Convergence Clubs Beyond GDP: A Non-Parametric Density Approach

Mendez-Guerra, Carlos

Kyushu University

18 October 2017

Convergence Clubs Beyond GDP: A Non-Parametric Density Approach

Carlos Mendez-Guerra*

Kyushu University

October 18, 2017

Abstract

In the study of convergence in living standards across countries, per-capita Gross Domestic Product (GDP) has been usually used as a proxy for the measurement of national well-being. However, other important welfare aspects—beyond GDP—should also be considered. This paper revisits the cross-country convergence hypothesis in a context beyond GDP. It focuses on a novel welfare index [Jones and Klenow (2016), American Economic Review, 106(9)] that incorporates measures of consumption, leisure, life expectancy, and inequality. Based on a sample of 128 countries over the 1980-2007 period, the paper first documents the lack of overall sigma and (absolute) beta convergence. Next, through the lens of a stochastic kernel density and a clustering algorithm, it documents the formation of three convergence clubs. Under this classification, the beta convergence coefficient 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 convergenceeven within convergence clubs and in a context beyond GDP.

*Contact Email: carlosmendez777@gmail.com

1. Introduction

A central topic in the study of economic growth and development is the empirical testing of the convergence hypothesis across countries. 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], economist are well aware that it is an incomplete measure of national welfare. In this context, a series of different alternatives have been proposed by economists and non-economists alike. 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 combine measures of national consumption (private and public), leisure, life expectancy, and inequality to construct a theoretically appealing new welfare index. Given the cross-country and time coverage of the index, it can be used as an alternative proxy variable for the study of the convergence hypothesis.

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 some classical tests of convergence (sigma and beta convergence). Then, its fundamental contribution relies on the implementation of a non-parametric density convergence approach (Quah, 1993, 1996, 1997; Magrini, 2009). Furthermore, it extends this density convergence framework by integrating a novel clustering algorithm (Azzalini and Menardi 2014a, 2014b) that allows the 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. Both the lack of overall convergence and the limited country mobility suggest the possible existence of convergence clubs. Moreover, the multiple modes of the estimated stochastic kernel density suggests that welfare differences appear to be characterized by three clubs. Under this classification, the beta convergence coefficient 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 reemphasize a central finding in the economic growth and development literature: beta con-

[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.

vergence 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 beyond GDP data and the different convergence frameworks. Section 3 introduces some stylized convergence facts about national welfare differences. Section 4 presents the convergence clubs results. Finally, Section 5 offers some concluding remarks with extended suggestions for further research.

2. Data and Methods

2.1 Beyond GDP Data

Jones and Klenow (2016) propose a summary statistic that aims to quantify the level of welfare of people in a country. This novel statistic incorporates measures of consumption (private and public), leisure, life expectancy, and inequality. To aggregate these measures, in the theoretically consistent way, they calibrate an expected utility function and evaluate the joint consumption-equivalent level of the four variables. Using this statistic these authors find that, on average, welfare is highly correlated with GDP per capita. However, they also report that there are often large deviations from GDP, in particular in developing countries. 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.

Jones and Klenow (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 subsample of 128 countries[2]. Given the welfare level data for the final year (2007) and the average growth rate (1980-2007 period), it is possible to compute the cross-sectional data for the initial year (1980). As a result of this calculation, this paper uses a sample of 128 countries for two time periods: 1980 and 2007.

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 time evolution of the cross-sectional dispersion of a variable. From this perspective, convergence

[2]The database can be accessed from https://web.stanford.edu/~chadj/BeyondGDP500.xls .

occurs when the cross-sectional dispersion declines over time, so the level of the variable under study becomes increasingly more similar across the countries of the world. (Baumol, 1986; Dowrick and Nguyen, 1989; Sala-i-Martin, 1996). Typically, sigma convergence is measured by the coefficient of variation or the standard deviation of the logarithm of a variable. In this research paper, the latter indicator is adopted and the measurement of sigma convergence is implemented as follows:

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

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 (absolute) beta, β, convergence describes the inverse relationship between the initial level of a variable and its average growth rate. From this perspective, if such 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 (Baumol, 1986; Barro and Sala-i-Martin, 1992; 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 this research paper, the measurement of beta convergence is implemented as follows:

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

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 related[3]. Keeping other variables constant, when poor countries tend to grow faster than rich ones (i.e., beta convergence occurs), the cross-country income or welfare dispersion declines over time (i.e., sigma convergence occurs). 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. As explained by Quah (1993) and Sala-i-Martin (1996), beta convergence is necessary but not sufficient to achieve sigma convergence.

[3] See Sala-i-Martin (1996) for further details.

2.3 Density Convergence and Clubbing

This density approach to convergence aims to capture the time 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 have only one mode over time. In this framework, the emergence of multiple modes is usually associated with existence of convergence clubs (Galor, 1996; Quah, 1993, 1996, 1997; Magrini, 2009). Typically, density convergence is measured by the shape and the number of modes of a stochastic kernel density and its corresponding ergodic density. In this research paper, only the former is adopted and the measurement of density convergence is implemented as follows:

- The variable under study (i.e., national welfare) is expressed relative to a benchmark economy (in this paper it is the United States). This normalization allows abstraction from forces that might simultaneously affect all countries.
- To facilitate comparison[4] and visualization, the natural logarithm of the relative variable is applied.
- The bivariate stochastic kernel is a conditional density that is calculated as follows:

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

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

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

  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 (AMISE[5])

The bivariate stochastic kernel is a tree dimensional object that is typically represented by

[4]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).

[5]See Johnson (2005) and Magrini (2007, 2009) and for further details.

a surface plot or a contour plot. If most countries are concentrated around the main diagonal of any of these graphs, then there is evidence of distributional persistence over time. Density convergence is found when most of the countries are located around zero in the the $(t + s)$-axis and parallel to the t-axis.

Finally, the density-based clustering algorithm developed by Azzalini and Menardi (2014a, 2014b) is applied with the previously described kernel functions and smoothing parameters. The implementation of this new clustering framework is useful for two purposes. First, to identify the location of each country under the stochastic kernel. Second, to 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 showed that, similar to the case of GDP per capita, welfare differences across countries are characterized by a lack of both sigma and (absolute) beta convergence. Figures 1 and 2 summarize this finding. Figure 1 shows that, although the distribution has moved in the right direction, 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. The standard deviation of the cross-country welfare distribution in the year 1980 was 1.22. By the year 2007, however, 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 Figure 1.

Figure 2 shows that, on average, welfare-poor countries are not growing faster than welfare-rich countries. As such, this result highlights the lack of beta convergence in the context of welfare differences across countries. Also note 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 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.

Figure 1: Lack of Sigma Convergence in Welfare Differences Across Countries

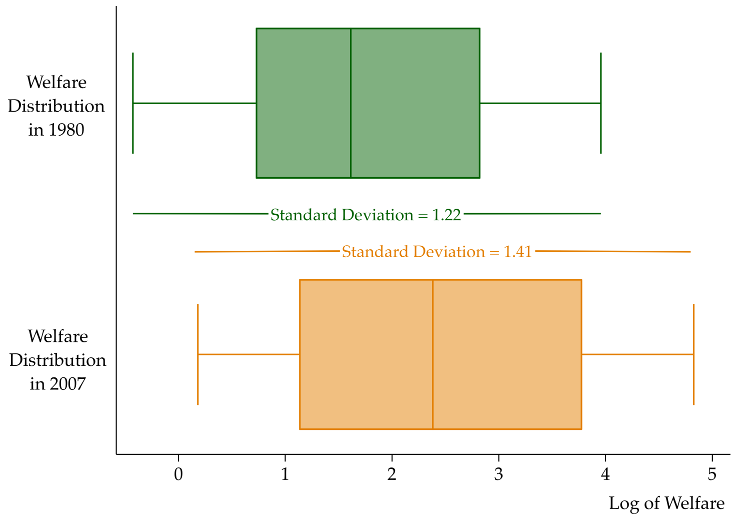
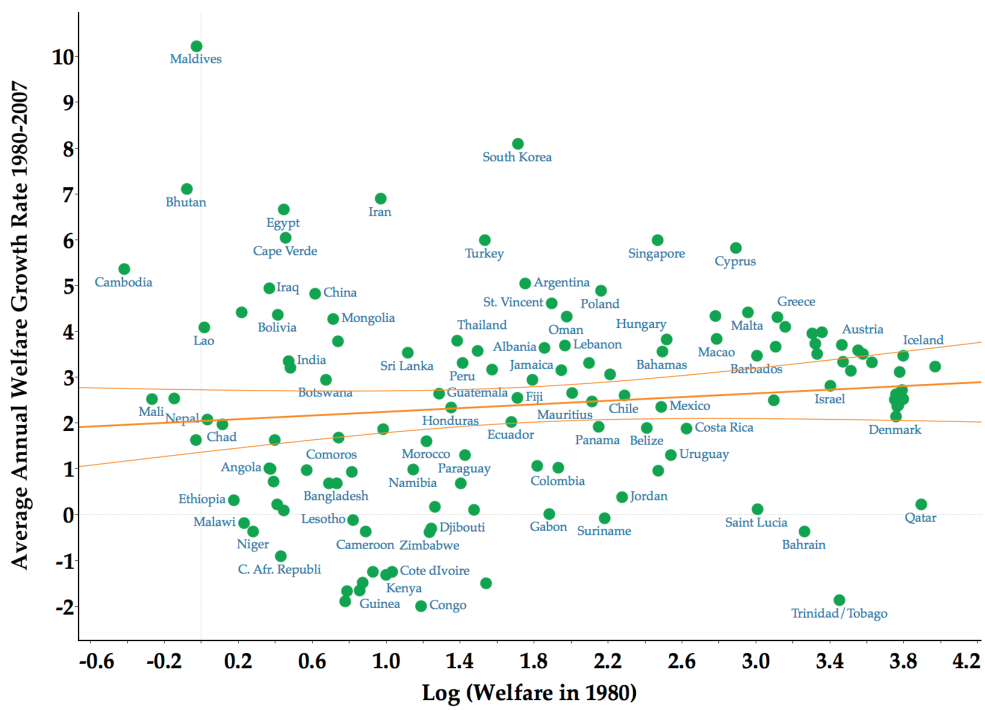


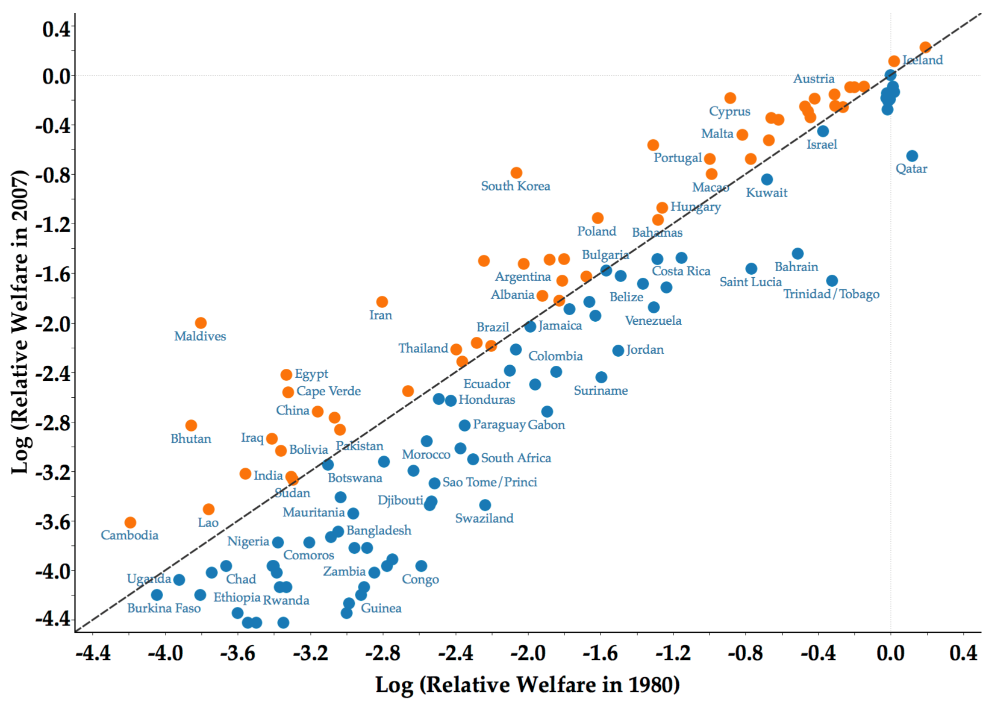
Figure 2: Lack of Beta Convergence in Welfare Differences Across Countries



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[6], since most countries are located around the 45 degree line. Indeed, the relation between the initial and the final (relative) level of welfare is summarized by a linear regression in which the coefficient of the slope is statistically equal to one[7]. 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 Welfare Mobility Across Countries



Another interesting mobility finding is that 76 out the 128 countries in the sample (i.e., almost 60 percent of the total sample) are moving backwards[8]. This divergence process is much more pronounced in the poorest countries in the sample. Finally, both the limited degree 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

[6]Note that the measurement of country mobility here is relative to that experienced by the frontier, which in this case is the United States.

[7]The 95% confidence interval for this slope coefficient is between 0.97 and 1.14.

[8]Note that a backward movement in relative terms, does not necessary imply that the country is experiencing less welfare in absolute terms. Even a country that improves its absolute level of welfare can move backwards in relative terms when the technological frontier (the United States in this case) moves forward at a higher speed.

presented in the next section.

4. Convergence Clubs Results

4.1 Stochastic Kernel Density

Figure 4 shows the 3D surface of the stochastic kernel density. The transitional dynamics between 1980 and 2007 are 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 (i.e., density modes) in the main diagonal are typically interpreted as suggestive evidence about the formation of convergence clubs. In this case, given the existence of significant three modes, then three different convergence clubs are likely to be present. Interestingly, the existence of tree convergence clubs in welfare is qualitatively consistent with the three convergence clubs in income reported in the recent work of Pittau, Zelli, and Massari (2016)[9].

Figure 4: Stochastic Kernel Density and Convergence Clubs (3D Surface Plot)

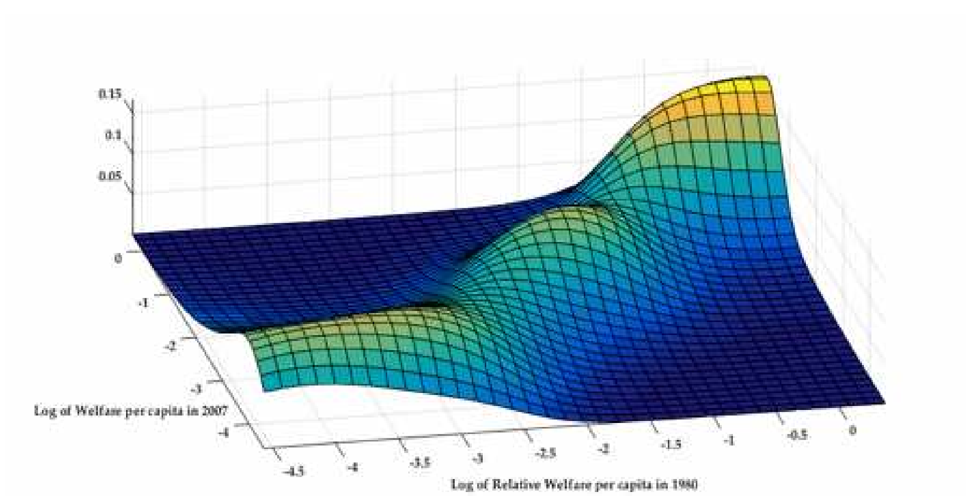
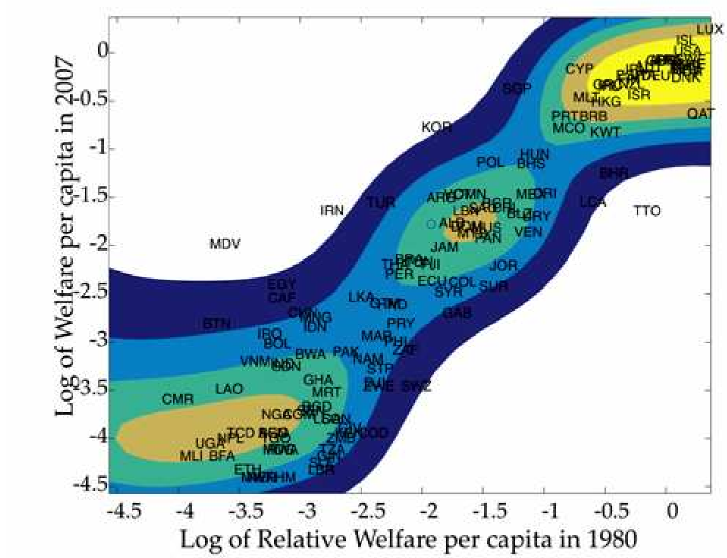


Figure 5 shows the contour plot representation stochastic kernel overlapped with a scatter plot of countries. This figure is a first attempt to understand the country composition

[9] Although qualitatively the number of clubs is the same, the methodology of Pittau, Zelli, and Massari (2016). They use a finite Gaussian mixture model to evaluate the existence of convergence clubs. This methodological approach is a semi-parametric alternative to the non-parametric framework implemented in this paper.

of each convergence club. From the figure, it clear that some countries are located at the center-core of each club. However, the position of other countries raise some doubts about their club membership. In particular, there seems to be a considerable number of countries between the bottom club and the middle club. For this subsample, some countries are moving forward and they might be transitioning towards a superior club. Other countries are moving backward and they might be transitioning towards an inferior club. To help clarify the membership of these kind of countries, a novel density-based clustering algorithm.

Figure 5: Stochastic Kernel Density and Convergence Clubs (Contour Plot)

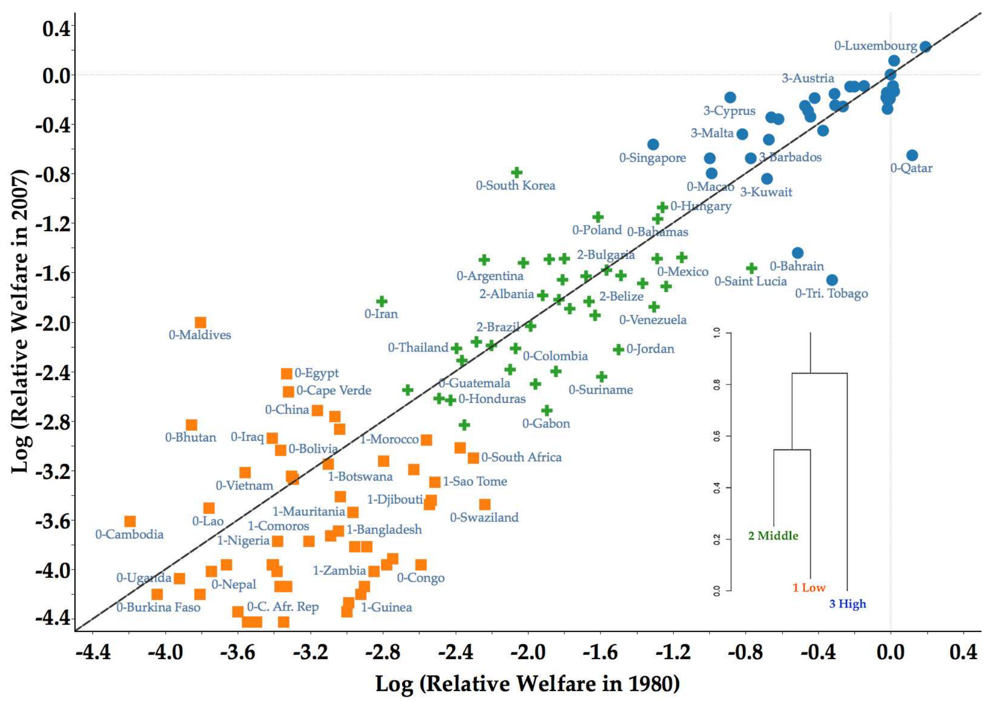


4.2 Core Clusters and Classification

In a series of papers, Azzalini and Torelli (2007) and Azzalini and Menardi (2014a, 2014b) have used the modes of a non-parametric kernel density as a criteria for identifying clusters. High-density observations group themselves in what Azzalini and Menardi call "cluster cores". Next, for the low-density observations, they implement a Delaunay triangulation to allocate those observations to their nearest club.

Figure 6 shows the results of the application of this clustering algorithm to the scatter plot of welfare mobility of Figures 3 and 5. Consistent with the stochastic kernel of Figure 4 and 5, three clusters or clubs are identified. The main advantage of this clustering framework is the endogenous identification of core club members (i.e., those with a "1", "2", or "3" prefix). In addition, low-density observations (i.e., those with a "0" prefix) are also identified and allocated to their more proximate core club. In this low-density context, cases such as Maldives or South Korea are interesting examples to be studied. Although both countries experienced relatively large forward mobility, the progress they made is not enough to be classified as a "core" member of the immediately superior club. In these cases, the clustering classification is still informative in the sense that it is also clear that these kind of countries do not belong to their immediately inferior "core" club either.

Figure 6: Density Based Clustering



Finally, the bottom right of Figure 56includes a cluster tree. It is meant to provide a measure of the robustness of the clubs to different density thresholds. Although for a considerable large set of density thresholds three clubs are identifiable, it is also possible that the countries of club 1 and club 2 could be converging to a similar steady state.

However, even in this case, the cross-country distribution of welfare is more likely to be characterized by more than one convergence club.

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 is only present for the core members of the high-welfare club. In other words, only the core members of the club 3 are increasingly becoming more similar in terms of its 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. For the other country clubs, 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 show the largest increase in dispersion.

Figure 7: (Lack of) Sigma Convergence in Welfare Differences within Country Clubs

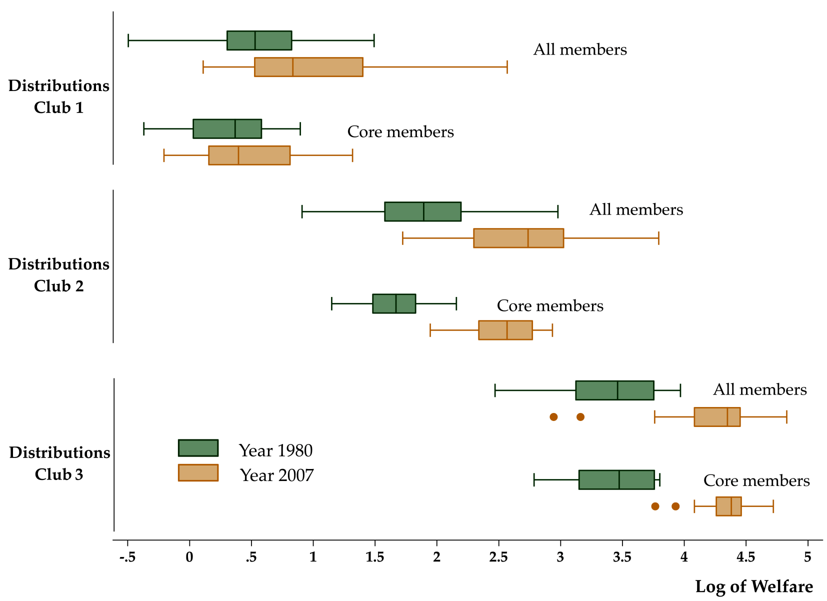


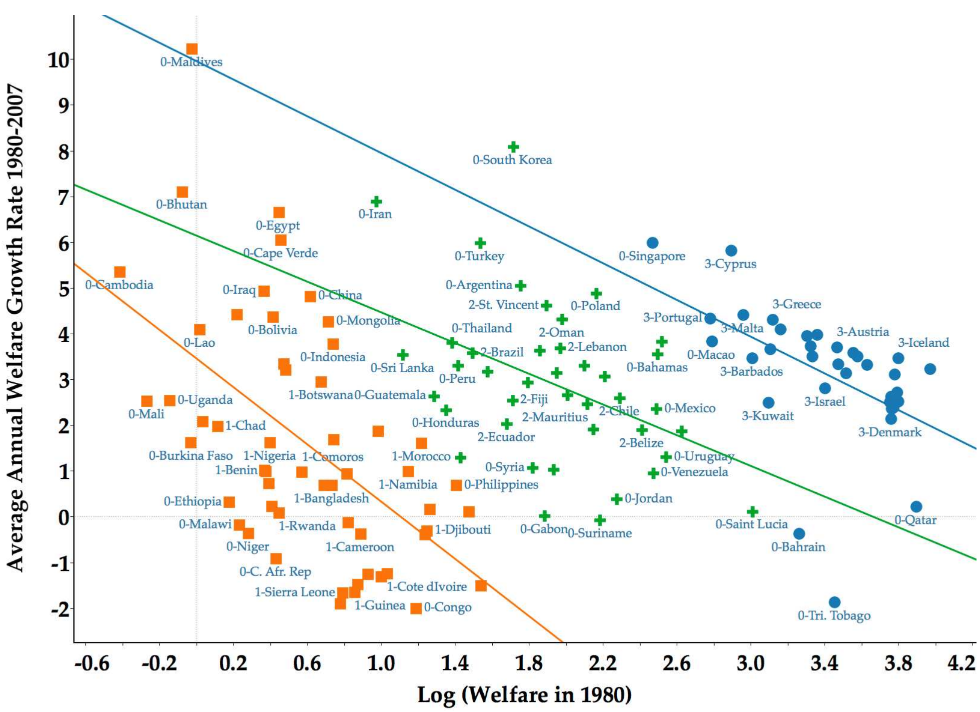
Figure 8 and Table 2 present a re-evaluation of the beta convergence test for each of the previously identified clubs. In contrast to the previous result, significant beta convergence is found for all clubs. This results means that welfare-poor countries—within each

Table 1: (Lack of) Sigma Convergence in Welfare Differences within Country Clubs

	Standard Deviation	
	Year 1980	Year 2007
Total (128 members)	1.22	1.41
Club 1 (All 54 members)	0.46	0.58
Club 2 (All 40 members)	0.45	0.49
Club 3 (All 34 members)	0.37	0.39
Club 1 (Core 30 members)	0.31	0.39
Club 2 (Core 17 members)	0.25	0.27
Club 3 (Core 28 members)	0.32	0.21

group—tend to grow faster than the welfare-rich countries (of the same group), so over time welfare-poor countries could catch up with the level of the rich ones.

Figure 8: Beta Convergence in Welfare Differences within Country Clubs



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 (i.e., sigma convergence). This phenomenon is consistent with the results shown in Figure 6, 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 show 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 the clubs 2 and 3.

Table 2: Beta Convergence in Welfare Differences within Country Clubs

Club 1			
Variable	Coefficient	t-statistic	p-value
constant	0.0337	6.9442	0.0000
log(y0)/T	-0.8252	-4.76	0.0000
R2	0.3		
Speed of convergence (β)	6.46%		
Half-life (periods)	11		

Club 2			
Variable	Coefficient	t-statistic	p-value
constant	0.0599	5.4013	0.0000
log(y0)/T	-0.4388	-2.9187	0.0059
R2	0.18		
Speed of convergence (β)	2.14%		
Half-life (periods)	32		

Club 3			
Variable	Coefficient	t-statistic	p-value
constant	0.0966	4.6296	0.0001
log(y0)/T	-0.5231	-3.1832	0.0032
R2	0.24		
Speed of convergence (β)	2.74%		
Half-life (periods)	25		

5. Concluding Remarks

This paper has tested the cross-country convergence hypothesis in a context beyond GDP. A novel index of national welfare has been used as a proxy for the measurement of living standards across countries. 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 has applied classical summary measures of convergence, such as sigma and beta convergence. Its fundamental contribution, however, relies on the implementation of a non-parametric density convergence approach. Furthermore, this paper has extended this density convergence framework by integrating a novel density-based clustering algorithm developed in the statistical computing literature.

The classical analysis of convergence showed that welfare differences across countries are characterized by a lack of both sigma and (absolute) beta convergence. Moreover, there is a limited degree of forward and backward mobility across countries, in particular in the welfare-rich countries. Together the lack of overall convergence and limited mobility suggest the possible existence of convergence clubs.

The density analysis of convergence showed that welfare differences appear to be characterized by different convergence clubs. In particular, the modes of the estimated stochastic kernel density and the related cluster analysis suggest the existence of three convergence clubs. Under this classification, a significant (absolute) beta convergence coefficient is recovered for each club; moreover, convergence within the poorest club shows the fastest convergence rate. In terms of sigma convergence, however, only the core members of the richest club appear to be reducing their welfare differences. Overall, these results highlight a central finding in the economic growth and development literature: beta convergence is necessary but not sufficient for sigma convergence—even within convergence clubs and in a context beyond GDP.

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 density, but also through an ergodic kernel density. 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 density provides a description of the transitional dynamics of the system; on the other, the ergodic density 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. Finally, analysis of each component of the welfare index could be useful for identifying the sources of convergence and divergence.

References

Azzalini, A. and Menardi, G. (2014a). An advancement in clustering via nonparametric density estimation. *Statistics and Computing*, 24(5):753–767.

Azzalini, A. and Menardi, G. (2014b). Clustering via nonparametric density estimation: The r package pdfcluster. *Journal of Statistical Software*, 57(11):1–26.

Azzalini, A. and Torelli, N. (2007). Clustering via nonparametric density estimation. *Statistics and Computing*, 17(1):71–80.

Barro, R. J. and Sala-i Martin, X. (1992). Convergence. *Journal of Political Economy*, 100(2):223–251.

Barro, R. T. and Sala-I-Martin, X. (1992). Regional growth and migration: A japan-united states comparison. *Journal of the Japanese and International Economies*, 6(4):312–346.

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

Dowrick, S. and Nguyen, D.-T. . T. (1989). Oecd comparative economic growth 1950-85: catch-up and convergence. *American Economic Review*, 79(5):1010–1030.

Galor, O. (1996). Convergence? inferences from theoretical models. *Economic Journal*, pages 1056–1069.

Johnson, P. A. (2005). A continuous state space approach to convergence by parts. *Economics Letters*, 86(3):317–321.

Jones, C. I. and Klenow, P. J. (2016). Beyond gdp? welfare across countries and time. *American Economic Review*, 106(9):2426–57.

Magrini, S. (2009). Why should we analyse convergence using the distribution dynamics approach? *Scienze Regionali*, 8(1):5–34.

Pittau, M. G., Zelli, R., and Massari, R. (2016). Evidence of convergence clubs using mixture models. *Econometric Reviews*, 35(7):1317–1342.

Quah, D. (1993). Galton's fallacy and tests of the convergence hypothesis. *The Scandinavian Journal of Economics*, pages 427–443.

Quah, D. T. (1996a). Empirics for economic growth and convergence. *European economic review*, 40(6):1353–1375.

Quah, D. T. (1996b). Twin peaks: growth and convergence in models of distribution dynamics. *Economic Journal*, pages 1045–1055.

Quah, D. T. (1997). Empirics for growth and distribution: Stratification, polarization, and convergence clubs. *Journal of Economic Growth*, 2(1):27–59.

Sala-i Martin, X. (1996). The classical approach to convergence analysis. *Economic journal*, pages 1019–1036.