



WP-EC 2020-01

A simple non-parametric catching-up and convergence index

Francisco J. Goerlich

Ivie

Working papers
Working papers
Working papers

Los documentos de trabajo del Ivie ofrecen un avance de los resultados de las investigaciones económicas en curso, con objeto de generar un proceso de discusión previo a su remisión a las revistas científicas. Al publicar este documento de trabajo, el Ivie no asume responsabilidad sobre su contenido.

Ivie working papers offer in advance the results of economic research under way in order to encourage a discussion process before sending them to scientific journals for their final publication. Ivie's decision to publish this working paper does not imply any responsibility for its content.

La edición y difusión de los documentos de trabajo del Ivie es una actividad subvencionada por la Generalitat Valenciana, Conselleria de Hacienda y Modelo Económico, en el marco del convenio de colaboración para la promoción y consolidación de las actividades de investigación económica básica y aplicada del Ivie.

The editing and dissemination process of Ivie working papers is funded by the Valencian Regional Government's Ministry for Finance and the Economic Model, through the cooperation agreement signed between both institutions to promote and consolidate the Ivie's basic and applied economic research activities.

La Serie EC, coordinada por Matilde Mas, está orientada a la aplicación de distintos instrumentos de análisis al estudio de problemas económicos concretos.

Coordinated by Matilde Mas, the EC Series mainly includes applications of different analytical tools to the study of specific economic problems.

Todos los documentos de trabajo están disponibles de forma gratuita en la web del Ivie <http://www.ivie.es>, así como las instrucciones para los autores que desean publicar en nuestras series.

Working papers can be downloaded free of charge from the Ivie website <http://www.ivie.es>, as well as the instructions for authors who are interested in publishing in our series.

Versión: enero 2020 / Version: January 2020

Edita / Published by:
Instituto Valenciano de Investigaciones Económicas, S.A.
C/ Guardia Civil, 22 esc. 2 1º - 46020 Valencia (Spain)

DOI: <http://dx.medra.org/10.12842/WPASEC-2020-01>

WP-EC 2020-01

A simple non-parametric catching-up and convergence index

Francisco J. Goerlich *

Abstract

This paper presents a simple and intuitive index that attempts to summarize two non-parametric kernel density estimators in order to quantify the degree of convergence between both distributions. Its derivation follows a similar argument as the Gini index, in the sense that it can be expressed as a ratio of areas under the corresponding densities. The index can be routinely computed once kernel density estimators have been obtained, and it is bounded between 0, absence of convergence, and 1, full convergence. In this way it complements numerically the common graphical device of inspecting non-parametric kernel density estimators in exploratory data analysis.

Keywords: Catching-up, Convergence, Distribution dynamics, Income distribution.

JEL classification: D32, O47.

Resumen

Este artículo presenta un índice simple e intuitivo que intenta resumir dos estimadores de densidad de kernel no-paramétricos para cuantificar el grado de convergencia entre ambas distribuciones. Su derivación sigue un argumento similar al del índice de Gini, en el sentido de que puede expresarse como una relación de áreas bajo las densidades correspondientes. El índice puede calcularse rutinariamente una vez que se han obtenido dichos estimadores de densidad, y está acotado entre 0, ausencia de convergencia y 1, convergencia completa. De esta forma, complementa numéricamente el recurso común de inspección visual de los estimadores de densidad kernel no-paramétricos en el análisis exploratorio de datos.

Palabras clave: *Catching-up*, convergencia, dinámica de las distribuciones, distribución de la renta.

Clasificación JEL: D32, O47.

* University of Valencia and Instituto Valenciano de Investigaciones Económicas (Ivie).

Correspondence to: Francisco J. Goerlich Gisbert, University of Valencia, Department of Economic Analysis, Campus de Tarongers, Av de Tarongers s/n, 46022-Valencia, Spain. E-mail: Francisco.J.Goerlich@uv.es. Tel.: +34 963 190 050.

Acknowledgements: The author acknowledges feedback from Lorenzo Serrano and financial support by the Spanish Ministry of Science and Technology, ECO2015-70632-R project, and the BBVA Foundation-Ivie research programme. Remaining errors are mine. An R function performing the calculations, Shiny R files for an interactive presentation and results mentioned, but not shown, are available from the author upon request.

1. Introduction.

At the end of the previous century a vast empirical literature developed around the concept of economic convergence. After the seminal contributions of Barro and Sala-i-Martin (1992, 1998), which developed the concepts of β and σ convergence, many authors introduced alternative metrics focusing on different aspects of convergence. Some of them came from the income distribution literature (Dagum 1980, Shorrocks 1982, Ebert 1984, Yitzhaki and Lerman 1991, Bishop, Formby and Thistle 1992, Yitzhaki 1994), and were developed within the context of measuring the distance between economic distributions, but others were introduced from different branches of statistics (Quah 1993, 1996a, 1996b).

Quah (1997) criticizes that β and σ convergences only pay attention to a single aspect of convergence, essentially related to the mean of the distribution, and proposed alternative instruments focusing on the entire distribution. Among these alternative measures were the non-parametric kernel density estimators, which have become very popular in applied studies since their introduction. Hence, inspecting non-parametric kernel density estimators has been a common way to do exploratory data analysis. However, this new metric is essentially graphical, and so convergence is checked visually by examining the temporal evolution of the graphs, and no single summary index comes from these.

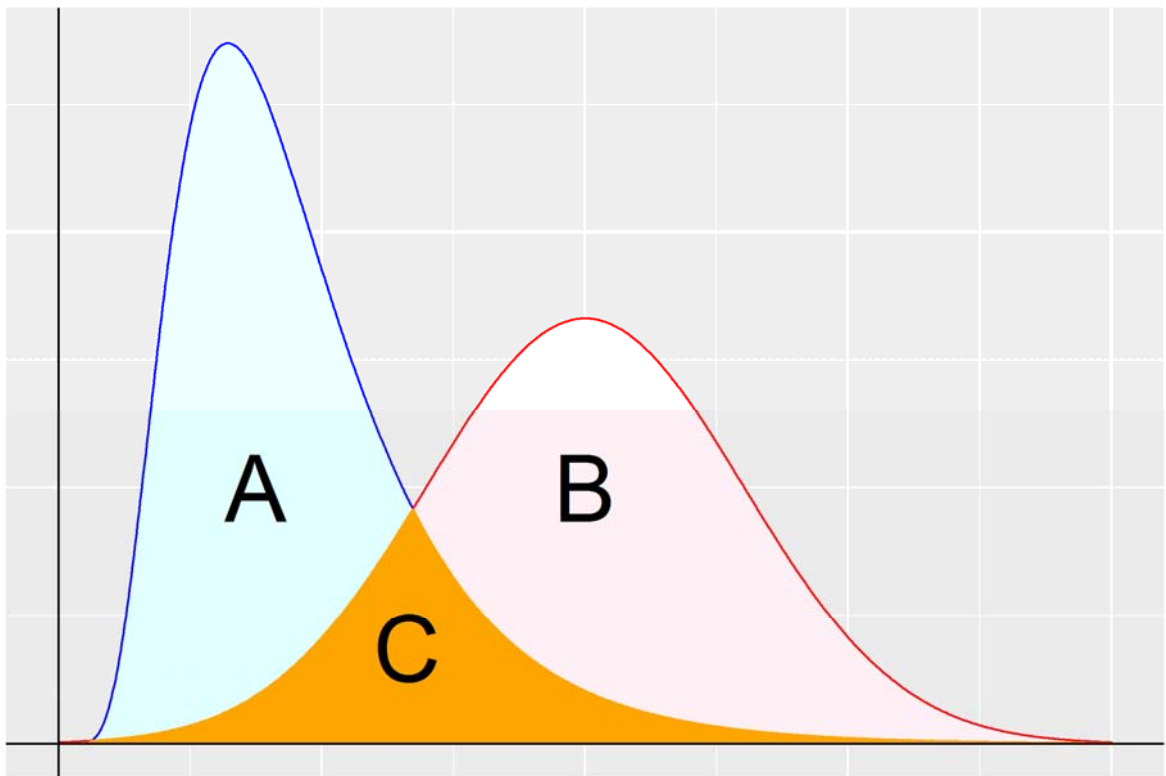
This research note attempts to fill in this gap by proposing a single index, bounded between zero and one, which can be routinely calculated after non-parametric kernel density estimators have been computed. In this way we can supplement these densities with a numerical measure of convergence, which is similar to how the Gini inequality index offers a bounded numerical summary of the Lorenz curve.

The structure of the paper is as follows. In the next section the index is represented graphically and computational details are provided. The following section presents a simple illustration which shows how poor countries have been catching up to rich ones since 1970, using the Penn World Tables (Feenstra, Inklaar and Timmer 2015). A brief last section offers concluding comments.

2. *A simple catching-up and convergence index.*

Non-parametric kernel density estimators offer a simple and graphical way to perform exploratory data analysis. By comparing various densities, across time or populations, we can discover interesting patterns in the evolution of the underlying data (Quah 1997). Figure 1 offers a simple illustration.¹ We can see two densities of a given variable that overlap by a given amount, as represented by C . It is worth remembering that the area under these curves represents probabilities which, for example, translate into shares of population if the variable in question is per capita income, and thus in this case the density shows the distribution of individual incomes among people in a given society. Hence, the full area under a given curve should be equal to 1. In terms of figure 1, this means that $A + C = B + C = 1$. From this it should be clear that $A + B + 2C = 2$, or $A + B + C = 2 - C$.

Figure 1: Graphical representation of the catching-up and convergence index



¹ An interactive Shiny application showing the intuition behind the index is available on the following website: https://goerlich.shinyapps.io/dist_convergence/.

Clearly, if the two densities are very far from one another, and they do not overlap, $C = 0$, the underlying societies have nothing in common regarding the variable under investigation. Using convergence terminology, we could say that we have two clubs, poor and rich, fully separated. If β -convergence is taking place, then the poor will, on average, grow faster than the rich. In terms of figure 1, this means that C should increase as time elapses, and consequently A and/or B should decrease. However, visual inspection of the full distribution may detect other behaviors, such as very poor or very rich countries, or individuals, departing from the average performance. Full catching-up of the rich by the poor means that $C \rightarrow 1$, so $A, B \rightarrow 0$, and eventually we would have a single density.

Developing a catching-up and convergence measure by using a non-parametric kernel density estimation is straightforward. Clearly this index should be related to area C in figure 1; in fact, this area constitutes a suitable candidate for such a measure, since it is bounded between 0, signaling absence of convergence, and 1, showing full convergence. Paralleling, and generalizing, the graphical derivation of the Gini index from areas under the Lorenz curve (Yitzhaki 1998), the catching-up and convergence index, δ_α , is calculated as

$$\delta_\alpha = \frac{\text{Intersection Area}}{(\text{Union Area})^\alpha} = \frac{C}{(A+B+C)^\alpha} \quad (1)$$

for $\alpha \geq 0$. Undoubtedly the index is bounded between 0, $C = 0$, and 1, $C = 1$ and $A = B = 0$. For $\alpha = 0$, δ_0 , the index is simply the intersection area in figure 1, C , simple and intuitive as the mixture between distributions. For $\alpha = 1$, δ_1 , the index is only the ratio of the intersection to the union area, which is similar to the ratio of areas under the Lorenz curve in the derivation of the Gini index (Yitzhaki 1998) and close in spirit to the spatial segregation index of O'Sullivan and Wong (2007).² Because the union area, $A + B + C$, is always greater than 1, the δ_α index for $\alpha = 1$, δ_1 , is always lower than the intersection area, δ_0 , and the index falls

as α increases. In fact, $\delta_0 = C$ and $\delta_1 = \frac{C}{A+B+C} = \frac{C}{2-C} = \frac{\delta_0}{2-\delta_0}$. Beyond the particular

² Essentially the segregation index is defined as the complement of the catching-up and convergence index, $1 - \delta_\alpha$.

value of the index, it should be clear that the poor catching up with the rich always translates into a higher value in the index, albeit at a lower rate as α increases.³

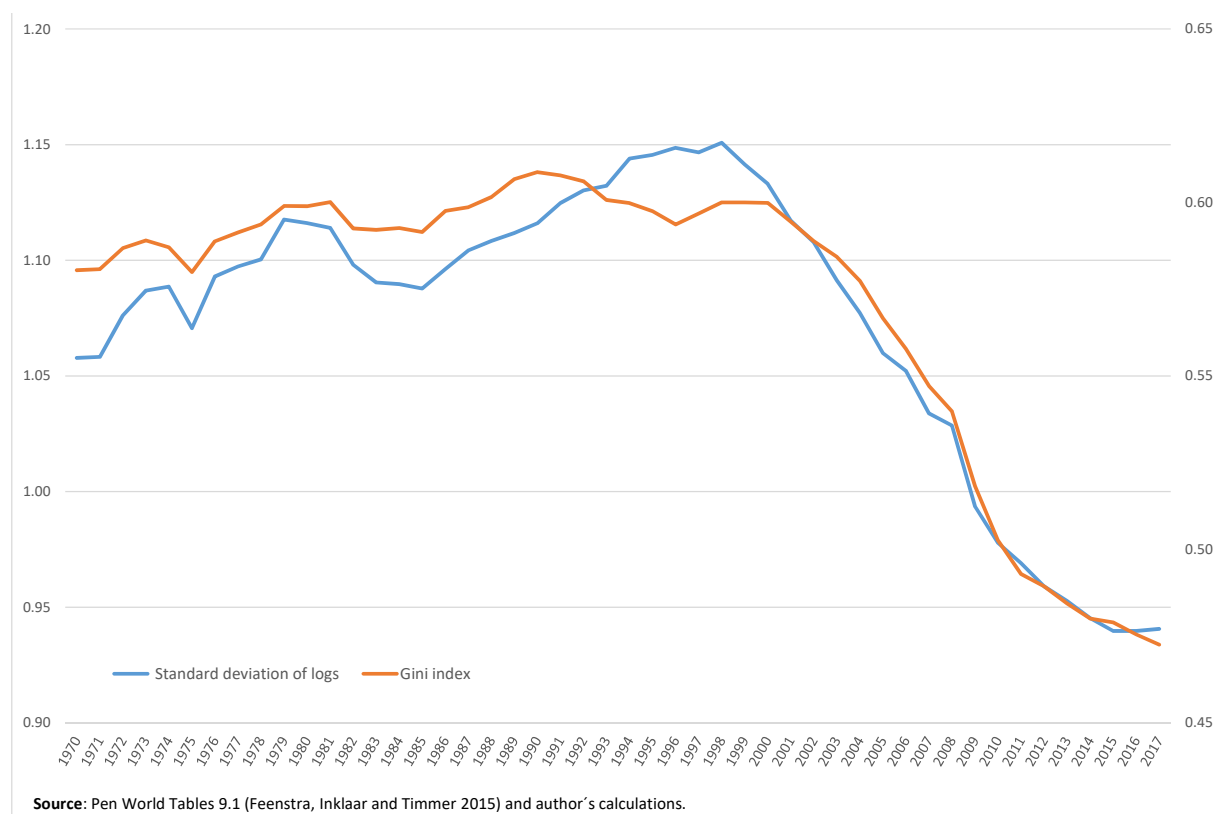
From a practical point of view, once the kernel densities have been estimated, the computation of the index is straightforward by means of numerical integration of the corresponding areas. If $f(x)$ and $g(x)$ are the ordinates of the two densities represented in figure 1, the intersection area is computed by numerically integrating, over the corresponding domain, $\min(f(x), g(x))$, whereas the union area is computed by numerically integrating, over the same domain, $\max(f(x), g(x))$.

³ In this sense, δ_α can be considered as an ‘equality’ index, instead of an ‘inequality’ index. If we want to recast it into an ‘inequality’ or ‘diverging index, then we can use the complement, $1 - \delta_\alpha$. For $\alpha = 0$, this is simply $1 - C$. For $\alpha = 1$, this is $1 - \delta_1 = \frac{A+B}{A+B+C}$.

3. *The index in practice. Are poor countries catching up with rich ones?*

As an illustration, we offer a simple application of the index to the convergence hypothesis at country level. Using per capita income from the Pen World Tables 9.1 (Feenstra, Inklaar and Timmer 2015), we examine whether poor countries have been catching up with rich ones from 1970 to the present day. In total, we have data –per capita income in US dollars (purchasing power parities, PPP), base year = 2011, and population– for 156 countries for almost half a century. Along this period, average world per capita income has been multiplied by a factor of 3, but inequality remains huge at country level, as the Gini index oscillates between 0.5 and 0.6. Standard deviation of logs, and σ -convergence statistics more generally, show a clear trend, with divergence from 1970 to the end of the century and convergence during the last 20 years –figure 2–. Income dispersion is now much lower than fifty years ago.

Figure 2: Sigma convergence statistics



β -convergence regressions tell us a similar story –table 1–. For the full period we find that β -convergence increases very slowly, 0.3% per year. But this is the result of the absence of β -convergence in the initial years, even divergence in some decades, together with a

stronger β -convergence along the 21st century. In general, β -convergence results are highly sensitive to the particular period studied, and when found it is never very strong.

Table 1: β -convergence for different periods

Decades	β (%)	t-statistic	Period	β (%)	t-statistic
1970-1980	0.35	1.46	1970-1996	-0.13	-0.60
1980-1990	-0.34	-1.21	1996-2017	-0.51	-3.77
1990-2000	0.58	2.84	Full Period	β (%)	t-statistic
2000-2010	-0.63	-3.03			
2010-2017	-0.57	-3.34			
			1970-2017	-0.28	-2.20

Note: β coefficient is the slope regression coefficient of the average (annualized) growth rate on the log of the initial per capita income, multiplied by 100, so it is comparable across periods.

Source: Pen World Tables 9.1 (Feenstra, Inklaar and Timmer 2015) and author's calculations.

Given this data set, we define the 'poor' club of countries as those with lower per capita income than the global average in the initial year, 1970. The remaining countries are labeled as 'rich'. With this definition, 96 countries belong to the 'poor' club. In that particular year, they represented $\frac{3}{4}$ of the total population, but accrued only $\frac{1}{4}$ of total income. The number of people in the 'poor' club has risen along the period, reaching 80% in 2017, but the share of total income of these countries has also increased up to 48%. As a result, per capita income of the 'poor' club has multiplied by 5, whereas for the remaining 60 countries, belonging to the 'rich' club, the scale factor is slightly lower than 3. On average, the poor seem to have somewhat caught up to the rich between 1970 and 2017. But what remains to be seen is the evolution of the 'poor' club as a whole, in relation to the 'rich' club.⁴

To inspect this process in much more detail, non-parametric kernel density estimators were calculated for all available years, for rich and poor countries separately, and the corresponding catching-up index, δ_1 , was computed. Starting from an almost zero value in the index, since the definitions of the rich and poor imply that they do not overlap in per capita income in the initial year, the catching-up and convergence index shows a clear up-ward trend, which means that the poor are, indeed, catching up to the rich, or at least, that some of the rich countries are falling behind in the development process.

⁴ This application implies that two clubs are defined at a given moment in time, in order to see the temporal evolution of both. Membership is fixed across periods. It could be technically possible to define more clubs, and extend the index (1) as a summary of the overlapping areas, but much of the simplicity and intuition concerning the index would get lost.

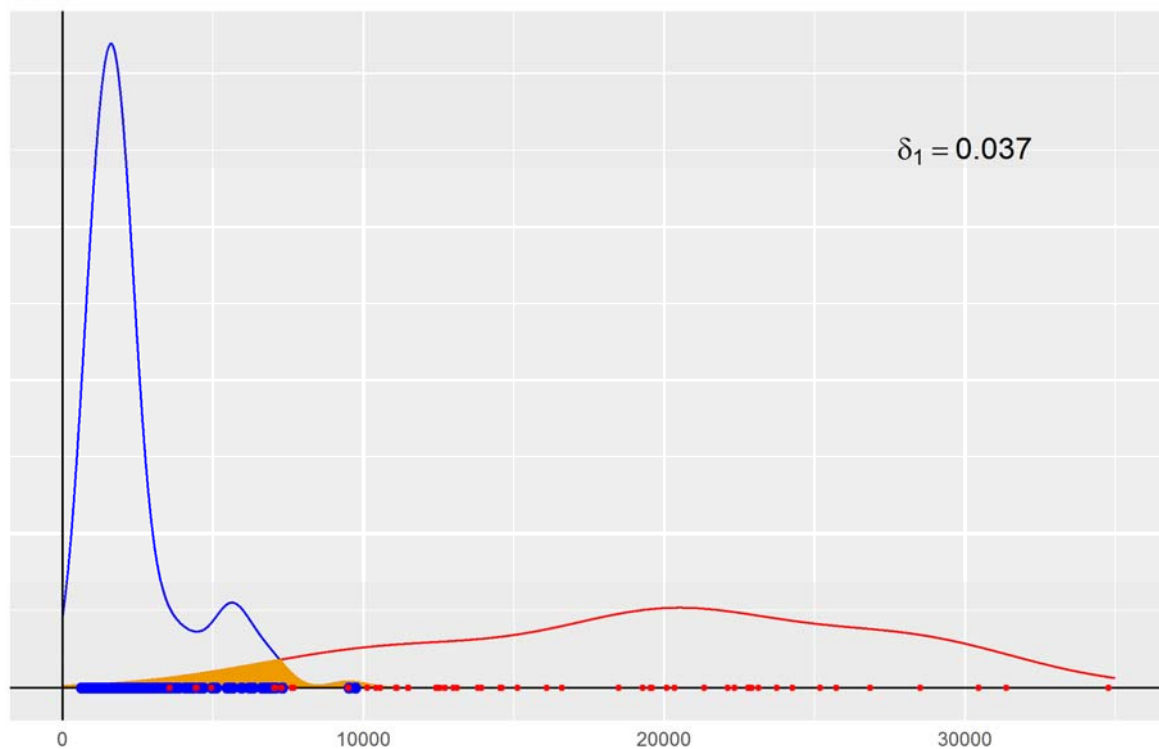
Results are depicted in figure 3 for the selected years: 1980, 1995, 2010 and 2017. We can see the ‘poor’ distribution –blue–, the ‘rich’ one –red–, and the overlapping area –orange–. The figure also shows the value of the catching-up index, δ_1 , and, on the x-axis, the income per capita for the different countries, and so we can inspect the convergence process, as the observations of the poor and of the rich are mixed together. The values for the overlapping area, δ_0 , for these years are 0.071 for 1980, 0.157 for 1995, 0.194 for 2000 and 0.183 for 2017, showing the same tendency as δ_1 .

Looking at almost five decades of development, we can observe some interesting facts. The distribution of the poor is quite narrow as compared with the rich one, and in 1980 a nascent twin-peakedness is clearly visible for these countries. The distribution of the rich is much wider, and peaks much less pronounced. From the end of the 20th century onwards, the distribution sequence of the ‘poor’ club shows three peaks instead of two. Given that these distributions are weighted by population, the middle peak, which shifts to the right as time elapses, is associated to the development of China, which has multiplied its per capita income along the period by a factor of almost 10.

The catching-up process of the rich by the poor is clearly visible from a graphical point of view until 2010. This is reflected in the value of δ_1 , which shows a slightly decreasing propensity in the last years, within the general tendency to increase. In this way, the δ_1 index offers a different perspective than the one shown by β and σ convergence statistics.

Figure 3: Weighted *kernel* density estimators and graphical representation of the catching-up index for selected years.

1980



1995

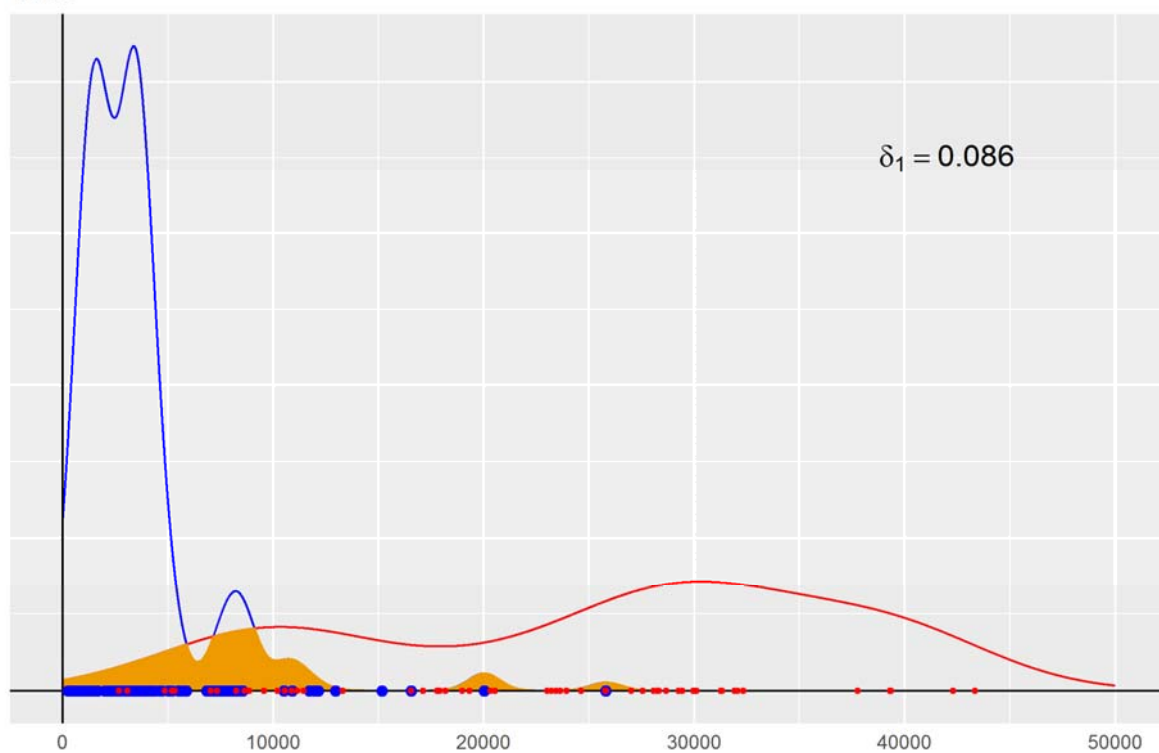
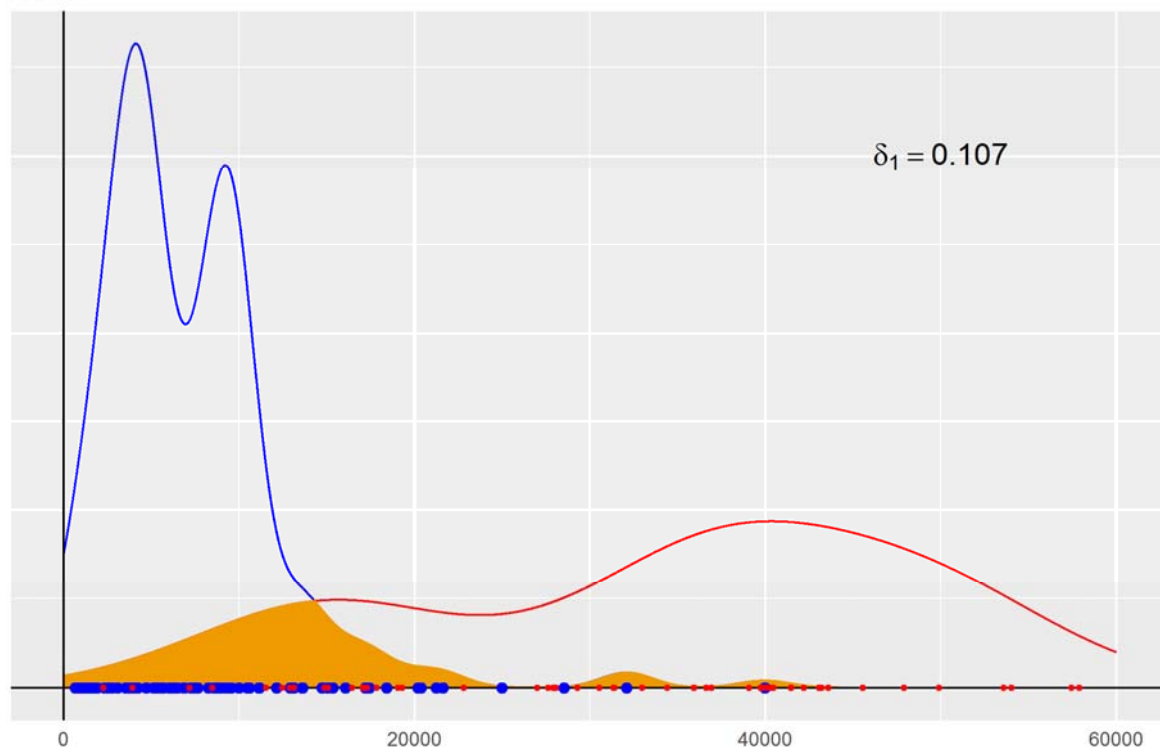
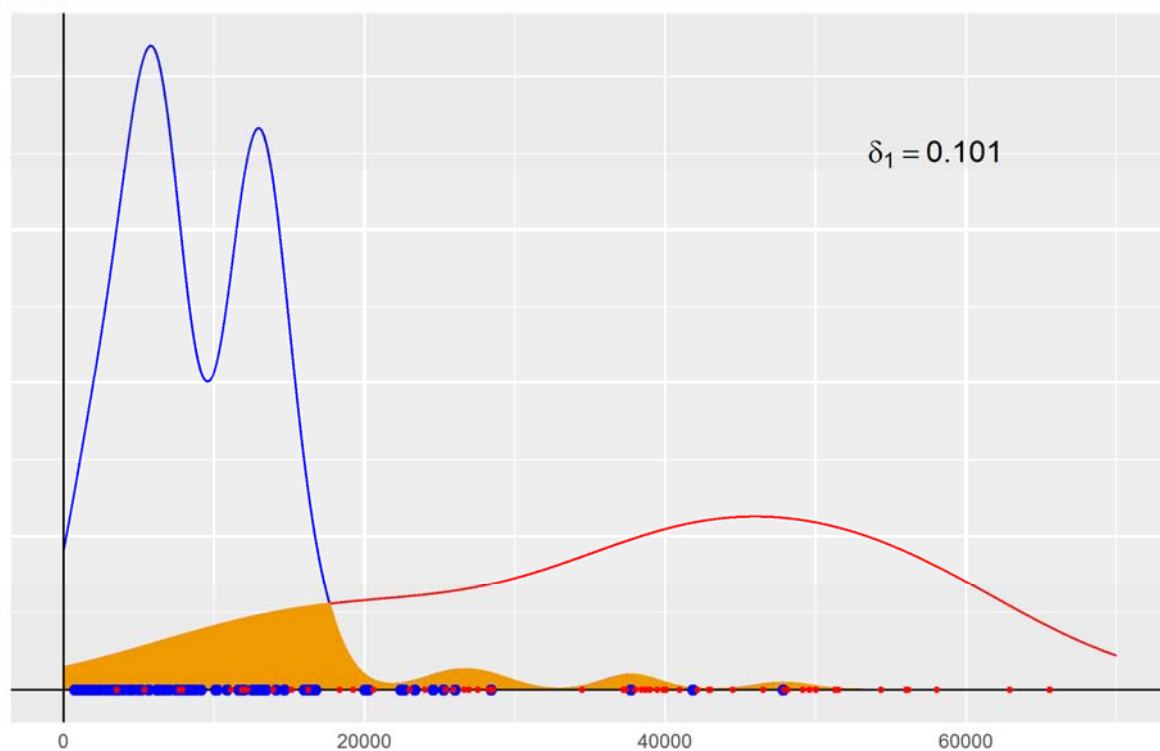


Figure 3: Weighted *kernel* density estimators and graphical representation of the catching-up index for selected years (cont.).

2010



2017



Source: Pen World Tables 9.1 (Feenstra, Inklaar and Timmer 2015) and author's calculations.

4. - *Final comments.*

This short note has introduced a simple non-parametric catching-up and convergence index that can be routinely computed once non-parametric kernel density estimators have been calculated. In this way, the exploratory data analysis commonly performed using kernel density estimators can be concisely summarized in a single number that complements this graphical device. The δ_α index can then supplement the β and σ convergence statistics. As an example, the usefulness of the index has been illustrated by the catching-up of rich countries by poor ones that has taken place in recent decades.

References

1. **Barro, R. J. and X. Sala-i-Martin (1992)** “Convergence”, *Journal of Political Economy*, 100, 21, (April), 223-251.
2. **Barro, R. J. and X. Sala-i-Martin (1998)** *Economic Growth*, MIT Press.
3. **Bishop, J. A., Formby, J. P. and Thistle, P. D. (1992)** “Convergence of the south and non-south income distributions, 1969-1979”. *The American Economic Review*, 82, 1, (March), 1969–1979.
4. **Dagum, C. (1980)** “Inequality measures between income distributions with applications”. *Econometrica*, 48, 7, (November), 1791-1803.
5. **Ebert, U. (1984)** “Measures of distance between income distributions”. *Journal of Economic Theory*, 32, 266-274.
6. **Feenstra, R. C.; Inklaar, R. and Timmer, M. P. (2015)** “The Next Generation of the Penn World Table”. *The American Economic Review*, 105, 10, (October), 3150–3182.
7. **O’Sullivan, D. and Wong, D. W. S. (2007)** “A surface-based approach to measuring spatial segregation”. *Geographical Analysis*, 39, 147–168. doi:10.1111/j.1538-4632.2007.00699.x
8. **Quah, D. T. (1993)** “Empirical cross-section dynamics in economic growth”. *European Economic Review*, 37, 2-3, 426–434. doi:10.1016/0014-2921(93)90031-5
9. **Quah, D. T. (1996a)** “Empirics for economic growth and convergence”. *European Economic Review*, 40, 6, 1353–1375. doi:10.1016/0014-2921(95)00051-8
10. **Quah, D. T. (1996b)** “Regional convergence clusters across Europe”. *European Economic Review*, 40, 3-5, 951–958. doi:10.1016/0014-2921(95)00105-0
11. **Quah, D. T. (1997)** “Empirics for growth and distribution: stratification, polarization, and convergence clubs”. *Journal of Economic Growth*, 2, 27–59. doi:10.1023/a:1009781613339
12. **Shorrocks, A. (1980)** “On the distance between income distributions”. *Econometrica*, 50, 5, (September), 1337-1339.
13. **Wong, D. W. S. (1999)** “Geostatistics as measures of spatial segregation”, *Urban Geography*, 20, 7, 635-647.
14. **Yitzhaki, S. (1994)** “Economic distance and overlapping of distributions”. *Journal of Econometrics*, 61, 147-159.
15. **Yitzhaki, S. (1998)** “More Than a Dozen Alternative Ways of Spelling Gini”. *Research on Economic Inequality*, 8, 13–30.
16. **Yitzhaki, S. and Lerman, R. (1991)** “Income stratification and income inequality”. *Review of Income and Wealth*, 37, 313-329.



Ivie

Guardia Civil, 22 - Esc. 2, 1º
46020 Valencia - Spain
Phone: +34 963 190 050
Fax: +34 963 190 055

Website: <http://www.ivie.es>
E-mail: publicaciones@ivie.es