurement of beta convergence is implemented as follows:

$$\left(\frac{1}{T} \right) \log \left(\frac{y_t}{y_0} \right) = c - \frac{(1 - e^{-\beta t})}{t} \log(y_0), \quad (2)$$

where the left side represents the average rate of welfare growth, which shows an inverse relationship to its initial level in log terms, $\log(y_0)$, β represents the speed of convergence to the steady-state, and c represents unobserved parameters, such as steady-state values.

These two measures of convergence are related⁵. Keeping other variables constant, when poor countries tend to grow faster than rich ones (that is, beta convergence), then the cross-country dispersion declines over time (that is, sigma convergence). In other words, beta convergence is one determinant of sigma convergence. However, the effect of beta convergence can be offset by other variables and shocks that increase the dispersion at any point in time. As pointed out by Quah (1993) and Sala-i Martin (1996), beta convergence is necessary, but not sufficient to achieve sigma convergence.

³ The database can be accessed at <https://web.stanford.edu/~chadj/BeyondGDP500.xls>.

⁴ Although one could use the latest versions of the Penn World Table, UNI-WIDER, and the World Bank databases to extend the analysis beyond 2007, the purpose of this paper is to be directly comparable with the paper of Jones and Klenow (2016). Thus, the reference period for comparison is still the 1980–2007 period. Further research, beyond the scope of this paper, could extend the period of analysis and evaluate the persistence of the convergence clusters.

⁵ See Sala-i Martin (1996) for further details.

2.3. Distributional Convergence and Convergence Clubs

The distributional convergence approach aims to capture the evolution of the entire cross-country distribution of a variable. From this perspective, convergence occurs when the shape of the cross-sectional distribution tends to one density peak (one mode) over time. In this framework, the emergence of multiple peaks or modes is usually associated with the existence of convergence clubs (Galor 1996; Magrini 2009; Quah 1996a 1997). Typically, distributional convergence is measured by the shape and the number of modes of a stochastic kernel distribution and its corresponding ergodic distribution. In the current paper, only the former is adopted, and the measurement of distributional convergence⁶ is implemented as follows:

1. The variable under study (that is, national welfare) is expressed relative to a benchmark economy, which in the literature is usually the United States. The purpose of this normalization is to abstract from systematic forces that might simultaneously affect all countries.
2. To facilitate comparison and visualization, the natural logarithm of the relative variable is applied. The log of a relative variable can be interpreted as the proportional difference between a country and the benchmark country (i.e., the convergence frontier).
3. The stochastic kernel is a conditional distribution that is calculated as follows:

$$G(y_{t+s} | y_t) = \frac{f_{t+s,t}(y_{t+s}, y_t)}{f_t(y_t)}, \quad (3)$$

where $f_t(\cdot)$ is the univariate kernel distribution of relative welfare in the initial year, t , and $f_{t+s,t}(\cdot)$ is the (inter-temporal) bivariate kernel distribution between the years.

4. The bivariate kernel distribution is estimated as follows:

$$f_{t+s,t}(y_{t+s}, y_t) = \frac{1}{nh_{t+s}h_t} \sum_{i=1}^n K_{t+s} \left(\frac{y_{t+s} - y_{t+s,i}}{h_{t+s}} \right) K_t \left(\frac{y_t - y_i}{h_t} \right), \quad (4)$$

where y_{t+s} and y_t denote the relative welfare of each country at time $t + s$ and t , respectively, K_{t+s} and K_t denote kernel functions, and h_{t+s} and h_t denote the smoothing parameters of y_{t+s} and y_t , respectively. Following the convention of the literature, the kernel functions adopt a Gaussian form, and the smoothing parameters are selected based on the minimization of the Asymptotic Mean Integrated Square Error (Magrini 2009).

The stochastic kernel is a three-dimensional object that is commonly represented by a surface plot or a contour plot. If most countries are concentrated around the main diagonal of these graphs, then there is evidence of distributional persistence over time. Global distributional convergence is found when most of the countries are located around zero in the $(t + s)$ -axis and parallel to the t -axis.

Finally, the distribution-based clustering algorithm developed by Azzalini and Menardi (2014a) is applied with the previously-described kernel functions and smoothing parameters. The implementation of this new clustering framework is useful for two purposes. First, it helps us identify the location of each country within the distribution. Second, it helps us allocate each country to its nearest convergence club.

3. Some Stylized Facts

3.1. Lack of Sigma and Beta Convergence

The classical analysis of convergence shows that, similar to patterns of GDP per capita documented by Barro and Sala-i Martin (1992), welfare differences across countries are characterized by a lack

⁶ For a more comprehensive and recent presentation of the distributional convergence approach, see Dal Bianco (2016), Durlauf et al. (2005), Epstein et al. (2003), or Mendez (2018)

of both sigma and beta convergence. Figures 1 and 2 summarize this finding. Figure 1 shows that, although the distribution has shifted to the right, the cross-country dispersion of welfare has increased between 1980 and 2007. As such, this result highlights the lack of sigma convergence in the context of welfare differences across countries.

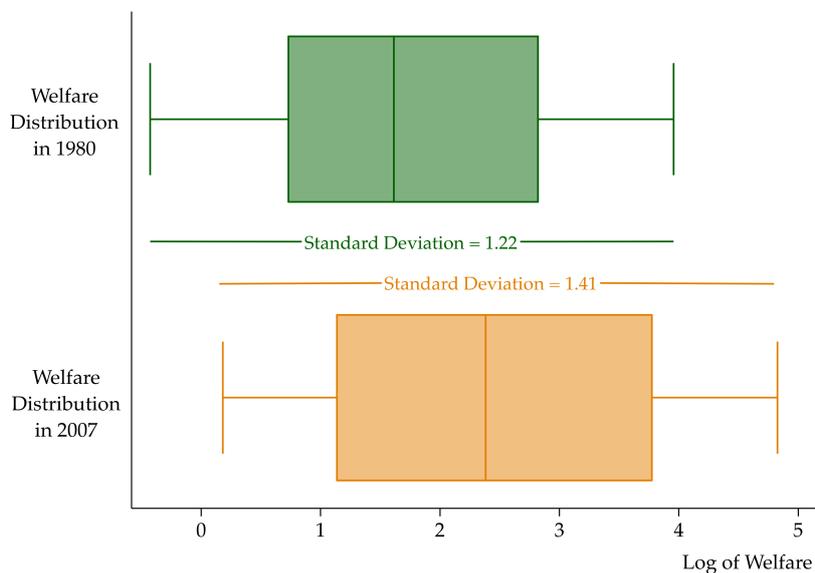


Figure 1. Lack of sigma convergence in welfare across countries.

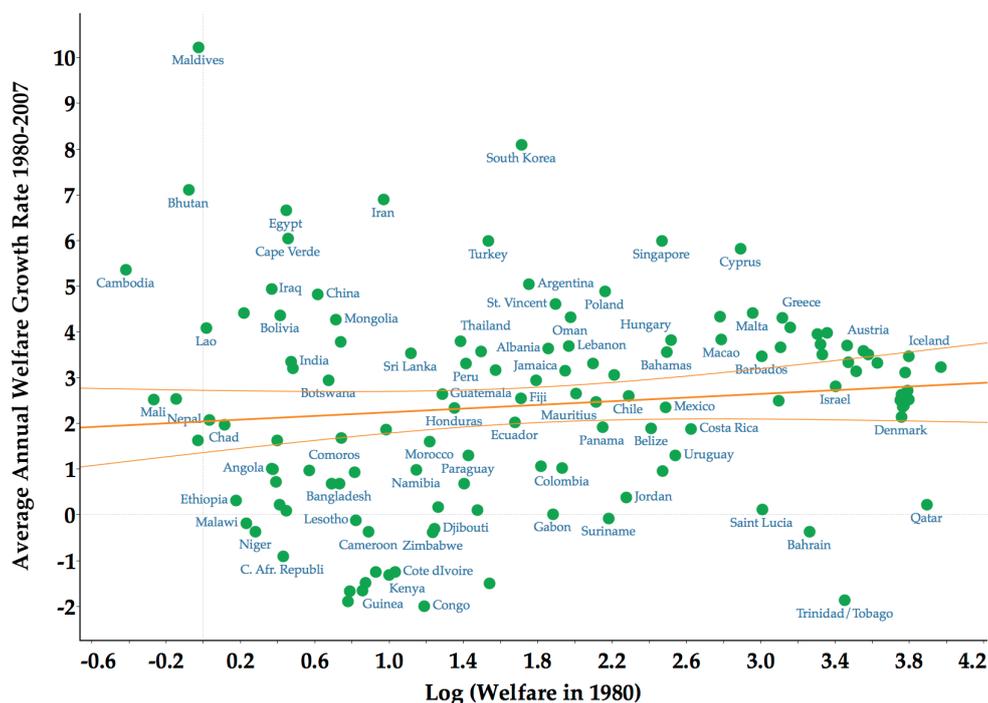


Figure 2. Lack of beta convergence in welfare across countries.

More specifically, the standard deviation of the cross-country welfare distribution in the year 1980 was 1.22. By the year 2007, this dispersion increased to 1.41. This higher degree of cross-country inequality in welfare is also observable in the increasing distance between the third quartile and the first quartile of the boxplot presented in Figure 1.

Figure 2 shows that, on average, welfare-poor countries are not growing faster than welfare-rich countries. As such, this result indicates a lack of beta convergence. Note also that, if anything, the positive (but not significant) slope of the regression line would suggest that welfare-rich countries

are growing faster than welfare-poor countries, and thus contributing to the increasing dispersion reported in Figure 1. Finally, it is also worth noting that the triangular shape of the scatter plot is highly similar to that reported in the studies that document a lack of beta convergence in income (Barro and Sala-i Martin 1992; Sala-i Martin 1996).

3.2. Limited Forward and Backward Mobility

Figure 3 highlights the fact that there is a limited degree of forward and backward mobility across countries. The measurement of country mobility here is relative to that experienced by the frontier, which in this case is the United States. Thus, the lack of relative mobility does not imply the lack of progress in absolute terms. In Figure 3, most countries are located around the 45-degree line. Indeed, the relation between the initial and the final level of welfare is summarized by a linear regression in which the coefficient of the slope is statistically equal to one. The 95% confidence interval for this slope coefficient is between 0.97 and 1.14. Moreover, the R-squared of this regression highlights that 83 percent of the cross-country welfare variation of the year 2007 is explained by the welfare variation of the year 1980.

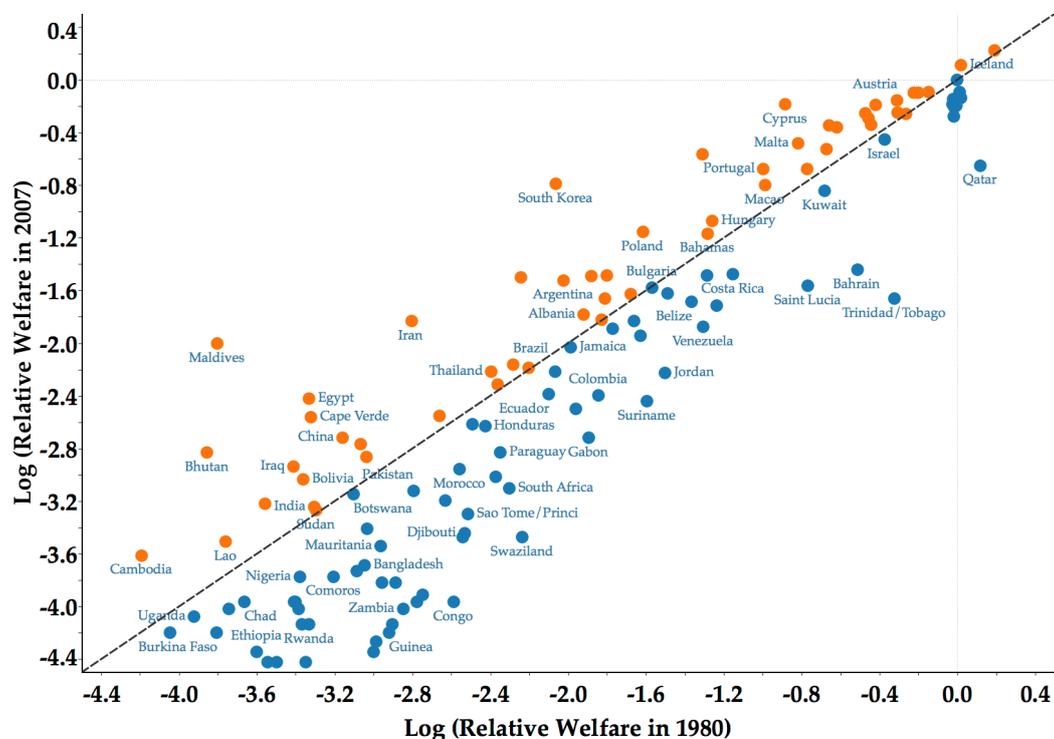


Figure 3. Limited forward and backward mobility in welfare across countries.

Another interesting mobility fact is that 76 out of the 128 countries in the sample (that is, almost 60 percent of the total sample) are moving backwards. In other words, relative to its initial position in 1980, 60 percent of the countries ended up with lower relative welfare in 2007. This backward mobility is more pronounced in the poorest countries in the sample. Finally, both the limited degree of country mobility and the overall lack of cross-country convergence suggest the possible existence of local convergence clubs. The results associated with this hypothesis are presented in the next section.

4. Results

4.1. Transitional Dynamics via the Stochastic Kernel Distribution

Figure 4 shows the 3D surface of the stochastic kernel distribution. The transitional dynamics between 1980 and 2007 were characterized by three density modes along the main diagonal. In this

convergence framework, the lack of country mobility and the emergence of multiple basins of attraction (that is, multiple density modes) in the main diagonal are typically interpreted as suggestive evidence of convergence clubs. In this case, given the existence of three modes, then three different convergence clubs are likely to be present. Interestingly, these three convergence clubs in welfare are qualitatively similar to the three convergence clubs in income reported in [Pittau et al. \(2016\)](#). Although the number of clubs is the same, the methodology of [Pittau et al. \(2016\)](#) is based on a finite Gaussian mixture model, which is a semi-parametric distributional alternative to the non-parametric distribution framework implemented in the current paper.

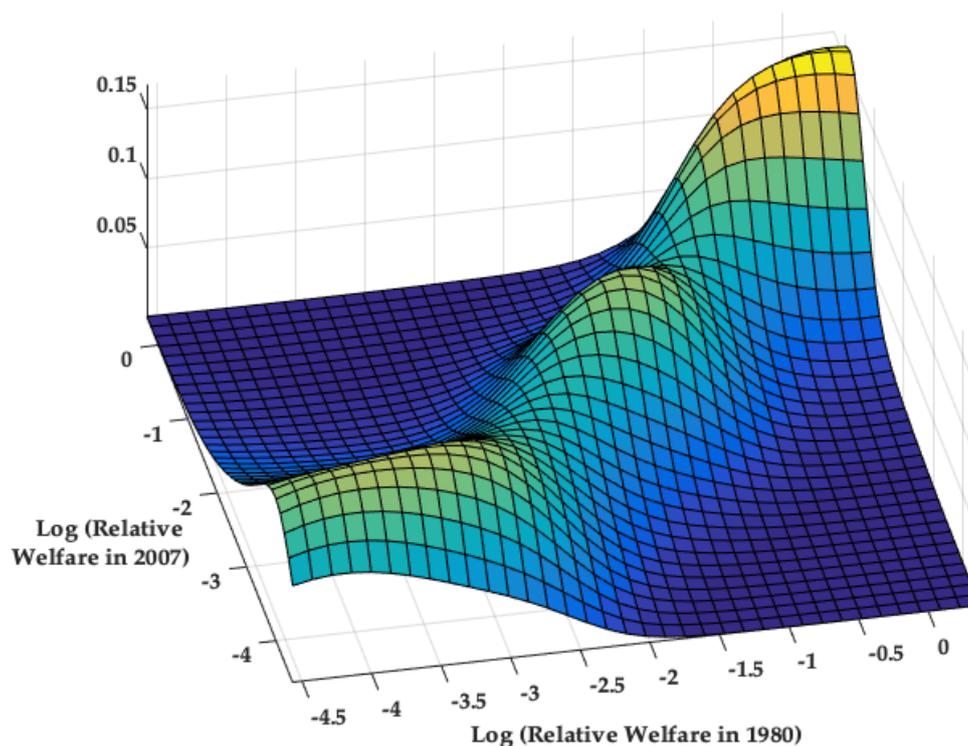


Figure 4. Stochastic kernel distribution and convergence clubs (3D surface plot).

Figure 5 shows the contour plot representation stochastic kernel overlapped with the scatter plot of the countries of Figure 3. It provides a rough overview of the country composition of each convergence club. From the figure, it is clear that some countries were located at the center-core of each club, while the position of other countries raises some doubts about their club membership. In particular, there seemed to be a considerable number of countries between the bottom club and the middle club. For this sub-sample, some countries were moving forward and transitioning towards a superior club. In contrast, other countries were moving backward and transitioning towards an inferior club. To help clarify the membership of these kind of countries, the novel distribution-based clustering algorithm of [Azzalini and Menardi \(2014a\)](#) is implemented in the next subsection.

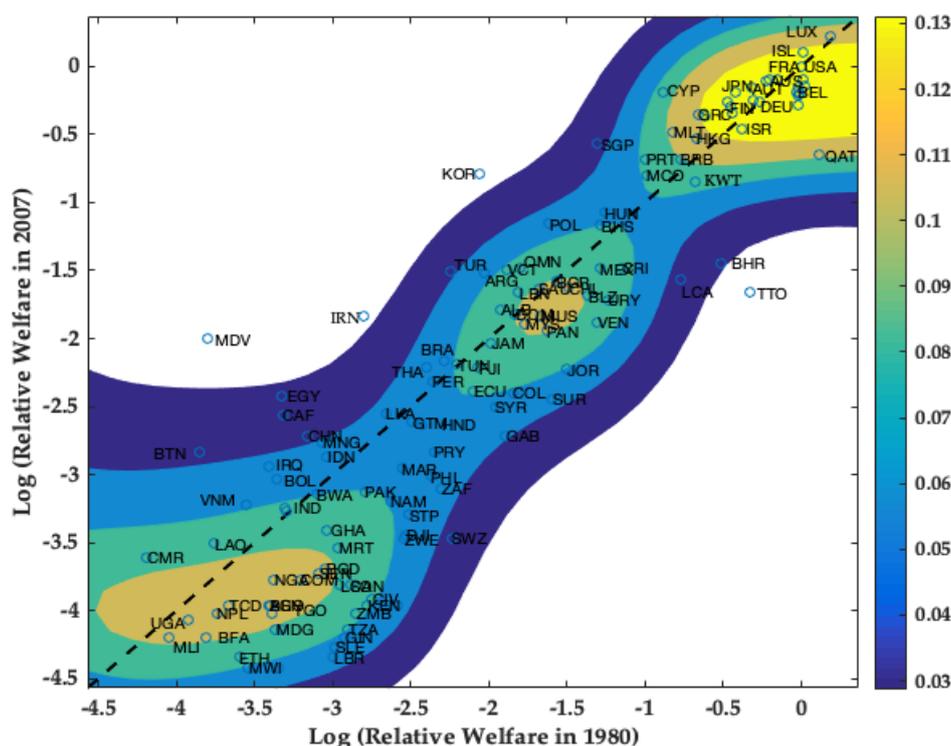


Figure 5. Stochastic kernel distribution and convergence clubs (contour plot).

4.2. Core Clusters and Classification of Countries

In a series of papers, [Azzalini and Torelli \(2007\)](#) and [Azzalini and Menardi \(2014a,2014b\)](#) used the modes of a non-parametric distribution as a criterion for identifying clusters. High-density observations group themselves in what Azzalini and Menardi call “cores clusters”. Low-density observations, on the other hand, could be allocated to a proximate cluster based on a Delaunay triangulation.

Figure 6 shows the results of this approach in the context of the mobility scatter plot of Figure 3. Consistent with the stochastic kernel of Figures 4 and 5, three clusters or clubs were identified. The main advantage of this clustering framework is the endogenous identification of core club members (those with a “1”, “2”, or “3” prefix). In addition, low-density observations (i.e., those with a “0” prefix) were also identified and allocated to their more proximate core club.⁷ In this low-density context, countries such as the Maldives or South Korea are interesting cases to be studied. Although both countries experienced forward mobility (catching up behavior), their progress has not been classified as that of a “core” member of the immediately superior club. In these cases, the clustering classification was still informative in the sense that it was also clear that these kinds of countries did not belong to their immediately inferior “core” club either.

Finally, the bottom right of Figure 6 includes a cluster tree. It is meant to provide a measure of the robustness of the clubs to different density thresholds. Although for a considerably large set of density thresholds, three clubs were identifiable, it is also possible that the countries of Club 1 and Club 2 could be converging to a similar steady state. Even in this case, the cross-country distribution of welfare is likely to be characterized by more than one convergence club.

⁷ See Appendix A for a list of countries their respective clubs.

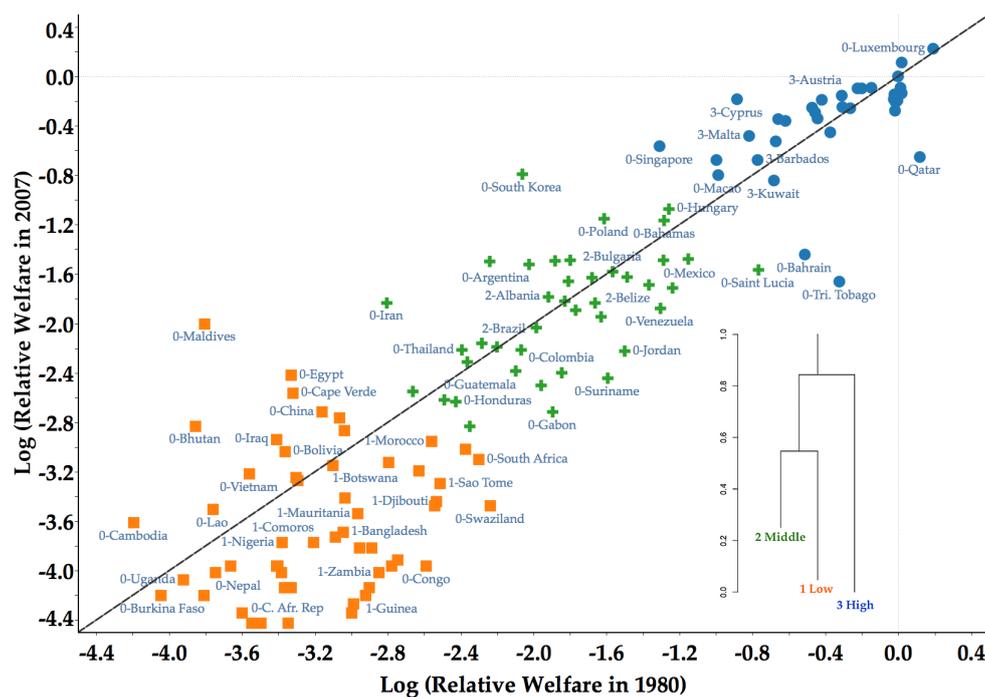


Figure 6. Local convergence clusters.

4.3. Sigma and Beta Convergence within Clubs

Figure 7 and Table 1 present a re-evaluation of the sigma convergence test for each of the previously-identified clubs. Sigma convergence was only present for the core members of the high-welfare club. In other words, only the core members of Club 3 were increasingly becoming more similar in terms of their level of national welfare. For these 28 country members, the standard deviation of the log of welfare was 0.35 in 1980, and by the year 2007, it reduced to 0.21. Moreover, as indicated by the p -value of the ANOVA test, this reduction in the cross-sectional dispersion was statistically significant. For the other country clubs, however, the cross-sectional dispersion increased in spite of the absolute improvements in the average and medium levels of welfare. It also worth noting that the countries in the poorest welfare club showed the largest increase in dispersion.

Figure 8 and Table 2 present a re-evaluation of the beta convergence analysis for each of the previously-identified clubs. In contrast to the sigma result, significant beta convergence was found for all clubs. This results means that, within each group, welfare-poor countries tended to grow faster than the welfare-rich countries. Thus, in the absence of random shocks, there was a tendency toward catching up within each convergence cluster.

However, as explained in Quah (1993) and Sala-i Martin (1996), beta convergence by itself is not a sufficient condition for reducing cross-country dispersion (that is, sigma convergence). This phenomenon is consistent with the results shown in Figure 7, where cross-country dispersion within Clubs 1 and 2 increased in spite of their beta convergence results. A final point worth noting is that countries in the poorest welfare club showed the fastest speed of beta convergence (6.46 percent; see Table 2 for details). In this case, however, it is important not to forget that the steady-state equilibrium to which these countries appear to be converging is lower than those of Clubs 2 and 3.

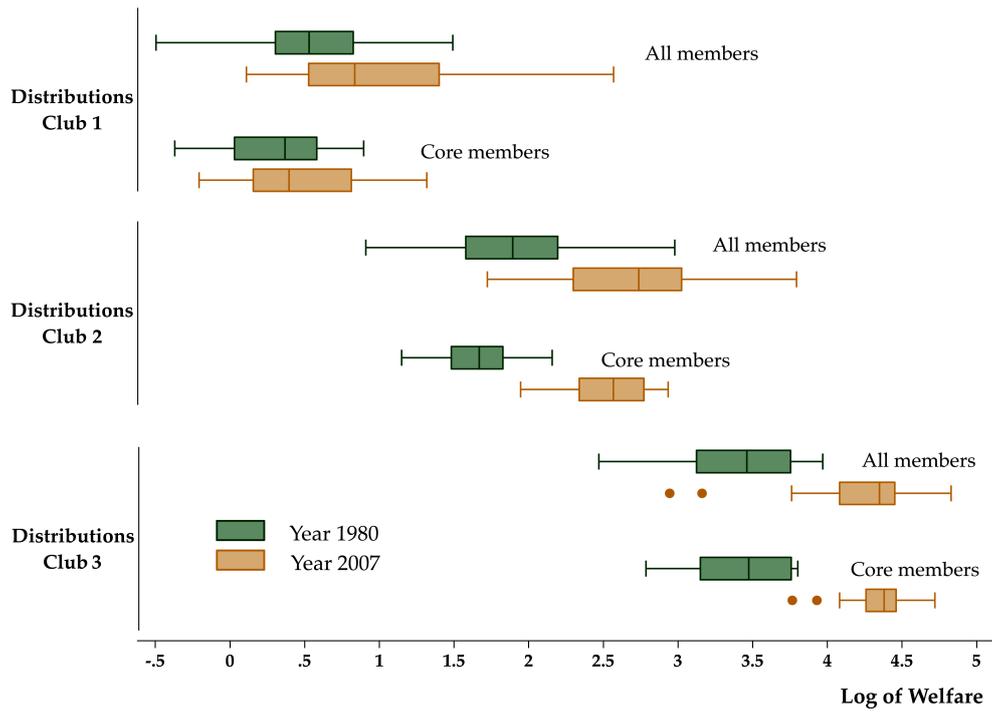


Figure 7. Sigma convergence in welfare within clubs.

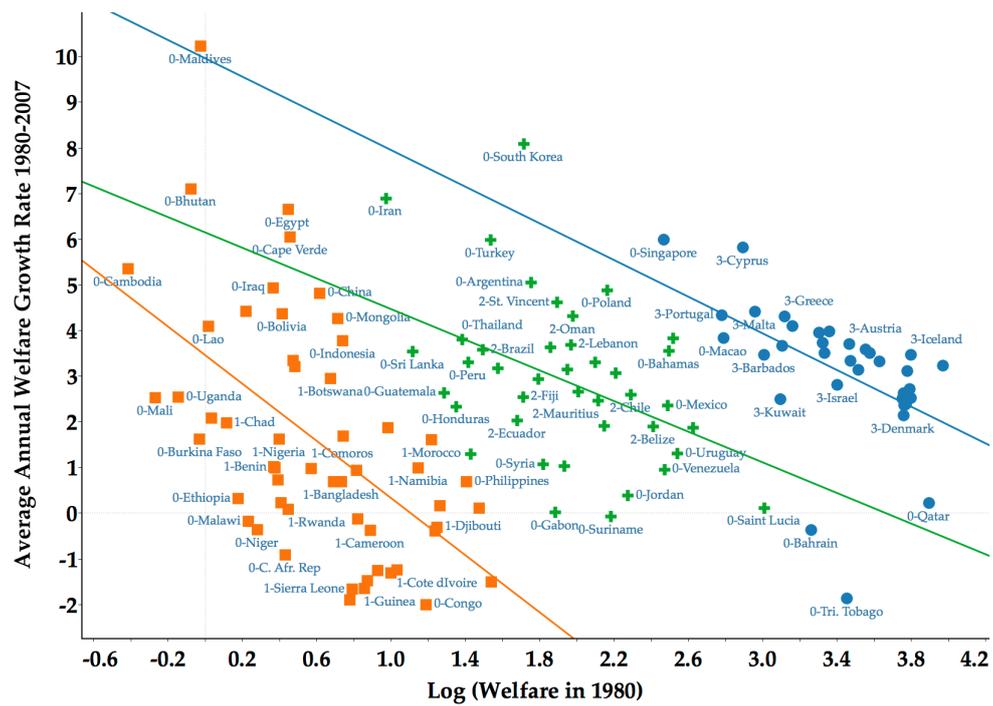


Figure 8. Beta convergence in welfare within clubs.

Table 1. Sigma convergence in welfare within clubs.

	Standard Deviation Log of Relative Welfare 1980	Standard Deviation Log of Relative Welfare 2007	Dispersion Ratio 1980/2007	ANOVA Test <i>p</i>-Value
Total (128 members)	1.22	1.41	0.86	0.10
Club 1 (All 54 members)	0.46	0.58	0.79	0.09
Club 2 (All 40 members)	0.45	0.49	0.92	0.62
Club 3 (All 34 members)	0.37	0.39	0.96	0.80
Club 1 (Core 30 members)	0.31	0.39	0.80	0.23
Club 2 (Core 17 members)	0.25	0.27	0.92	0.75
Club 3 (Core 28 members)	0.32	0.21	1.47	0.05

Table 2. Beta convergence in welfare within clubs.

Club 1			
Variable	Coefficient	<i>t</i> -Statistic	<i>p</i> -Value
constant	0.0337	6.9442	0.0000
log(<i>y</i> ₀)/ <i>T</i>	−0.8252	−4.76	0.0000
R ²	0.3		
Speed of convergence (β)	6.46%		
Half-life (periods)	11		
Club 2			
Variable	Coefficient	<i>t</i> -Statistic	<i>p</i> -Value
constant	0.0599	5.4013	0.0000
log(<i>y</i> ₀)/ <i>T</i>	−0.4388	−2.9187	0.0059
R ²	0.18		
Speed of convergence (β)	2.14%		
Half-life (periods)	32		
Club 3			
Variable	Coefficient	<i>t</i> -Statistic	<i>p</i> -Value
constant	0.0966	4.6296	0.0001
log(<i>y</i> ₀)/ <i>T</i>	−0.5231	−3.1832	0.0032
R ²	0.24		
Speed of convergence (β)	2.74%		
Half-life (periods)	25		

5. Discussion

To contextualize the results of the previous section with some recent findings in the convergence literature, it is important to discriminate between two types of cluster analyses. Some classical studies evaluate the formation of clusters in a static setting (see, for instance, the seminal work of [Berlage and Terweduwe \(1988\)](#)). More recent studies, however, have moved to a more dynamic (panel-data based) setting [Battisti and Parmeter \(2012, 2013\)](#). Static-setting studies tend to evaluate multiple variables at one point in time. For instance, the work of [Berlage and Terweduwe \(1988\)](#) evaluated 20 socioeconomic variables (including GDP) and found that the world can be partitioned into nine clusters. Dynamic-setting studies, however, tend to evaluate single or a few variables over time. For instance, most studies focused on the clustering of GDP per capita. In this context, the work of [Battisti and Parmeter \(2012\)](#) found that the world can be partitioned into three clusters.

The results of the current paper are based on a dynamic-setting clustering framework; and although [Battisti and Parmeter \(2012\)](#) focused on the clustering dynamics of GDP per capita, their clustering framework was relatively similar.⁸ Furthermore, to ensure an additional level of comparability, one must note that the correlation between the welfare index presented in the current paper and GDP per capita is 0.95. Thus, at least from this perspective, the number of clusters identified in the current paper was equal to that in [Battisti and Parmeter \(2012\)](#). There are some noticeable differences, however, in the composition of the clusters. Compared to the high-GDP cluster, there are more European countries in the high-welfare cluster. As noted by [Jones and Klenow \(2016\)](#), the reason for this difference could be that, when compared to the United States (the benchmark country), European countries have a longer life expectancy, additional leisure, and lower inequality. Furthermore, compared to the low-GDP cluster, there are more developing countries in the low-welfare cluster. Similarly, the work in [Jones and Klenow \(2016\)](#) argued that, relative to GDP (and to the

⁸ Although the frameworks are similar, they are not identical. The work in [Battisti and Parmeter \(2012\)](#) used a semi-parametric density-based clustering framework, while this paper uses a non-parametric density-based clustering framework.

benchmark country), a decrease in the welfare ranking of most developing countries is expected given their relatively higher mortality, lower consumption, and higher inequality.

The previous findings point to some promising directions for further research. First, convergence clubs could be identified not only through the lens of a stochastic kernel, but also through an ergodic distribution. In fact, these two approaches are complementary, and together, they provide a more complete characterization of a dynamic system. On the one hand, the stochastic kernel provides a description of transitional dynamics; on the other, the ergodic distribution provides a description of the long-run equilibrium of the system. It would be interesting to know whether the three identified transitional clubs are still observable in a long-run equilibrium. Second, to further test the robustness of the clubs, an alternative convergence-club framework could be used. A close alternative could be the estimation of a finite Gaussian mixture density. In this semi-parametric framework, each component of the mixture can be interpreted as a convergence club. Third, it would be interesting to evaluate the role of geographic neighbors in the formation of clusters. For instance, one could evaluate to what extent the three clusters persist across continents. Finally, an analysis of each component of the welfare index could be useful for identifying the sources of convergence. The identification of the factors driving the formation of welfare clusters is a necessary extension that could provide valuable insights in the design of welfare-enhancing policies.

6. Concluding Remarks

A central topic in long-run macroeconomics is the study of the cross-country convergence hypothesis. This paper tested this hypothesis in a context beyond GDP and through the lens of an unsupervised machine learning algorithm. The novel welfare index of [Jones and Klenow \(2016\)](#) was used as a proxy for the measurement of living standards across countries. From a conceptual standpoint, this welfare index aggregates, in a theoretically-consistent way, measures of consumption (private and public), leisure, life expectancy, and inequality for a sample of 128 countries over the 1980–2007 period. From a methodological standpoint, this paper started with the evaluation of classical summary measures of convergence, such as sigma and beta convergence ([Barro and Sala-i Martin 1992](#)). Next, it applied the non-parametric distributional convergence approach of ([Quah 1997](#)). Finally, it extended this distributional framework by integrating the recent distribution-based clustering algorithm of [Azzalini and Menardi \(2014a\)](#).

The classical analysis of convergence showed that welfare differences across countries are characterized by a lack of both global sigma and global beta convergence. The distributional convergence analysis showed that welfare differences appeared to be characterized by multiple local convergence clubs. In particular, the three modes of the estimated stochastic kernel distribution and the related cluster analysis suggested the existence of three convergence clubs. Under this classification, beta convergence was recovered for each club; moreover, convergence within the poorest club showed the fastest convergence rate. In terms of sigma convergence, however, only the core members of the richest club appeared to be reducing their welfare differences. Overall, these results highlighted a central finding of the long-run macroeconomics literature: beta convergence is necessary, but not sufficient for sigma convergence, even within convergence clubs and in a context beyond GDP.

Funding: This research project was supported by JSPS KAKENHI Grant Number 19K13669.

Acknowledgments: A previous version of this paper was circulated under the title “Convergence Clubs Beyond GDP: A Non-Parametric Distributional Approach”. For further reference, see [Mendez \(2017\)](#).

Conflicts of Interest: The authors declare no conflict of interest.

Appendix A

Table A1. List of countries and clubs.

ID	Name	ISO Code	Clubs	Core Club	Relative Welfare in 1980	Relative Welfare in 2007
1	Luxembourg	LUX	High	NA	52.98	125
2	Iceland	ISL	High	High	44.59	111.7
3	United States	USA	High	High	43.74	100
4	Sweden	SWE	High	High	44.3	91.2
5	France	FRA	High	High	37.72	91.1
6	Australia	AUS	High	High	35.83	90.7
7	United Kingdom	GBR	High	High	34.97	90.4
8	Switzerland	CHE	High	High	44.6	87.1
9	The Netherlands	NLD	High	High	42.88	86.2
10	Austria	AUT	High	High	32.06	85.5
11	Belgium	BEL	High	High	42.61	83
12	Cyprus	CYP	High	High	18.07	83
13	Japan	JPN	High	High	28.8	82.6
14	Canada	CPV	High	High	43.5	82.3
15	Norway	NOR	High	High	43.15	81
16	Italy	ITA	High	High	32.21	78
17	Spain	ESP	High	High	27.26	77.6
18	Germany	DEU	High	High	33.63	77.3
19	Denmark	DNK	High	High	42.91	75.8
20	Finland	FIN	High	High	27.75	74.4
21	New Zealand	NZL	High	High	28.05	71
22	Greece	GRC	High	High	22.62	70.5
23	Ireland	IRL	High	High	23.58	69.6
24	Israel	ISR	High	High	30.08	63.4
25	Malta	MLT	High	High	19.27	61.8
26	Hong Kong	HKG	High	High	22.35	59
27	Singapore	SGP	High	NA	11.82	56.7
28	Qatar	QAT	High	NA	49.23	52.1
29	Barbados	BRB	High	High	20.24	50.7
30	Portugal	PRT	High	High	16.14	50.7
31	South Korea	KOR	Middle	NA	5.56	45.3
32	Macao	MCO	High	NA	16.28	44.9
33	Kuwait	KWT	High	High	22.13	43
34	Hungary	HUN	Middle	NA	12.43	34.2
35	Poland	POL	Middle	NA	8.7	31.5
36	Bahamas	BHS	Middle	NA	12.13	31.1
37	Bahrain	BHR	High	NA	26.16	23.6
38	Costa Rica	CRI	Middle	NA	13.83	22.8
39	Mexico	MEX	Middle	NA	12.07	22.6
40	Oman	OMN	Middle	Middle	7.23	22.6
41	St. Vincent	VCT	Middle	Middle	6.66	22.5
42	Turkey	TUR	Middle	NA	4.65	22.3
43	Argentina	ARG	Middle	NA	5.78	21.8
44	Saint Lucia	LCA	Middle	NA	20.29	20.9
45	Bulgaria	BGR	Middle	Middle	9.13	20.6
46	Chile	CHL	Middle	Middle	9.88	19.7
47	Saudi Arabia	SAU	Middle	Middle	8.16	19.6
48	Trinidad/Tobago	TTO	High	NA	31.63	19
49	Lebanon	LBN	Middle	Middle	7.16	19
50	Belize	BLZ	Middle	Middle	11.16	18.5
51	Uruguay	URY	Middle	NA	12.7	18
52	Albania	ALB	Middle	Middle	6.42	16.8
53	Dominican Rep.	DOM	Middle	Middle	7.03	16.2

Table A1. Cont.

ID	Name	ISO Code	Clubs	Core Club	Relative Welfare in 1980	Relative Welfare in 2007
54	Mauritius	MUS	Middle	Middle	8.3	16
55	Iran	IRN	Middle	NA	2.65	16
56	Venezuela	VEN	Middle	NA	11.85	15.3
57	Malaysia	MYS	Middle	Middle	7.45	15.1
58	Panama	PAN	Middle	Middle	8.58	14.3
59	Maldives	MDV	Low	NA	0.98	13.5
60	Jamaica	JAM	Middle	Middle	6.01	13.1
61	Brazil	BRA	Middle	Middle	4.46	11.5
62	Tunisia	TUN	Middle	Middle	4.84	11.2
63	Fiji	FJI	Middle	Middle	5.54	10.9
64	Thailand	THA	Middle	NA	3.99	10.9
65	Jordan	JOR	Middle	NA	9.75	10.8
66	Peru	PER	Middle	NA	4.12	9.9
67	Ecuador	ECU	Middle	Middle	5.36	9.2
68	Colombia	COL	Middle	NA	6.92	9.1
69	Egypt	EGY	Low	NA	1.56	8.9
70	Suriname	SUR	Middle	NA	8.89	8.7
71	Syria	SYR	Middle	NA	6.17	8.2
72	Sri Lanka	LKA	Middle	NA	3.06	7.8
73	Cape Verde	CAF	Low	NA	1.58	7.7
74	Guatemala	GTM	Middle	NA	3.62	7.3
75	Honduras	HND	Middle	NA	3.87	7.2
76	Gabon	GAB	Middle	NA	6.58	6.6
77	China	CHN	Low	NA	1.86	6.6
78	Mongolia	MNG	Low	NA	2.04	6.3
79	Paraguay	PRY	Middle	NA	4.17	5.9
80	Bhutan	BTN	Low	NA	0.93	5.9
81	Indonesia	IDN	Low	NA	2.1	5.7
82	Iraq	IRQ	Low	NA	1.45	5.3
83	Morocco	MAR	Low	Low	3.39	5.2
84	Philippines	PHL	Low	NA	4.08	4.9
85	Bolivia	BOL	Low	NA	1.52	4.8
86	South Africa	ZAF	Low	NA	4.38	4.5
87	Pakistan	PAK	Low	Low	2.68	4.4
88	Botswana	BWA	Low	Low	1.97	4.3
89	Namibia	NAM	Low	Low	3.15	4.1
90	Vietnam	VNM	Low	NA	1.25	4
91	India	IND	Low	Low	1.61	3.9
92	Sudan	SDN	Low	Low	1.62	3.8
93	Sao Tome/Princi	STP	Low	Low	3.54	3.7
94	Ghana	GHA	Low	Low	2.1	3.3
95	Djibouti	DJI	Low	Low	3.48	3.2
96	Swaziland	SWZ	Low	NA	4.67	3.1
97	Zimbabwe	ZWE	Low	Low	3.44	3.1
98	Lao	LAO	Low	NA	1.02	3
99	Mauritania	MRT	Low	Low	2.26	2.9
100	Cambodia	CMR	Low	NA	0.66	2.7
101	Bangladesh	BGD	Low	Low	2.08	2.5
102	Senegal	SEN	Low	Low	2	2.4
103	Comoros	COM	Low	Low	1.77	2.3
104	Nigeria	NGA	Low	Low	1.49	2.3
105	Cameroon	CAN	Low	Low	2.44	2.2
106	Lesotho	LSO	Low	Low	2.28	2.2
107	Cote d'Ivoire	CIV	Low	Low	2.81	2
108	Congo	COD	Low	NA	3.29	1.9
109	Kenya	KEN	Low	Low	2.72	1.9
110	Benin	BEN	Low	Low	1.46	1.9

Table A1. Cont.

ID	Name	ISO Code	Clubs	Core Club	Relative Welfare in 1980	Relative Welfare in 2007
111	Angola	AGO	Low	Low	1.45	1.9
112	Chad	TCD	Low	Low	1.12	1.9
113	Zambia	ZMB	Low	Low	2.53	1.8
114	Togo	TGO	Low	Low	1.48	1.8
115	Nepal	NPL	Low	NA	1.04	1.8
116	Uganda	UGA	Low	NA	0.87	1.7
117	Tanzania	TZA	Low	Low	2.4	1.6
118	Rwanda	RWA	Low	Low	1.57	1.6
119	Madagascar	MDG	Low	Low	1.51	1.6
120	Guinea	GIN	Low	Low	2.36	1.5
121	Burkina Faso	BFA	Low	NA	0.97	1.5
122	Mali	MLI	Low	NA	0.77	1.5
123	Sierra Leone	SLE	Low	Low	2.21	1.4
124	Liberia	LBR	Low	Low	2.18	1.3
125	Ethiopia	ETH	Low	NA	1.2	1.3
126	C. Afr. Republic	KHM	Low	NA	1.54	1.2
127	Niger	NER	Low	NA	1.33	1.2
128	Malawi	MWI	Low	NA	1.26	1.2

References

- Athey, Susan. 2018. The impact of machine learning on economics. In *The Economics of Artificial Intelligence: An Agenda*. Chicago: University of Chicago Press.
- Athey, Susan, and Guido W. Imbens. 2019. Machine learning methods that economists should know about. *Annual Review of Economics* 11: 685–725. [\[CrossRef\]](#)
- Azzalini, Adelchi, and Giovanna Menardi. 2014a. An advancement in clustering via nonparametric density estimation. *Statistics and Computing* 24: 753–67.
- Azzalini, Adelchi, and Giovanna Menardi. 2014b. Clustering via nonparametric density estimation: The r package pdfcluster. *Journal of Statistical Software* 57: 1–26. [\[CrossRef\]](#)
- Azzalini, Adelchi, and Nicola Torelli. 2007. Clustering via nonparametric density estimation. *Statistics and Computing* 17: 71–80. [\[CrossRef\]](#)
- Barro, Robert J., and Xavier Sala-i Martin. 1992. Convergence. *Journal of Political Economy* 100: 223–51. [\[CrossRef\]](#)
- Battisti, Michele, and Christopher F. Parmeter. 2012. Gdp clustering: A reappraisal. *Economics Letters* 117: 837–40. [\[CrossRef\]](#)
- Battisti, Michele, and Christopher F. Parmeter. 2013. Clustering and polarization in the distribution of output: A multivariate perspective. *Journal of Macroeconomics* 35: 144–62. [\[CrossRef\]](#)
- Baumol, William J. 1986. Productivity growth, convergence, and welfare: What the long-run data show. *The American Economic Review* 76: 1072–85. [\[CrossRef\]](#)
- Becker, Gary S., Tomas J. Philipson, and Rodrigo R. Soares. 2005. The quantity and quality of life and the evolution of world inequality. *American Economic Review* 95: 277–91. [\[CrossRef\]](#)
- Berlage, Lodewijk, and Dirk Terweduwe. 1988. The classification of countries by cluster and by factor analysis. *World Development* 16: 1527–45. [\[CrossRef\]](#)
- Cordoba, Juan Carlos, and Genevieve Verdier. 2008. Inequality and growth: Some welfare calculations. *Journal of Economic Dynamics and Control* 32: 1812–29. [\[CrossRef\]](#)
- Dal Bianco, Silvia. 2016. Going clubbing in the eighties: Convergence in manufacturing sectors at a glance. *Empirical Economics* 50: 623–59. [\[CrossRef\]](#)
- Dowrick, Steve, and Duc-Tho . T. Nguyen. 1989. Oecd comparative economic growth 1950–1985: Catch-up and convergence. *American Economic Review* 79: 1010–30.
- Durlauf, Steven N., Paul A. Johnson, and Jonathan R. W. Temple. 2005. Growth econometrics. *Handbook of Economic Growth*. Amsterdam: North Holland.
- Epstein, Philip, Peter Howlett, and Max-Stephan Schulze. 2003. Distribution dynamics: Stratification, polarization, and convergence among oecd economies, 1870–1992. *Explorations in Economic History* 40: 78–97. [\[CrossRef\]](#)

- Fleurbaey, Marc. 2009. Beyond gdp: The quest for a measure of social welfare. *Journal of Economic Literature* 47: 1029–75. [\[CrossRef\]](#)
- Fleurbaey, Marc, and Didier Blanchet. 2013. *Beyond GDP: Measuring Welfare and Assessing Sustainability*. Oxford: Oxford University Press.
- Fleurbaey, Marc, and Guillaume Gaulier. 2009. International comparisons of living standards by equivalent incomes. *Scandinavian Journal of Economics* 111: 597–624. [\[CrossRef\]](#)
- Galor, Oded. 1996. Convergence? Inferences from theoretical models. *Economic Journal* 106: 1056–69. [\[CrossRef\]](#)
- Hartigan, John A. 1975. *Clustering Algorithms*. New York: John Wiley and Sons.
- Johnson, Paul, and Chris Papageorgiou. 2018. *What Remains of Cross-Country Convergence?* MPRA Working Paper 89355. Munich: University Library of Munich.
- Jones, Charles I., and Peter J. Klenow. 2016. Beyond gdp? Welfare across countries and time. *American Economic Review* 106: 2426–57. [\[CrossRef\]](#)
- Kaufman, Leonard, and Peter J. Rousseeuw. 2010. *Finding Groups in Data: An Introduction to Cluster Analysis*. New York: John Wiley and Sons.
- Magrini, Stefano. 2009. Why should we analyse convergence using the distribution dynamics approach? *Scienze Regionali* 8: 5–34.
- Mendez, Carlos. 2017. *Convergence Clubs Beyond GDP: A Non-Parametric Density Approach* MPRA Working Paper 82048. Munich: University Library of Munich.
- Mendez, Carlos. 2018. On the distribution dynamics of human development: Evidence from the metropolitan regions of bolivia. *Economics Bulletin* 38: 2467–75.
- Pittau, Maria Grazia, Roberto Zelli, and Riccardo Massari. 2016. Evidence of convergence clubs using mixture models. *Econometric Reviews* 35: 1317–42. [\[CrossRef\]](#)
- Quah, Danny. 1993. Galton's fallacy and tests of the convergence hypothesis. *The Scandinavian Journal of Economics* 95: 427–43. [\[CrossRef\]](#)
- Quah, Danny T. 1996a. Empirics for economic growth and convergence. *European Economic Review* 40: 1353–75. [\[CrossRef\]](#)
- Quah, Danny T. 1996b. Twin peaks: Growth and convergence in models of distribution dynamics. *Economic Journal* 106: 1045–55. [\[CrossRef\]](#)
- Quah, Danny T. 1997. Empirics for growth and distribution: Stratification, polarization, and convergence clubs. *Journal of Economic Growth* 2: 27–59. [1009781613339. \[CrossRef\]](#)
- Sala-i Martin, Xavier. 1996. The classical approach to convergence analysis. *Economic Journal* 106: 1019–36. [\[CrossRef\]](#)
- Stiglitz, Joseph E., Jean-Paul Fitoussi, and Martine Durand. 2019. *Measuring What Counts: A New Dashboard for Well-Being*. New York: The New Press.
- Stiglitz, Joseph E., Amartya Sen, and Jean-Paul Fitoussi. 2010. *Mismeasuring Our Lives: Why GDP Doesn't Add Up*. New York: The New Press.



© 2019 by the authors. 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 (<http://creativecommons.org/licenses/by/4.0/>).