Région et Développement

n° 47-2018

www.regionetdeveloppement.org

Sigma-Convergence revisited

Dimitris KALLIORAS*, **Marie-Noëlle DUQUENNE****,
Stevi VAFEIADOU**

Abstract - *Referring to the diachronic decrease of the overall dispersion of a (regional) dataset, σ-convergence is a dominant concept in the empirical regional convergence / divergence literature. The paper revisits the σ-convergence concept, expressing the, "classical", coefficient of variation and weighted coefficient of variation formulas against the backdrop of the median. To this end, the paper specifies and proposes a pair of, "alternative", formulas for apprehending the σ-convergence concept. Such an endeavor stems from the, purely, statistical rationale that the mean is a central tendency measure highly sensitive to the eventual presence of outliers. The theory-driven propositions of the paper are supported from an illustrative empirical analysis of regional inequalities in France, at the NUTS III spatial level, for the period 2001-2015. The findings of the analysis provide valuable insight to both theory and policy-making, indicating that different expressions of the σ-convergence concept may lead to different inferences with respect to regional inequalities.*

Classification JEL
C13, C18, R11

Key Words
Regional inequalities
σ-convergence
Coefficient of variation
Weighted coefficient of variation
Mean
Median

* Department of Planning and Regional Development, University of Thessaly ; dkallior@uth.gr
** Department of Planning and Regional Development, University of Thessaly ; mdyken@uth.gr
** Department of Planning and Regional Development, University of Thessaly ; stvafeia@uth.gr

1. INTRODUCTION

Fueling the relative academic debate and providing insight to the evaluation of the relative policies, the evolution of regional inequalities is an issue of utmost importance (Islam 2003, Kostov and Le Gallo 2015, Artelaris and Petrakos 2016). Hence, the study of regional inequalities - in particular, the study of regional convergence[1] / divergence in terms of per capita Gross Domestic Product (GDP)[2] - is at the heart of regional science. From the policy viewpoint, the study of regional convergence / divergence may interpret as a sign with respect to the evaluation of the effectiveness and the efficiency of the implemented regional policy mix. Regional policy, aims, precisely, at reducing the level of regional inequalities in a growth-enhancing economic environment. From the theory viewpoint, the study of regional convergence / divergence may serve as an empirical exercise with respect to the affirmation of regional development theories. Questioning the position of the neoclassical theory that (regional) inequalities are bound to diminish with growth through the activation of market-emanating convergence mechanisms[3] in a policy-free environment, theories with sharply different policy implications, such as the endogenous (new) growth theory (Romer 1986, Lucas 1988, *inter alia*) and the new economic geography theory (Krugman 1991, Fujita 1993, *inter alia*), stress the argument[4] that growth is a spatially selective and cumulative process.

Referring to the diachronic decrease of the overall dispersion of a (regional) dataset, σ-convergence (Barro and Sala-i-Martin 1992) is a dominant concept in the empirical regional convergence / divergence literature.[5] Σ-convergence may, usually, apprehend through the coefficient of variation (CV) and the weighted coefficient of variation (wCV) formulas.[6] CV is a standardized (relative) measure of dispersion and may express as the ratio of the standard deviation of a (regional) dataset to the corresponding arithmetic mean (henceforth: mean), at a given date (Barro and Sala-i-Martin 1992). Including a weighting factor in the CV formula (Petrakos and Artelaris 2009), so as to account for the corresponding relative (regional) size[7] in the treatment of the (regional) dataset, allows for the compilation of the wCV formula, the weighted CV counterpart.

The paper revisits the σ-convergence concept, expressing the, "classical", CV and wCV formulas against the backdrop of the median (i.e. the central tendency measure

---

[1] Intuitively, the term (regional) "convergence" suggests a process whereby poor(er) (regional) economies catch-up to rich(er) ones (Abreu et al. 2005).

[2] Customarily, this is the variable under consideration in the empirical regional convergence / divergence literature.

[3] Particularly, the neoclassical convergence mechanisms are the diminishing marginal productivity of capital (Solow 1956, Swan 1956, *inter alia*), the comparative advantage in interregional trade (Heckscher 1919/1991, Ohlin 1933/1966, *inter alia*) and the interregional production factors movement (Borjas 1979, Greenwood et al. 1991, *inter alia*).

[4] Bringing earlier theories regarding the operation of economic space (Perroux 1955, Myrdal 1957, Hirschman 1958, *inter alia*) back to the forefront.

[5] Usually, σ-convergence is examined together with β-convergence. The concept of β-convergence refers to the relation between the levels of a (regional) dataset at a given date and the consequent corresponding growth rates for a given period, either in an unconditional (i.e. absolute) or in a conditional (i.e. *ceteris paribus*) fashion (Baumol 1986, Barro 1991, Barro and Sala-i-Martin 1992, Sala-i-Martin 1996). Β-convergence is a necessary (though not sufficient) condition for σ-convergence (Barro and Sala-i-Martin 1995).

[6] Σ-convergence may, also, apprehend through the mean logarithmic deviation (Dalgaard and Vastrup 2001). Gini coefficient (Gini 1912), Theil index (Theil 1967) and Atkinson index (Atkinson 1970) fall, also, within the σ-convergence rationale.

[7] Customarily, in terms of relative population i.e. the ratio of the regional to the corresponding country population. In any case, the sum of regional weights is equal to 1 (i.e. 100%).

that separates the higher half of the (regional) dataset from the corresponding lower one). To this end, the paper specifies and proposes a pair of, "alternative", formulas for apprehending the σ-convergence concept. Particularly, next to the CV formula, the paper specifies and proposes the CV-median (CVmd) formula. CVmd is a standardized measure of dispersion that may express as the ratio of the standard deviation of a (regional) dataset to the corresponding median, at a given date. Correspondingly, next to the wCV formula the paper specifies and proposes the wCV-median (wCVmd) formula. Apparently, wCVmd is the weighed CVmd counterpart. Such an endeavor stems from the, purely, statistical rationale that the mean is a central tendency measure highly sensitive to the eventual presence of outliers. The theory-driven propositions of the paper are supported from an illustrative empirical analysis of regional inequalities in France, at the NUTS[8] III spatial level, for the period 2001-2013, on the basis of per capita GDP and (relative) population data obtained from EUROSTAT.

The paper proceeds as follows. The next section highlights the drawback of the "classical" CV and wCV formulas. The third section introduces the, "alternative", CVmd and wCVmd formulas for the apprehension of σ-convergence. The fourth section provides the empirical assessment of the level and the evolution of regional inequalities in France, at the NUTS III spatial level, for the period 2001-2013, on the basis of both the "classical" and the "alternative" expressions of σ-convergence. The last section offers the conclusions and discusses the inferences with respect to regional inequalities.

2. THE CLASSICAL APPROACH OF Σ-CONVERGENCE

CV and wCV are the "classical" formulas for apprehending the σ-convergence concept (Barro and Sala-i-Martin 1992, Petrakos and Artelaris 2009).

CV (equation 1.1) is a standardized (relative) measure of dispersion and may express as the ratio of the standard deviation of a (regional) dataset to the corresponding mean, at a given date. CV takes values within the interval $[0, \sqrt{n-1}]$[9], from perfect (regional) equality to perfect (regional) inequality.[10] Increasing (decreasing) values of CV diachronically, evince an increase (a decrease) of (regional) inequality.

$$\boldsymbol{CV}_{c,t} = \frac{\sqrt{\sum_{r=1}^{n}\left[\frac{\left(\boldsymbol{Y}_{r,t}-\overline{\boldsymbol{Y}_{r,t}}\right)^2}{\boldsymbol{n}}\right]}}{\overline{\boldsymbol{Y}_{r,t}}} \qquad \textit{(equation 1.1)}$$

where $\boldsymbol{CV}$ is the CV(-mean), $\boldsymbol{Y}$ is the variable under consideration, $\overline{\boldsymbol{Y}}$ is the mean of the variable under consideration, $\boldsymbol{\Sigma}$ denotes sum, $\boldsymbol{r}$ stands for regions, $\boldsymbol{n}$ is the number of regions, $\boldsymbol{t}$ stands for time (date), $\boldsymbol{c}$ stands for country.

wCV (equation 2.1) is a standardized (relative) measure of dispersion and may express as the ratio of the standard deviation of a (regional) dataset to the corresponding mean, at a given date, accounting for an included weighting factor in the

---

[8] NUTS (Nomenclature des Unités Territoriales Statistiques; Nomenclature of Territorial Units for Statistics) is a EUROSTAT geocode standard for referencing the subdivisions of European Union (EU) countries for statistical purposes.

[9] In essence, given than the number of regions may reach infinity, the upper value of CV may reach infinity as well. In any case, the upper value of CV is not 1, as Monfort (2008) mistakenly (or inadvertently) refers.

[10] When the mean is equal to 0, CV is not defined. In the case of using per capita GDP data, this means that CV is not defined in the theoretical case that each region under consideration has zeroed per capita GDP.

treatment of the (regional) dataset. wCV takes values within the interval [0, n-1][11], from perfect (regional) equality to perfect (regional) inequality.[12] Increasing (decreasing) values of wCV diachronically, evince an increase (a decrease) of (regional) inequality.

$$\boldsymbol{wCV}_{c,t} = \frac{\sqrt{\sum_{r=1}^{n}\left[\boldsymbol{w}_{r,t} \times (\boldsymbol{Y}_{r,t} - \overline{\boldsymbol{Y}_{r,t}})^2\right]}}{\overline{\boldsymbol{Y}_{r,t}}} \qquad \textit{(equation 2.1)}$$

where $\boldsymbol{wCV}$ is the weighted CV(-mean), $\boldsymbol{Y}$ is the variable under consideration, $\overline{\boldsymbol{Y}}$ is the mean of the variable under consideration, $\boldsymbol{\Sigma}$ denotes sum, $\boldsymbol{r}$ stands for regions, $\boldsymbol{n}$ is the number of regions, $\boldsymbol{t}$ stands for time (date), $\boldsymbol{c}$ stands for country, $\boldsymbol{w}$ is the weighting factor.

Even though both CV and wCV comply with the vast majority of the axioms (properties) of inequality indicators[13], they are subject to the criticism, stemming from a, purely, statistical rationale, that the mean is a central tendency measure highly sensitive to the eventual presence of outliers (i.e. (regional) dataset values extremely distant from the other corresponding values). This is so as the calculation of the mean is not based on any measure concerning position, and this is not without effect on standard deviation. Particularly, extremely high values connote fat- (i.e. with positive kurtosis) and right-tailed (i.e. with positive skewness) distributions[14], whereas extremely low values connote thin- (i.e. with negative kurtosis) and left-tailed (i.e. with negative skewness) distributions[15]. In other words, because the mean value gives equal weight to all observations it can be highly sensitive to outliers (Brewer et al. 2005), especially when considering regional economic data, and as a result, the distribution of the observations utilized to measure σ-convergence may exhibit skewness and/or kurtosis. Thus, given the fact that such outliers represent actual, and not erroneous regional values (that, usually, correspond to metropolitan and to outermost regions), turning to the use of the truncated mean (i.e. discarding the outliers and then taking the mean of the remaining regional dataset values) is not a risk-free, and beyond critique, methodological choice. Similarly, assuming "a priori" a normal distribution of the regional dataset values is quite often an erroneous assumption with serious consequences as regards the adequacy of the statistical treatment and their robustness: statistical parameters "appropriate with normal distributions may be misleading when applied to non-normal distributions" (Sainani, 2012:1005).

---

[11] In essence, given than the number of regions may reach infinity, the upper value of wCV may reach infinity as well.

[12] When the mean is equal to 0, wCV is not defined. In the case of using per capita GDP data, this means that wCV is not defined in the theoretical case that each region under consideration has zeroed per capita GDP.

[13] An inequality indicator should comply with the axioms (properties) of (Litchfield 1999, Monfort 2008, Cowell 2011, *inter alia*): a) the Pigou-Dalton transfer principle: the inequality indicator increases in response to a mean-preserving spread (Pigou 1912, Dalton 1920), b) the income scale independence: the inequality indicator is invariant to uniform proportional increases or decreases (Cowell 1999), c) the principle of population: the inequality indicator is invariant to replications of the population (Dalton 1920), d) anonymity (symmetry): the inequality indicator is dependent only on the variable in terms of which inequalities are measured (Amiel and Cowell 1994), and e) decomposability: the inequality indicator may be broken down into constituent parts (Bourguignon 1979).

[14] Such as the t-student, the Poisson and the Laplace distributions.

[15] Such as the Bernoulli distribution.

Given that the variable under consideration is continuous and not discrete, the mean value of distributions with positive kurtosis and positive skewness is generally higher than the median value, whereas the mean value of distributions with negative kurtosis and negative skewness is generally lower than the median value. Briefly, "the mean lies towards the direction of skew (the longer tail) relative to the median" (Agresti and Finlay 1997:50)[16]. Normal distributions in contrast, represent perfectly symmetrical distributions so as the mean value is equal to the median (and the mode) value (Gunver et al. 2017). Even though normal distributions rarely exist in nature (Pearson 1920) – this is so especially with economics and social data – it is the most frequently-used distribution for explaining continuous variables. Assuming normality is effectively easier because in such a case, it is possible to only consider the mean value as measure of central tendency. But this may be "a trap" (Klatzmann 1996). It is precisely from this point of view that the classical approach of σ-convergence is subject to criticism: the measure of inequalities - as proposed through the equations 1.1 and 2.1 - is relevant to the extent that the mean value is also relevant. Consequently, when the data set is not following normal distribution, it seems useful to consider an alternative measure of σ-convergence based on the median. Following Brewer et al. (2005), it can be stipulated that because of their potential differences, it is useful to consider both mean and median in the measure of regional inequalities, thus avoiding what Klatzmann (1996) denominates "the trap of unique control". In such a context, alternative measures of regional inequalities are suggested and finally applied to the case of GDP per capita at NUTS 3 level in France, after verifying that the distribution of GDP per capita is clearly not following the normal distribution.

3. Σ-CONVERGENCE REVISITED: THE "ALTERNATIVE" FORMULAS FOR APPREHENDING S-CONVERGENCE

The paper specifies and proposes a pair of "alternative" formulas for apprehending the σ-convergence concept, expressing the, "classical", coefficient of variation (CV) and weighted coefficient of variation (wCV) formulas against the backdrop of the median. Particularly, next to the CV formula, the paper specifies and proposes the CVmd formula, and, correspondingly, next to the wCV formula, the paper specifies and proposes the wCVmd formula.

CVmd (equation 1.2) is a standardized measure of dispersion that may express as the ratio of the standard deviation of a (regional) dataset to the corresponding median, at a given date. CVmd takes values greater than (or equal to) 0, from perfect (regional) equality to perfect (regional) inequality.[17] Increasing (decreasing) values of CVmd diachronically, evince an increase (a decrease) of (regional) inequality.

$$\boldsymbol{CVmd_{c,t}} = \frac{\sqrt{\sum_{r=1}^{n}\left[\frac{\left(Y_{r,t} - \widetilde{Y_{r,t}}\right)^2}{n}\right]}}{\widetilde{Y_{r,t}}} \qquad \textit{(equation 1.2)}$$

[16] Von Hippel (2005) demonstrates that this rule is quite often violated with discrete variables, while this can happen, even though continuous violations are rarer, with continuous variables, in case of multimodal continuous densities, as well.

[17] When the median is equal to 0, CVmd is not defined. In the case of using per capita GDP data, this means that CVmd is not defined in the theoretical case that the majority of regions under consideration has zeroed per capita GDP.

where $\boldsymbol{CV_{md}}$ is the CV-median, $\boldsymbol{Y}$ is the variable under consideration, $\boldsymbol{\widetilde{Y}}$ is the median of the variable under consideration, $\boldsymbol{\Sigma}$ denotes sum, $\boldsymbol{r}$ stands for regions, $\boldsymbol{n}$ is the number of regions, $\boldsymbol{t}$ stands for time (date), $\boldsymbol{c}$ stands for country.

wCVmd (equation 2.2) is a standardized measure of dispersion that may express as the ratio of the standard deviation of a (regional) dataset to the corresponding median, at a given date, accounting for an included weighting factor in the treatment of the regional dataset. wCVmd takes values greater than (or equal to) 0, from perfect (regional) equality to perfect (regional) inequality.[18] Increasing (decreasing) values of wCVmd diachronically, evince an increase (a decrease) of (regional) inequality.

$$\boldsymbol{wCVmd_{c,t}} = \frac{\sqrt{\sum_{r=1}^{n}\left[\boldsymbol{w_{r,t}} \times (\boldsymbol{Y_{r,t}} - \widetilde{\boldsymbol{Y_{r,t}}})^2\right]}}{\widetilde{\boldsymbol{Y_{r.t}}}} \qquad \textit{(equation 2.2)}$$

where $\boldsymbol{wCV_{md}}$ is the weighted CV-median, $\boldsymbol{Y}$ is the variable under consideration, $\boldsymbol{\widetilde{Y}}$ is the median of the variable under consideration, $\boldsymbol{\Sigma}$ denotes sum, $\boldsymbol{r}$ stands for regions, $\boldsymbol{n}$ is the number of regions, $\boldsymbol{t}$ stands for time (date), $\boldsymbol{c}$ stands for country, $\boldsymbol{w}$ is the weighting factor.

Both CVmd and wCVmd overcome the drawback of the "classical" CV and wCV formulas. Being in line with the fact that international organizations (EUROSTAT 1999, OECD 2007, *inter alia*) perceive the median – and not the mean – as the central tendency measure for defining thresholds, especially risk of poverty, such a methodological suggestion aims at offering an alternative perspective with respect to the empirical assessment of the level and the evolution of regional inequalities.

4. REGIONAL INEQUALITIES IN FRANCE: AN ILLUSTRATIVE EMPIRICAL ANALYSIS

4.1. General tendencies of per capita GDP during the period 2001-2015

The theory-driven propositions of the paper are supported from an illustrative empirical analysis of regional inequalities in France, at the NUTS III spatial level, for the period 2001-2015, on the basis of per capita GDP[19] and (relative) population data obtained from EUROSTAT.

France, spanning 643,801 km^2, comprises of 101 NUTS III regions (Table A1 and Figure A1 in the Annex), with Mayotte, La Réunion, Guyane, Guadeloupe, and Martinique having the status of overseas regions[20]. During the period under consideration, France exhibits a per capita GDP level ranging from 24,000 PPS/inh. (2001) to 30,600 PPS/inh. (2015) (Table 1). Systematically, Mayotte and Guyane are the poorest French regions. Paris, the capital region of France, is the richest region up to year 2006 followed by Hauts-de-Seine (i.e. the western inner suburbs of Paris) and Rhône while the Hauts-de Seine[21] is the richest French region onwards, Rhône remaining stable at the 3rd place.

[18] When the median is equal to 0, wCVmd is not defined. In the case of using per capita GDP data, this means that wCVmd is not defined in the theoretical case that the majority of regions under consideration has zeroed per capita GDP.

[19] Per capita GDP is expressed in Purchasing Power Standars per inhabitant (PPS/inh.).

[20] Overseas French regions are integral parts of France and have similar powers to the regions of metropolitan (i.e. European) France while Saint-Pierre and Miquelon hast he status of territorial collectivity.

[21] Even though it exhibits a high level of intra-regional inequalities (Feltin 2000), Hauts-de-Seine is well-known as the richest region in France including the business district of La Défense and communes such as Neuilly-sur-Seine and Saint-Cloud.

The mean value of per capita GDP ranges from 20,967 PPS/inh. (2001) to 26,300 PPS/inh. (2015), whereas the corresponding median value presents systematically lower level from 20,100 (years 2001 and 2003) to 23,800 (2016). Depending on the year, the median value in itself corresponds to different regions while some of them are throughout the whole period around the median values (between the 4th and 6th deciles), especially the departments of Cher, Loir et Cher, Morbihan Oise and Somme. Twelve other departments are also very often but not systematically near from the median: Ain, Corrèze, Haute-Marne, Hautes-Pyrénées, Haute-Vienne, Jura, Loire, Maine-et-Loire, Moselle, Sâone-et-Loire, Var and Vienne (Figure A2 in Annex).

As expected, Paris and Hauts-de-Seine are, during the entire period under consideration, high-level outliers and to a lesser extend Rhône. Val-de-Marne (2008 to 2015 except 2013), Yvelines (2008, 2015) and Essonne (2008) – all three of them are suburbs of Paris – join, sporadically, the group of high-level outliers. In contrast, Mayotte is a low-level outlier during the entire period under consideration with Guyane, also a low-level outlier thrice (2004, 2005 and 2007).

Table 1. Level and evolution of per capita GDP in France: analysis at the NUTS 3 spatial level (2001-2015)

	2001	2002	2003	2004	2005	2006	2007	2008	2009	2010	2011	2012	2013	2014	2015
p.c. GDP (PPP / inh.)	24,000	24,600	24,000	24,700	26,00	26,900	28,200	27,800	26,400	27,500	28,300	28,500	29,000	29,600	30,600
minimum	*4,200*	*4,300*	*4,500*	*4,900*	*5,500*	*6,100*	*6,400*	*7,000*	*6,700*	*7,200*	*7,300*	*7,500*	*7,900*	*8,600*	*9,400*
	FRA50	*FRA50*	*FRA50*	*FRA50*	*FRA50*	*FRA50*	*FRA50*	*FRA50*	*FRA50*	*FRA50*	*FRA50*	*FRA50*	*FRA50*	*FRA50*	*FRA50*
maximum	*68,700*	*69,200*	*66,100*	*66,600*	*70,700*	*71,300*	*76,100*	*79,200*	*74,600*	*81,100*	*81,900*	*83,600*	*86,600*	*88,300*	*90,400*
	FR101	*FR101*	*FR101*	*FR101*	*FR101*	*FR101*	*FR101*	*FR105*	*FR105*	*FR105*	*FR105*	*FR105*	*FR105*	*FR105*	*FR105*
mean	*20,967*	*21,565*	*21,035*	*21,681*	*22,805*	*23,579*	*24,569*	*23,828*	*22,710*	*23,363*	*24,129*	*24,157*	*24,604*	*25,199*	*26030*
median	*20,100*	*20,700*	*20,100*	*20,700*	*21,700*	*22,300*	*23,200*	*22,400*	*21,000*	*21,300*	*22,000*	*22,00*	*22,300*	*23,100*	*23,800*
	FR301 FR613 FR813	*FR241 FR432 FR512 FR534*	*FR432 FR711*	*FR263*	*FR825 FR626*	*FR222 FR413 FR524 FR613*	*FR222 FR263 FR512 FR711 FR715*	*FR222*	*FR245 FR263 FR631*	*FR263 FR631*	*FR245 FR263 FR633*	*FR633*	*FR263 FR414*	*FR223 F414*	*FR214 FR512*
outliers (high-level)	*3*	*3*	*3*	*3*	*3*	*3*	*3*	*6*	*4*	*4*	*4*	*4*	*3*	*5*	*2*
	FR101 FR105 FR716	*FR101 FR105 FR716*	*FR101 FR105 FR716*	*FR101 FR105 FR716*	*FR101 FR105 FR716*	*FR101 FR105 FR716*	*FR101 FR105 FR716*	*FR101 FR105 FR103 FR104 FR107 FR716*	*FR101 FR105 FR107 FR716*	*FR101 FR105 FR107 FR716*	*FR101 FR105 FR107 FR716*	*FR101 FR105 FR107 FR716*	*FR101 FR105 FR716*	*FR101 FR105 FR107 FR623 FR716*	*FR101 FR105*
outliers (low-level)	*1*	*1*	*1*	*2*	*1*	*1*	*2*	*1*	*1*	*1*	*1*	*1*	*1*		
	FRA50	*FRA50*	*FRA50*	*FRA30 FRA50*	*FRA30 FRA50*	*FRA50*	*FRA30 FRA50*	*FRA50*	*FRA50*	*FRA50*	*FRA50*	*FRA50*	*FRA50*	*FRA50*	*FRA50*

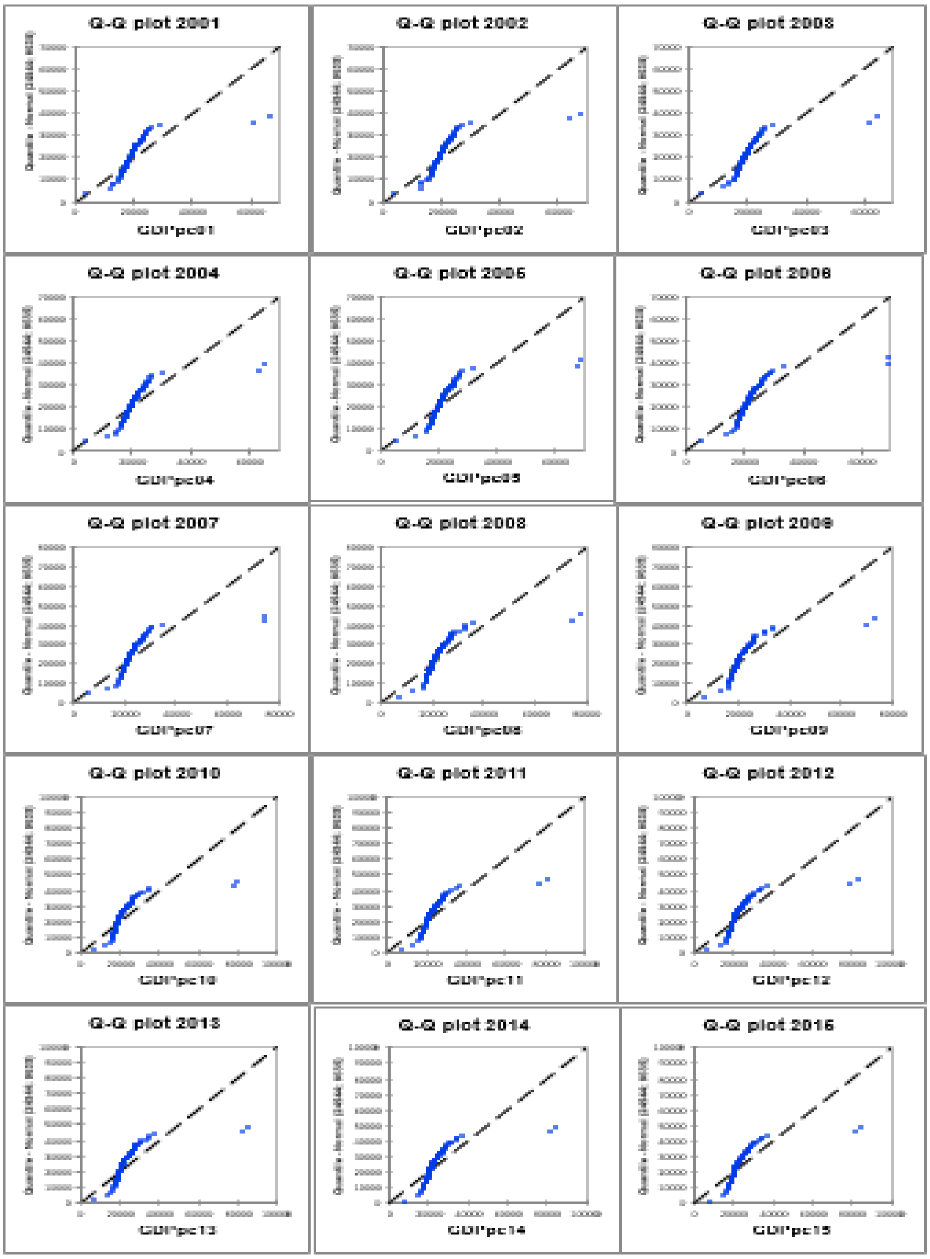
Sources: EUROSTAT / Authors' elaboration.

Observing the evolution of the mean and the median values, it comes that, between the years 2001 and 2015, the former increases by 24,1% (average annual rate of 1,6%) whereas the latter increases by 18.4% (respectively 1,2%). Apparently, the highest increase of the mean value comparatively to the median value is due to the impact of the high-level outliers. Indeed, the annual Q-Q plots[22] of the per capita GDP level of the NUTS III French regions (Figure 1) demonstrate that data do not in any case follow the normal distribution. Instead, the annual distributions of the data show positive kurtosis and positive skewness, indicating that the majority of regions considered exhibit per capita GDP level lower than the mean value.

Such an observation is considered to be perfectly awaited taking into consideration that systematically, during the entire period under consideration, the mean values are higher than the corresponding median values. This is so as the number of high-level outliers is systematically higher than the corresponding number of low-level outliers.

[22] Q-Q plot provides a graphical way to determine the level of normality (Wilk and Gnanadesikan 1968).

Figure 1: The annual Q-Q plots of the per capita GDP level of the NUTS III French regions, years 2001-2015



Source: EUROSTAT / Authors' elaboration.

Diachronically, the linear correlation between the per capita GDP level and the (relative) population level of the NUTS III French regions is positive, a frequent and quite logical pattern in case of countries - as France - with relatively high population's concentration in a few number of regions. Particularly, Pearson correlation coefficient[23] (Table 2) ranges from 0.544 (2002) to 0.605 (2012). Yet, it is still quite far from being characterized as perfectly positive. Such an observation indicates that the inclusion of the relative population as weighting variable in the assessment of the level and the evolution of regional inequalities may impact on the results.

Table 2. Linear correlation between the per capita GDP and the relative population of the NUTS 3 French regions (2001-2015)

	2001	2002	2003	2004	2005	2006	2007	2008	2009	2010	2011	2012	2013	2014	2015
Pearson correlation	0.545	0.544	0.551	0.552	0.555	0.568	0.567	0.597	0.598	0.601	0.600	0.605	0.597	0.579	0.582

Source: EUROSTAT / Authors' elaboration.

Towards performing the empirical analysis, both the "classical" (i.e. CV, wCV) and the "alternative" (i.e. CVmd, wCVmd) expressions of σ-convergence having be estimated (Table 3). The annual CV's variations are relatively limited with values ranging from 0.327 (2004 & 2006) to 0.394 (2013). The wCV reflects higher level of inequalities ranging from 0.514 (2006) to 0.621 (2013 & 2015), confirming that the spatial distribution of the population within the country clearly impacts the measure of inequalities. Obviously, the higher the degree of population's concentration in a limited number of metropoles, the greater is the measure of pcGDP' inequalities through the wCV coefficient, especially when the correlation between pcGDP and relative population is highly positive as it is the case for France. The two "classical" coefficient CV and wCV show the same pattern of annual fluctuations (Figure 2): during the first six years (2001-2006), they present a decreasing trend of regional inequalities about -1.0% and -1.2% per year, respectively while the next period (2006-2010) is characterized by a net increase of regional inequalities about +4,5% and +4,8% per year. Finally, from 2011, a relative stabilization can be observed.

Table 3. Regional inequalities among the NUTS 3 regions of France, estimations on the basis of both the "classical" (CV, wCV) and the alternative (CVmd, wCVmd) measures of σ-convergence (2001-2015)

		2001	2002	2003	2004	2005	2006	2007	2008	2009	2010	2011	2012	2013	2014	2015
non-weighted	CV	0.343	0.339	0.332	0.327	0.334	0.327	0.338	0.367	0.360	0.391	0.376	0.387	0.394	0.392	0.393
	CVmd	0.361	0.356	0.351	0.346	0.354	0.351	0.363	0.396	0.398	0.439	0.424	0.436	0.447	0.437	0.439
weighted	wCV	0.546	0.538	0.525	0.514	0.526	0.514	0.532	0.576	0.563	0.619	0.592	0.609	0.621	0.620	0.621
	wCVmd	0.582	0.573	0.563	0.553	0.569	0.562	0.582	0.634	0.638	0.713	0.683	0.704	0.722	0.702	0.711

Source: EUROSTAT / Authors' elaboration.

Considering the proposed "alternative" formulas, the values of the coefficient of variation based on the median (CVmd), are somewhat higher comparatively to the classical CV, ranging from 0.346 (2004) to 0.447 (2013) while the wCVmd varies from 0.553 (2004) to 0.722 (year 2013). Once again, these two alternative measures

[23] Pearson correlation coefficient is a measure of the linear correlation of two variables (Pearson 1895). It takes values in the interval [-1, 1], where -1 indicates perfectly negative linear correlation, 0 indicates no linear correlation, and 1 indicates perfectly positive linear correlation.

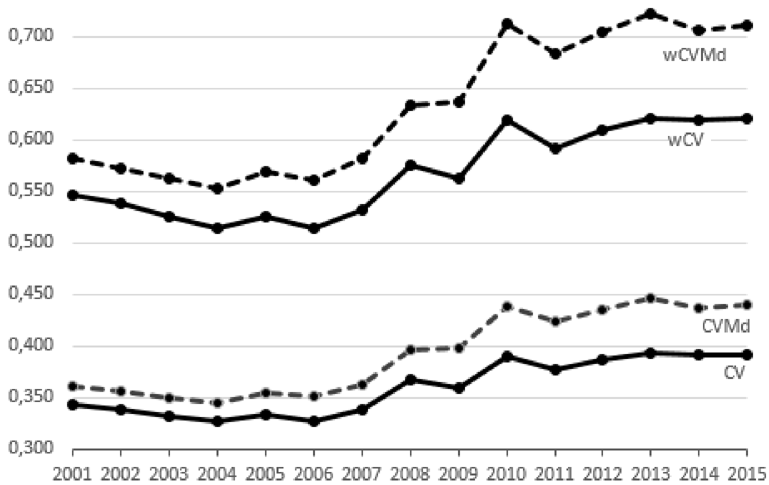
follow the same pattern of annual fluctuations (Figure 2) but it appears clearly that the intensity of their fluctuations differs from the classical coefficients (Table 4). Finally, with the alternative coefficients, it appears that the measure of the decline in regional inequalities is lower, while the increase in inequalities is significantly higher.

Table 4. Average annual rates of change of the "classical" (CV, wCV) and the "alternative" (CVmd, wCVmd) measures of σ-convergence (2001-2015)

Periods	"Classical"		"Alternative"	
	CV	wCV	CVmd	wCVmd
2001-2006	-0,96%	-1,22%	-0,57%	-0,72%
2006-2010	4,53%	4,76%	5,79%	6,13%
2010-2015	0,10%	0,07%	0,00%	-0,04%

Source: EUROSTAT / Authors' elaboration.

Figure 2. Regional inequalities among the NUTS III regions of France, estimation on the basis of both the "classical" (CV, wCV) and the "alternative" (CVmd, wCVmd) measures of σ-convergence period 2001-2013.



Source: EUROSTAT / Authors' elaboration.

The results of the empirical analysis evince that the level of regional inequalities in France appears to be higher in terms of the "alternative" formulas of the σ-convergence concept, comparing to the corresponding "classical" ones.[24] Particularly, the level of regional inequalities is, systematically, higher in terms of CVmd, comparing to the corresponding level in terms of CV, and in terms of wCVmd, comparing to

[24] The results of the empirical analysis, also, evince that, whilst the corresponding evolution seems to follow a similar pattern, the level of regional inequalities is higher in terms of wCV and wCVmd, comparing to the corresponding CV and CVmd formulas, respectively. In line with the considerations and the arguments of the corresponding literature (Firebaugh 2003, Sala-i-Martin 2003, Tortosa-Ausina et al. 2005, Petrakos and Artelaris 2009), it comes that the inclusion of a weighting factor (the variable of relative population, in particular) in the assessment of the level and the evolution of regional inequalities, indeed, impacts on the results.

the corresponding level in terms of wCV. Concerning the evolution of regional inequalities, it comes that the pattern is quite similar irrespective of the formula considered. Yet, a closer look indicates that the results are more sensitive, in given per capita GDP and (relative) population changes, against the backdrop of the median. Particularly, even though the median is a central tendency measure not sensitive to outliers, CVmd and wCVmd exhibit, during the whole period, higher increase than CV and wCV, respectively. The gap between CV and CVmd (as well as wCV and wCVmd) tends to widen over the years. In a nutshell, the illustrative empirical analysis of regional inequalities in France indicates that the estimation of the level and the evolution of regional inequalities with the use of the "classical" formulas of σ-convergence may mask the actual regional problem. This is so as regional inequalities appear to be lower and less sensitive against the backdrop of the mean. Even though sometimes it might be useful to policy-makers, such a discrepancy with the "alternative" formulas of σ-convergence may lead to conclusions scientifically misleading.

5. CONCLUSIONS AND INFERENCES

The paper revisits the σ-convergence concept and specifies the "alternative" CVmd and wCVmd formulas, expressing the "classical" CV and wCV formulas against the backdrop of the median. Such an endeavor stems from the, purely, statistical rationale that the mean is a central tendency measure highly sensitive to the eventual presence of outliers, and is in line with the fact that international organizations perceive the median as the central tendency measure for defining thresholds. The illustrative empirical analysis that supports the theory-driven propositions of the paper, indicates – as expected - the systematic presence of outliers and consequently the non-normal distribution of the data-sets, justifying the interest of reformulating or at least proposing a complementary measure of σ-convergence. This allows to avoid falling into "the trap of unique control" (Klatzmann, 1996). Moreover, the empirical analysis confirms that regional inequalities in France appear to be lower and less sensitive against the backdrop of the mean. Even though, sometimes, such results might be useful to policy-makers, the discrepancy between the "classical" and the "alternative" formulas of σ-convergence may mask the magnitude of the actual regional problem.

Indicating that different expressions of the σ-convergence concept may, in fact, lead to different inferences with respect to regional inequalities, the findings of the paper provide valuable insight to both theory and policy-making. Revisiting the σ-convergence concept, the paper casts strong doubts on the ability of the "classical" formulas to offer results not leading to conclusions scientifically misleading. Even though further empirical research is needed before the marginalization of the "classical" formulas of σ-convergence, the paper sets the ground for provoking the relative debate.

REFERENCES

Abreu M., de Groot H.L.F. and Florax R.J.G.M., 2005, A meta-analysis of b-convergence: The legendary 2, *Journal of Economic* Surveys, 19(3), 389-420.

Agresti A. and Finlay B., (ed), 1997, *Statistical Methods for the Social Sciences*, 3rd ed., Upper Saddle River, NJ: Prentice Hall.

Amiel Y. and Cowell F. A., 1994, Income inequality and social welfare, in Creedy J. (ed.), *Taxation, poverty and income distribution*, Cheltenham: Edward Elgar, 193-219.

Artelaris P. and Petrakos G., 2016, Intraregional spatial inequalities and regional income level in the European Union: Beyond the inverted-U hypothesis, *International Regional Science Review*, 39(3), 291-317.

Atkinson A.B., 1970, On the measurement of inequality, *Journal of Economic Theory*, 2, 244-263.

Barro R.J. (1991), Economic growth in a cross-section of countries, *Quarterly Journal of Economics*, 106, 407-443.

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

Barro R.J. and Sala-i-Martin X.X., (eds.), 1995, *Economic growth*, Boston: McGraw-Hill.

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

Borjas G., 1989, Economic theory and international migration, *International Migration Review*, 23(3), 457-485.

Bourguignon F., 1979, Decomposable income inequality measures, *Econometrica*, 47(4), 901-920.

Brewer M., Goodman A., Shaw J. and Shephard A., 2005, Poverty and inequality in Britain: 2005, *IFS*, C099.

Cowell F. A., (ed.), 2011, *Measuring inequality* (*3rd edition*), New York: Oxford University Press.

Cowell F.A., 1999, Measurement of inequality, in Atkinson A. B. and Bourguignon F. (eds.), *Handbook of income distribution*, Amsterdam: North Holland, 87-166.

Dalgaard C.-J. and Vastrup J., 2001, On the measurement of σ-convergence, *Economics Letters*, 70, 283-287.

Dalton H., 1920, The measurement of inequality of incomes, *Economic Journal*, 30, 348-361.

EUROSTAT, *http://ec.europa.eu/eurostat/web/regions/overview* (*last access: January 2018*).

EUROSTAT, (ed.), 1999, *European Community Household Panel Survey: Selected Indicators from the 1995 Wave*, Luxembourg: Office for Official Publications of the EC.

Feltin M., 2002, Where do you live in the Hauts-de-Seine? 36 Cities on the test bench, *L'Express*, 03.01.2002, https://www.lexpress.fr/informations/ou-vit-on-le-mieux-dans-les-hauts-de-seine-36-villes-au-banc-d-essai_646678.html (*in French*) (*last access: January 2018*).

Firebaugh G., (ed.), 2003, *The new geography of global income inequality*, Cambridge MA: Harvard University Press.

Fujita M., 1993, Monopolistic competition and urban systems, *European Economic Review*, 37, 308-315.

Gini C. (ed.), 1912, *Variability and mutability*, C. Cuppini: Bologna (*in Italian*).

Greenwood M.J., Hunt G.L., Rickman D.S. and Treyz G.I., 1991, Migration, regional equilibrium and the estimation of compensating differentials, *American Economic Review*, 81, 1382-1390.

Gunver M.G., Senocak M.S. and Vehid S., 2017, To determine skewness, mean and deviation with a new approach on continuous data, *PONTE International Journal of Sciences and Research*, 73(2), 30-44.

Heckscher E., 1991, The effect of foreign trade on the distribution of income, in Flam H. and Flanders M. (eds.), *Heckscher - Ohlin trade theory*, Cambridge MA: MIT Press, 43-69.

Hirschman A., 1958, *The strategy of economic development*, New Haven: Yale University Press.

Islam N., 2003, What have we learned from the convergence debate? *Journal of Economic Surv*eys, 17, 309-362.

Klatzmann J., 1996, *Attention Statistiques ! Comment en déjouer les pièges*, La Découverte, Paris, Poches, Essais No 17.

Kostov P. and Le Gallo J., 2015, Convergence: A story of quantiles and spillovers, *Kyklos*, 68(4), 552-576.

Krugman P.R., 1991, Increasing returns and economic geography, *Journal of Political Economy*, 99, 183-199.

Litchfield J.A., 1999, Inequality: Methods and tools,
http://siteresources.worldbank.org/INTPGI/Resources/Inequality/litchfie.pdf (*last access: January 2018*)

Lucas R., 1988, On the mechanics of economic development, *Journal of Monetary Economics*, 22(1), 3-42.

Monfort P., 2008, Convergence of EU regions: Measures and evolution, *EU DG for Regional Policy Working Papers*, 1.

Myrdal G. (ed.), 1957, *Economic theory and underdeveloped regions*, London: Hutchinson.

OECD (ed.), 2007, *Society at a glance: OEDC Social indicators*, Paris: OECD Publishing.

Ohlin B. (ed.), 1966, *Interregional and international trade*, Cambridge MA: Harvard University Press.

Pearson K., 1895, Notes on regression and inheritance in the case of two parents, *Proceedings of the Royal Society of London*, 58, 240-242.

Pearson K., 1920, The fundamental problem of practical statistics, *Biometrika*, 13(1), 1-16.

Perroux F., 1955, Note on the notion of growth poles, *Economic Appliquée*, 1(2), 307-320 (*in French*).

Petrakos G. and Artelaris P., 2009, European regional convergence revisited: A Weighted Least Squares approach, *Growth and Change*, 40(2), 314-331.

Pigou A. F., 1912, *Wealth and welfare*, London: Macmillan.

Romer P.M., 1986, Increasing returns and long-run, *Journal of Political Economy*, 94(5), 1002-1037.

Sala-i-Martin X.X., 1996, The classical approach to convergence analysis, *Economic Journal*, 106, 1019-1036.

Sala-i-Martin X.X., 2003, Keynote speech: Convergence and divergence- theoretical underpinnings, in Tumpel-Gugerell G. and Mooslechner P. (eds.), *Economic convergence and divergence in Europe: Growth and regional development in an enlarged European Union*, Cheltenham: Edward Elgar, 117-127.

Sainani K.L., 2012, Dealing With Non-normal Data, *Physical Medicine and Rehabilitation PM&R Journal*, 4, 1001-1005.

Solow R., 1956, A contribution to the theory of economic growth, *Quarterly Journal of Economics*, 70, 65-94.

Swan T., 1956, Economic growth and capital accumulation, *Economic Record*, 32, 334-361.

Theil H., 1967, *Economics and information theory*, Amsterdam: North Holland.

Tortosa-Ausina E., Pérez F., Mas M. and Goerlich F., 2005, Growth and convergence: Profiles in the Spanish Provinces (1965–1997), *Journal of Regional Science*, 45, 147-182.

Von Hippel P.T., 2005, Mean, median and skew: Correcting a textbook rule, *Journal of Statistics Education*, 13(2).

Wilk M.B. and Gnanadesikan R., 1968, Probability plotting methods for the analysis for the analysis of data, *Biometrika*, 55(1), 1-17.

ANNEXES

Figure A1. The nomenclature of the NUTS 3 French regions (Départements)

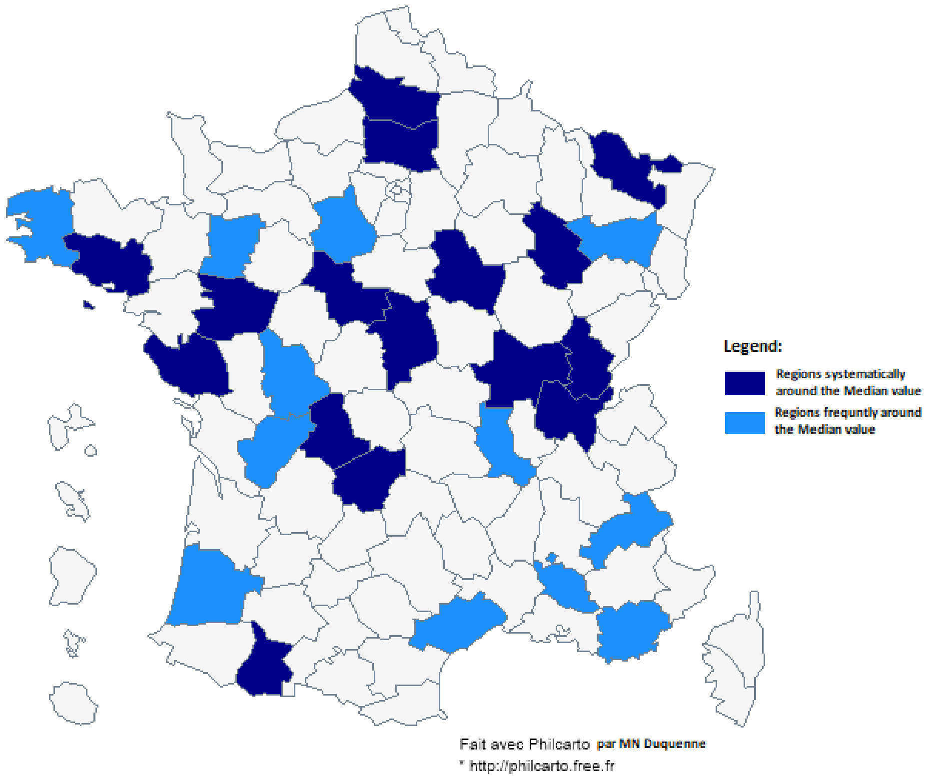
Sources: EUROSTAT / Authors' Elaboration.

Table A1. List of the 101 NUTS 3 Regions of France

FR101	Paris	FR413	Moselle	FR631	Corrèze
FR102	Seine-et-Marne	FR414	Vosges	FR632	Creuse
FR103	Yvelines	FR421	Bas-Rhin	FR633	Haute-Vienne
FR104	Essonne	FR422	Haut-Rhin	FR711	Ain
FR105	Hauts-de-Seine	FR431	Doubs	FR712	Ardèche
FR106	Seine-Saint-Denis	FR432	Jura	FR713	Drôme
FR107	Val-de-Marne	FR433	Haute-Saône	FR714	Isère
FR108	Val-d'Oise	FR434	Territoire de Belfort	FR715	Loire
FR211	Ardennes	FR511	Loire-Atlantique	FR716	Rhône
FR212	Aube	FR512	Maine-et-Loire	FR717	Savoie
FR213	Marne	FR513	Mayenne	FR718	Haute-Savoie
FR214	Haute-Marne	FR514	Sarthe	FR721	Allier
FR221	Aisne	FR515	Vendée	FR722	Cantal
FR222	Oise	FR521	Côtes-d'Armor	FR723	Haute-Loire
FR223	Somme	FR522	Finistère	FR724	Puy-de-Dôme
FR231	Eure	FR523	Ille-et-Vilaine	FR811	Aude
FR232	Seine-Maritime	FR524	Morbihan	FR812	Gard
FR241	Cher	FR531	Charente	FR813	Hérault
FR242	Eure-et-Loir	FR532	Charente-Maritime	FR814	Lozère
FR243	Indre	FR533	Deux-Sèvres	FR815	Pyrénées-Orientales
FR244	Indre-et-Loire	FR534	Vienne	FR821	Alpes-de-Haute-Provence
FR245	Loir-et-Cher	FR611	Dordogne	FR822	Hautes-Alpes
FR246	Loiret	FR612	Gironde	FR823	Alpes-Maritimes
FR251	Calvados	FR613	Landes	FR824	Bouches-du-Rhône
FR252	Manche	FR614	Lot-et-Garonne	FR825	Var
FR253	Orne	FR615	Pyrénées-Atlantiques	FR826	Vaucluse
FR261	Côte-d'Or	FR621	Ariège	FR831	Corse-du-Sud
FR262	Nièvre	FR622	Aveyron	FR832	Haute-Corse
FR263	Saône-et-Loire	FR623	Haute-Garonne	FRA10	Guadeloupe
FR264	Yonne	FR624	Gers	FRA20	Martinique
FR301	Nord	FR625	Lot	FRA30	Guyane
FR302	Pas-de-Calais	FR626	Hautes-Pyrénées	FRA40	La Réunion
FR411	Meurthe-et-Moselle	FR627	Tarn	FRA50	Mayotte
FR412	Meuse	FR628	Tarn-et-Garonne		

Source: EUROSTAT.

Figure A2. Regions with per capita GDP around the Median during the period 2001-2015



La sigma-convergence revisitée

Résumé - En se référant à l'évolution dans une période donnée de la dispersion globale d'un ensemble de données régionales, la sigma-convergence est une méthode couramment utilisée dans les analyses sur la convergence/divergence des régions. L'article revisite le concept de σ-convergence, en proposant de remplacer la moyenne arithmétique par la médiane dans les formules classiques de calcul du coefficient de variation et du coefficient de variation pondéré. À cette fin, l'article spécifie et propose une paire de formules alternatives pour appréhender le concept de σ-convergence. Une telle proposition découle de la logique purement statistique selon laquelle la moyenne est une mesure de tendance centrale très sensible à la présence éventuelle de valeurs extrêmes ou aberrantes. Cette proposition est illustrée par une analyse empirique des inégalités régionales en France à l'échelle des départements (NUTS 3) pour la période 2001-2015. Les résultats de l'analyse fournissent des indications précieuses tant sur le plan théorique que des politiques, dans la mesure où les différentes expressions du concept de σ-convergence peuvent conduire à des évaluations différentes en matière d'inégalités régionales.

Mots-clés
Inégalités régionales
Sigma-convergence
Coefficient de variation
Coefficient de variation pondéré
Moyenne arithmétique
Médiane