Universidad de Oviedo
Facultad de Economía y Empresa

Trabajo Fin de Máster en Economía: Instrumentos del Análisis

The role of spatial scale in convergence studies:

Exploring the consequences of MAUP in empirical analysis

Autor: Alberto Díaz Dapena

Tutores: Prof. Dr. Esteban Fernández Vázquez
Prof. Dr. Fernando Rubiera Morollón

Septiembre 2012

The role of spatial scale in convergence studies:

Exploring the consequences of MAUP in empirical analysis

Master´s thesis presented by:

Alberto Díaz Dapena

Supervised by:

Prof. Dr. Esteban Fernández Vázquez
Prof. Dr. Fernando Rubiera Morollón

Abstract

The empirical analysis of regional convergence is normally based on data collected at the administrative divisions of the territory level. However, as endogenous growth theories highlighted, large cities and central areas usually grow faster than rural and peripheral areas. So, it could be more realistic considering that the dynamics that generate economic growth take place at a smaller scale. These potential differences across sub-regional areas might be lost if the analysis is made at the NUTS II or III level (in the case of the EU countries), generating the so-called Modifiable Area Unit Problem (MAUP). The objective of this paper is to explore to which extent the MAUP bias could affect convergence analysis. By means of a Montecarlo simulation, we study how the beta-convergence results could be affected by using different levels of geographical aggregation. Several hypotheses about the generating process of local growth are considered in the simulation, according with the neo-classical and the New Economic Geography models, and the relevance of the MAUP is evaluated under different theoretical scenarios.

I. Introduction.

Since the contributions of Abramovizt (1986) and Baumol (1986), studies on the patterns of GDP per capita convergence among territories -nations or regions, have been one of the central issues in the economic literature. This interest is logical, because from the point of view of economic policy design as well as economic theory, finding empirical evidence is fundamental about where and under what conditions is possible to observe processes of economic convergence or divergence. For example, identifying clear patterns of divergence or very slow convergence among European regions provides the empirical evidence to support active regional policies, e.g., the European Union Cohesion Policy, which is now the most expensive policy in the EU budget. From a more academic point of view, the identification of divergence processes in different scenarios, even among territories with strong integration, implies finding an empirical evidence for rejecting neoclassical models of growth and evidence in favor of the theoretical framework depicted in the New Economic Geography (NEG).

Therefore, this is one of the subjects in which the conclusions have more relevant political and theoretical implications. Consequently, economics has put its attention at improving as much as possible the empirical models and the inference techniques of convergence analysis, which in turn has been leading to better estimation procedures or more detailed and complete modeling. Nevertheless, one point that does not receive all the attention that probably it deserves is wondering about if changes in the spatial level of disaggregation in the analysis could significantly affect the conclusions. This potential problem could emerge due to intra-regional effects that we cannot detect with aggregated information. When the conclusion of an analysis depends on the spatial scale we use, we have what in the literature is defined as a Modifiable Areal Unit Problem (MAUP).

This issue has been partially studied by Miller and Genc (2005). They apply standard convergence analysis for the U.S. economy from data aggregated at different scales according to different administrative and analytical divisions provided by the Bureau of Economic Analysis (BEA). In their paper, they find the same rate of convergence for almost every regional specification, being the biggest difference found between economic areas and a specification in states. Even when this is an interesting result, it could happen that these results cannot be automatically extended to other countries with bigger sources of divergence and a labor market defined by `rooted workers´ (without mobility).

The aim of out paper is to obtain more general conclusions and to measure the effect of small-scale process that can be hidden when data are studied at a more aggregated scale.

With this aim in mind, this paper is divided as follows. Section II makes a review about the literature of convergence. Section III describes different techniques and results that other authors have estimated for the specific case of Spain. Section IV explains briefly the theoretical –statistical- issue of MAUP. Section V explores the practical consequences of MAUP in convergence studies by conducting different numerical simulations. Finally, section VI closes the paper with some remarks and potential future research lines.

II. What is the expected result in convergence analysis?

The studies of convergence among regions are extremely relevant in order to design the appropriate regional policy. Convergence or divergence processes are theoretically possible within models of regional growth. This implies that empirical studies are required to check which hypothesis is supported by the observed data. If the outcome of an empirical analysis is that there is an underlying process of convergence, this means that the poor territories grow faster than the rich ones. The regions will converge to the same level of GDP per capita in the long run, so the differences between the regions are going to vanish or, at least, to reduce significantly with time. In other words, there is no need of redistribution policies across regions: we only need either to ensure that this result is achieved and speed the process if possible, or to balance asymmetric shocks that can change this result. Oppositely, if we obtain as outcome that a process of divergence is present, the differences between territories will be bigger along time and the role of the regional policy will be fundamental to push the development of poorer regions and reduce the differences across space.

As many authors argue (see, for example, Polese (2009) for a review) in the process of growth of a territory -regions, cities...- *centrifugal* and *centripetal forces* are always present. The set of *centrifugal forces* tend to push economic activity out from richer to poorer areas as a consequence of saturation, spreading economic growth across space (convergence). *Centripetal forces* act just in the opposite direction making richer territories more attractive than poorer areas and generating processes of concentration (divergence).

Neoclassical economics emphasizes how and why the *centrifugal forces* operate. Their explanation is fundamentally based on the existence and effect of decreasing returns in the different production factors. In the Solow's model (1956) we can see that there is convergence among economies when we have taken into account different relevant factors of the economy. We depart from an aggregated production function with constant returns to scale and accumulation of capital. We also assume a closed economy where all the agents save part of the production in a constant proportion (s) and the population growth is constant (n):

$$Y = K^{\alpha}L^{1-\alpha} \quad 0 < \alpha < 1 \tag{1}$$

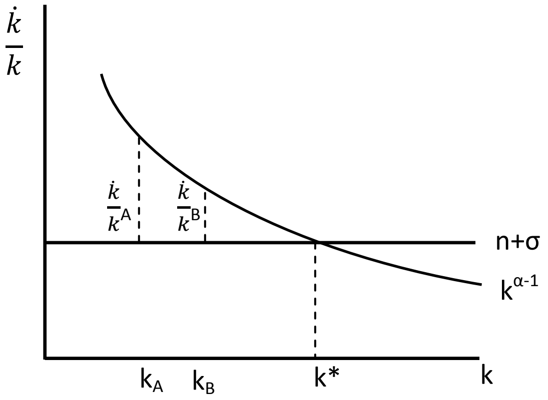
In terms of GDPpc: $Y/L = (K/L)^{\alpha}$ we can write it as y=k$^{\alpha}$ where $y = Y/L$ and $k = K/L$. In the steady state this variable does not grow, so $\frac{\dot{k}}{k} = \frac{\dot{K}}{K} - \frac{\dot{L}}{L} = 0$ where the accumulation of capital is given by the function $\dot{K} = sY - \sigma K$ and being σ the rate of depreciation.

Replacing $\dot{K}$ for its expression, we obtain $\frac{\dot{k}}{k} = \mathrm{s}\frac{Y}{K} - (\mathrm{n} + \sigma)$. We can simplify this expression with a division: $\frac{Y}{K} = \frac{Y/L}{K/L} = \mathrm{y/k} = \mathrm{k}^{\alpha-1}$ because $y = k^{\alpha}$. Finally, the growth of the different regions in the steady state is:

$$\frac{\dot{k}}{k} = \mathrm{s}k^{\alpha-1} - (\mathrm{n} + \sigma) = 0 \tag{2}$$

This equation is illustrated in Figure 1. It shows how every country tends to the same level of capital and income per capita because of the decreasing returns. Moreover, the economies which are further from the steady state grow faster than the closer ones. This means that they will follow a general pattern of conditional convergence. In the long run all territories will converge to the same level of GDPpc, as we have taken into account the relevant factors of an economy, like the rate of saving or the growth of the population.

Figure 1: Representation of convergence to the steady state.

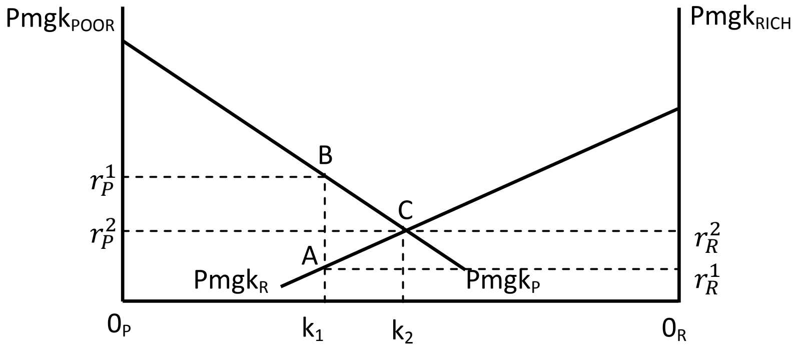


This approach was originally proposed for countries but can be extended to regions or smaller spatial units. However, territories that belong to the same country can be really similar in relevant factors. This fact justifies that sometimes it is assumed that all of them have the same relevant factors and, in consequence, an unconditional convergence process is analyzed.

This model also predicts that there is no growth of the GDPpc in the steady state. We can complete it with exogenous technological progress, which can explain this growth of the GDPpc with this exogenous factor. Empirical estimates of the Solow model generally obtain the expected signs, but the magnitude of these coefficients was not the expected by the theory. In order to improve the model, Mankiw, Romer and Weil (1992) add the human capital as a relevant factor of the economy. Moreover, there are other factors that have been emphasized. As De la Fuente (1996) explains, the spread of the knowledge and the reallocation of the resources of the economy from agriculture to industry or services are both mechanisms which generate convergence.

Barro and Sala (1991) also explain that, in models with labor mobility, factors go to the places with higher compensation rates. That movement of factors is considered an important centrifugal force in the Neoclassical Economics. For example, a model that assumes rich territories having larger proportions of capital per labor unit than poor regions, will predict that its price rises due to the decreasing returns produce by an intense use in capital. In other words, soothe relative abundance of capital in relation with the population that an additional machine would only increase the production a little. However, poorer regions have the opposite situation. As a result, the capital goes to the place where there is the lowest relation between capital and labor, and all the territories of the country tend to obtain the same level of well-being. The conclusion of this model is that there is no need of redistributional regional policies. We can illustrate that idea with Figure 2.

Figure 2: Representation of equilibrium with factors mobility.



In Figure 2, two regions are represented so that in the first one (*R*) there is more capital than in the second one (*P*). The starting point of this reasoning is the situation *A*. In the poor region the salary and the productivity of capital are bigger than in the rich one because the lack of this factor. As a result we move to the situation *B*, where these two regions obtain the same level of productivity. Surface of the triangle *ABC* measures the gains generated by this effect, because in this new situation the capital is used where it can produce higher output and, therefore, the society enjoys a better assignation of the factors than at the initial situation. In consequence, a process of convergence happens due to decreasing returns and the movement of inputs. This result can also be stronger if supported by additional elements, such as trade among territories.

Nevertheless, the empirical analysis of differences in GDPpc among countries or regions frequently find processes of divergence -or very slow convergence- that were not in line with this neoclassical theoretical framework. Alternative research lines contributed to understand how the other set or forces -*centripetal forces*- operate explaining why convergence is so moderate or directly nonexistent.

First, as already stated in the early works by Marshall (1920), they explain how much relevant the scale is (*economies of scale*). The gains derived from large-scale production and from the positive externalities associated with *size* lead to the concentration of economic activity in central locations from which the largest possible market is accessible. Consequently, those activities that (i) are tradable over broader distances, not requiring proximity to the point of consumption; and/or (ii) are less frequently demanded, will concentrate their production in a limited number of *central* locations. As distance costs fall and trade increases, larger concentrations should normally grow.

When a strong concentration of population and economic activity is formed in a particular place -a large city or metropolis- the externalities associated with *size* -*agglomeration economies*- start to work. There are two types of positive agglomeration externalities. First, the concentration of similar companies generates positive effects external to the company but internal to the industry that are named as *location economies*. Firms within the same industry benefit from lower recruitment and training costs (shared labor-force), knowledge spillovers, lower industry-specific information costs and increased competition (Rosenthal and Strange 2001, Beardsell and Henderson 1999, Porter 1990). Moreover, the increasing *size* of the metropolis makes certain infrastructures possible: international airports, post-graduate universities, research hospitals… the recent literature also stresses the positive link between productivity and the presence of a diversified, highly-qualified and versatile labor pool (Duraton and Puga 2002, Glaeser 1998, 1994, and Quigely 1998) in large cities. As highlighted by Hall (2000) and Castells (1996), large metropolises stimulate the exchange of knowledge. Activities that are characterized by the need for high creativity and innovation will in general choose to locate in major metropolitan areas or close to them. All this constitutes the second type of *agglomeration economies*: the *urbanization economies.*

If we identify large metropolis in which *scale* and *agglomeration economies* are strong in opposition with the small size places located far away from this large metropolis we can distinguish among *central* and *peripheral* areas, which is one of the essentials of the New Economic Geography (NEG) models. According to NEG literature: (i) there are incentives to concentrate the production largely in the central areas, and (ii); the intra-regional and inter-nations processes of specialization and commerce reinforce the processes of concentration and, in consequence, of divergence.

The *core-periphery model* illustrates an economy in the special case with two regions. Some assumptions need to be made in order to simplify this problem. In this economy we have two sectors: on the one hand there is a competitive agricultural sector with an exogenous part of the population; on the other hand there is a monopolistically competitive manufacturing sector with a labor force that moves to the region with biggest higher wage. With the correct units we can represent the total population of manufacturing workers as μ and the agricultural population as $1-\mu$. In addition to this, in the first region there is a share of manufacturing workers λ while in the second region it is $1-\lambda$. T represents the transport costs between both economies. If we add the salary of the agricultural sector as the numeraire, we finally have four equations for each region.

- Income equations are shown in (3). The total income is just the population in each sector, weighted by the share of the region and multiplied by its salary.

$$\begin{aligned} Y_1 &= \mu\lambda w_1 + \frac{1-\mu}{2} \\ Y_2 &= \mu(1-\lambda)w_2 + \frac{1-\mu}{2} \end{aligned} \qquad (3)$$

- The equations in (4) show the price index. The manufacturing goods in this model are subject to iceberg transport cost. As a result, it is necessary to multiply the price of the item by the transport cost in order to obtain the delivery price. According to this equation, a higher share of manufacturing implies a lower price in that region, ceteris paribus the salary. This is the forward link.

$$\begin{aligned} G_1 &= [\lambda w_1^{1-\sigma} + (1-\lambda)(w_2 T)^{1-\sigma}]^{1/1-\sigma} \\ G_2 &= [\lambda(T w_1)^{1-\sigma} + (1-\lambda)w_2^{1-\sigma}]^{1/1-\sigma} \end{aligned} \qquad (4)$$

- Nominal wage equations are shown in (5). These equations derive from the profit maximization of each producer. Assuming a similar price index, the higher the income in the nearest region, the higher its nominal wage. This is called the backward link.

$$\begin{aligned} \mathrm{w}_1 &= [\mathrm{Y}_1 G_1^{\sigma-1} + \mathrm{Y}_2 G_2^{\sigma-1} T^{1-\sigma}]^{1/\sigma} \\ \mathrm{w}_2 &= [\mathrm{Y}_1 G_1^{\sigma-1} T^{1-\sigma} + \mathrm{Y}_2 G_2^{\sigma-1}]^{1/\sigma} \end{aligned} \tag{5}$$

Proof: To obtain this equation we need the demand of a product in the region r, $q_r^m = \sum_s Y_s (p_r^M T)^{-\sigma} \, G_s^{\sigma-1} T$, and the profit of a producer, $\pi_r = p_r^m q_r^m - w_r^m(\mathrm{F} + c^m q_r^m)$. The producer chooses the price, ceteris paribus the price index of the region, so:

$$\frac{\pi_r}{p_r^m} = q_r^m + p_r^m \frac{\partial q_r^m}{\partial p_r^m} - w_r^m c^m \frac{\partial q_r^m}{\partial p_r^m} = 0 \quad \text{with } \frac{\partial q_r^m}{\partial p_r^m} = -\sigma \frac{q_r^m}{p_r^m} \tag{6}$$

$$\frac{\sigma - 1}{\sigma c^m} p_r^m = w_r^m \tag{7}$$

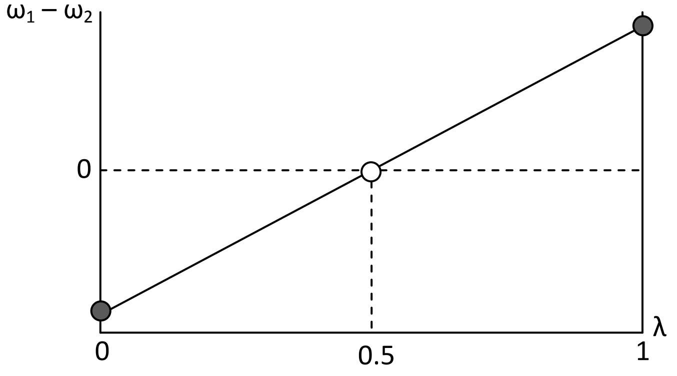
As the level of production with no profits has to follow $\mathrm{q}^* = \sum_s Y_s (p_r^M T)^{-\sigma} \, G_s^{\sigma-1}$, the price needed in this condition is $(p_r^m)^\sigma = \frac{\mu}{q*} \sum_s Y_s (T)^{-\sigma} \, G_s^{\sigma-1}$. Therefore, with the correct units we can normalize $\mu = \mathrm{q}^*$ and $c^m = \frac{\sigma-1}{\sigma}$ so the nominal wage is $w_r^m = [\sum_s Y_s (T)^{-\sigma} \, G_s^{\sigma-1}]^{1/\sigma}$.

- Finally, we have the real wage equations, which are represented in (8). The nominal wage is deflated here taking into account that the price of agriculture is one in all regions.

$$\begin{aligned} \omega_1 &= \mathrm{w}_1 G_1^{-\mu} \\ \omega_2 &= \mathrm{w}_2 G_2^{-\mu} \end{aligned} \tag{8}$$

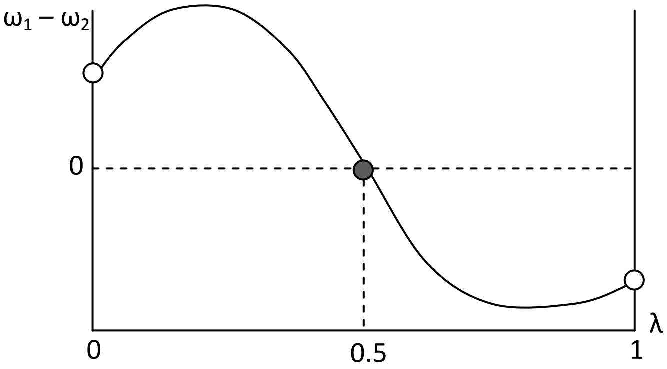
In order to understand the result of this model, the difference in real wage ($\omega_1 - \omega_2$) and the share of manufacturing (λ) are represented in Figures 3, 4 and 5. When the first economy is defined by a positive difference, the share of manufacturing grows due to the mobility of the labor force. As a consequence, $\omega_1 - \omega_2$ can be seen as a proxy of $\dot{\lambda}$.The results of this model can be grouped in three main cases.

Figure 3: Representation of core-periphery pattern.



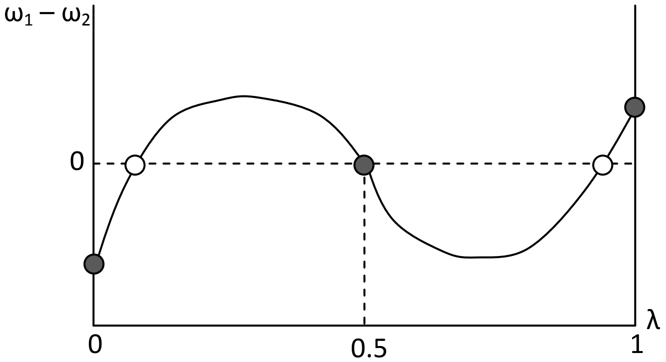
In this first scenario there is a low level of transport costs. Due to the backward and forward linkages, a concentration of the manufacturing in one economy generates a bigger salary and lower prices. As a result of the mobility, any situation slightly different from 0.5, which is an unstable equilibrium, generates an extreme concentration in one of the two economies.

Figure 4: Representation of symmetric equilibrium.



In this second case, there are high transport costs. This barrier between the two economies causes that any λ bigger than 0.5 is not attractive for the workers. The consequence is a symmetric equilibrium, with the industry divided between both regions.

Figure 5: Representation of the intermediate case.



Finally, an intermediate level of transport costs generates three stable equilibriums. The result depends on the initial situation of the two economies. When the initial situation is close to λ=0.5, the economy tends to the symmetric equilibrium. However, if the initial situation is far enough, the larger the concentration, the higher the positive difference in real salary. As a result , the labor force moves to the region with more industry, and a core-periphery pattern can be found. This model represents how the agglomeration economies generate concentration, according to Marshall (1920).[1]

---

[1] A non core-periphery pattern is possible as well: high transport costs can prevent concentrating in one region all the manufacturing activities. However, the weight of transportation costs over total cost of production has been gradually decreasing in many activities, being simply irrelevant for many industries.

In summary, theories that focus on explaining *centrifugal forces*, such the classical approaches or the NEG models, have a more local perspective than their neoclassical counterparts, which mainly focus on regional or even national analysis. This is motivated by the higher attention paid to factors that produce local endogenous growth. Under a NEG perspective, cities and metropoles (local areas) are put in the center of the analysis. NEG draws the attention to cities as the missing link between the macroeconomic theories of growth and the spatial empirical analysis.[2]

The basic hypothesis of this paper is that we should account for the spatial scale at which our empirical analysis takes place. The centripetal forces have an effect at the local level, so the aggregation into packages of regional data, which is normal in convergence analysis, can hide all this intraregional information. The possibility of this error cast doubts on the empirical evidence found in convergence studies. This is the main motivation in this paper to investigate the effect of aggregating data: the so-called MAUP issue.

III. Estimation techniques and some results.

III.1. Estimation strategies for convergence analysis.

Different procedures are applicable in empirical studies that aim to measure convergence. This section gives a only very general description of the basic versions of the most commonly used: the so-called sigma and beta convergence analysis.

Sigma convergence

When the approach followed in the empirical study is the sigma (σ)convergence, we aim at quantifying the dispersion or variability of GDPpc in different moments in time. Although different specifications are also possible, when a sigma convergence analysis is conducted at regional scale, GDPpc levels are normally measured in logs and compared with the GDP_{pc} in the whole country by means of the expression:

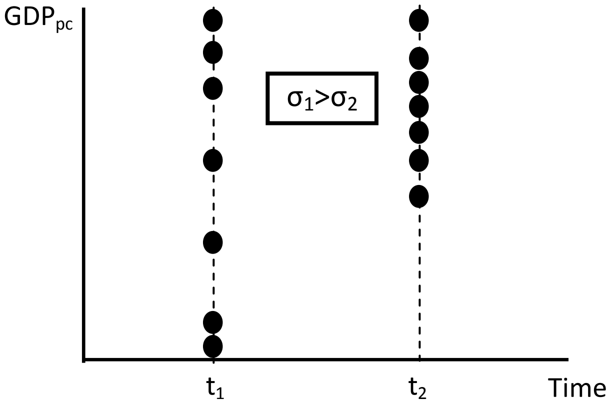
$$\sigma_t = \sqrt{\frac{\sum(lnGDPpc_{it} - lnGDPpc_t)^2}{n}} \qquad (9)$$

So, as a consequence, if σ_t decreases along time this will indicate that there is convergence. Oppositely, if σ_t goes up, this is a clear signal of divergence. Figure 6 shows the level of GDPpc at different regions of the same hypothetic country in two moments in time. It shows less variability in the second moment than in the first one, leading to the conclusion that an underlying process of convergence is present under the σ-convergence criterion.

---

[2] See, for example, the empirical analysis of Ciccone (1993) who found a positive relation between density and productivity.

Figure 6: Example of σ-convergence in two periods.



Beta convergence

Baumol (1986) introduced a new perspective of the convergence analysis using an alternative approach: a simple Ordinary Least Squares regression of the GDPpc growth rate in a territory on the initial level of GDPpc. So, this model is based in the following equation:

$$\frac{\dot{y}}{y_i} = \alpha + \beta ln y_{0i} + u_i \tag{10}$$

where $\frac{\dot{y}}{y}$ is the growth rate of GDPpc during a period of time of the *i* spatial unit and y_{0i} is the GDPpc in the initial moment of the period. When no other regressor is considered, we talk about an analysis of *unconditional* β-convergence, whereas if other explanatory variable is included we conduct a *conditional* β-convergence analysis. With this estimation framework we can see if poorer areas tend to obtain grow faster or not than the rich ones. If the parameter β is estimated with negative sign, this indicates that lower levels of GDPpc produce higher growth rates, leading to a process of convergence in the long-run. A positive estimate of β would reveal a process of divergence.

As an alternative to GDPpc, Baumol (1986) estimated an equation like (10) but considering output per worker for a dataset of industrial countries. He regressed productivity growth from 1870 to 1979 on labor productivity in 1870. The result is shown in equation (11):

$$\frac{\hat{\dot{y}}}{y_{i(1870-1979)}} = 5.25 - 0.75 \mathrm{lny}_{\mathrm{i}(1870)} \tag{11}$$

Baumol estimated (11) in his dataset and obtained a result of unconditional convergence. However, a conditional convergence analysis is possible if we include information of the *relevant factors* which can explain the steady-state of an economy (Barro and Sala, 1991). According to a Solow model, these factors could be the percentage of savings, the population growth or the technologic growth. When we do not have data of these relevant factors, we can overcome this lack of information with panel data: we can use a constant term for each region in a model with fixed effects to calculate the influence of these relevant factors. A result of conditional convergence has a different conclusion for regional policy. In this case, there is convergence when we take into account the relevant factors of the different economies, so the regional policy should change them in the poor regions. Alternatively, we can group sets of regions and estimate a dummy variable for each group instead of each region, as Dolado, Gonzalez-Paramo and Roldan (1994) proposed. In this case we assume that there are shocks which are common for all the regions of the same group.

As stated previously, the description of empirical tools for convergence analysis we presented here is far from being exhaustive. There are other ways to measure convergence than the ones depicted in this section. For example, the γ-convergence focuses on relative rankings of the GDPpc of the territories. The stochastic convergence conducts a time series test for unit roots, which make possible to check if there are persistent differences in the series of GDP. However, the σ and, specifically, the β-convergence analysis are the most commonly applied approaches in empirical studies on regional convergence. This is the reason why in the subsequent sections of this paper we will limit our discussion to traditional β-convergence analysis.

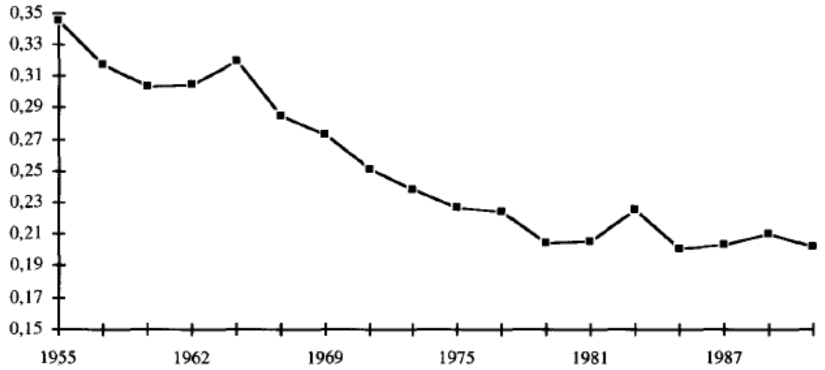
III.2. A selection of results in previous empirical analysis at regional level.

There is some consensus on the idea that a 2% in rate of convergence is almost a ‘magic number’ in the literature. In order to see it we can use some examples of international researches with cross-section information. Barro and Sala (1994) found a 3% of convergence in the U.K. (1950-1990), 1.7% in U.S.A. (1880-1990), 1.6% in France (1950-1990), 1.4% in Germany (1950-1990) and 1% in Italy (1950-1990). In addition to this, other authors also find similar results for other countries: Columbe and Lee (1993) found a 2.4% for Canada (1961-1991) and Shioji (1992) estimated a 1.9% for Japan (1955-1990). In other words, all they conclude that generally poor regions grow faster than the rich ones. Moreover, the regions in all the samples needed about 35 years to reduce to one half the initial differences in income. The relationship between the β-convergence and the rate of convergence (λ) is shown in (12):

$$\beta = \frac{-(1-e^{-\lambda T})}{T}, \lambda = \frac{-ln(1+\beta T)}{T} \qquad (12)$$

For the Spanish case, probably the most prominent work is De la Fuente (1996). This paper studied convergence at regional NUTS II level. In his work, he conducts initially a σ-convergence that shows an important reduction in the differences between regions during the period 1955-1991. As we can see in Figure 4, the main reduction of the divergence took place especially in the period 1955-1981, while in the second period the process of convergence stops.

Figure 4: σ-convergence of Spanish Autonomous Communities (1955-1991).



Source: De la Fuente (1996).

Assuming that all the regions of Spain had similar levels of the same relevant factors, an unconditional beta-convergence equation like (10) was estimated as well, obtaining as result a λ coefficient of 0.0295 (a convergence speed of 2.95%).

A conditional convergence analysis is possible by regressing some indicator of economic growth on the relevant factors of the economy, or with a model with regional fixed effects if information by region of the steady state is not available. In De la Fuente (1996), the fixed effects were always significant and the estimate for beta convergence coefficient was 0.1273. These results indicate that Spain is defined by a phenomenon of conditional β-convergence. A summary of other convergence analysis for Spain is presented in Table 1.

Table 1: Summary of speed of convergence in analysis for Spain.

	Unconditional	Conditional
Cuadrado., García-Greciano and Raymond (1999); 17 regions 1955-1993	0.032 (-6.481)	0.107 (-5.999)
De la Fuente (1996); 17 regions 1955-1991	0.0295 (4.78)	0.1273 (6.23)
Dolado, González-Páramo and Roldán, (1994); 50 provinces 1955-1989	0.0199 (5.527)	0.0443 (7.91)
Goerlich, Mas and Perez (2002); 17 regions 1955-2000	0.0113 (-9.043)	

Note: t-Statistics in brackets.

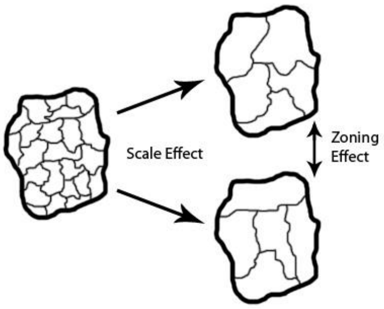
IV. The Modifiable Areal Unit Problem: (MAUP) a general introduction and some potential effect on convergence analysis.

A pervasive statistical problem when dealing with data available at some specific spatial scale: could the changes in this geographical scale be affecting the conclusions and inferences in our analysis? If the answer is yes, we call this kind of problems *Change of support problems* (Gotway and Young, 2004), which comprises several specific measurement issues. Within this set of potential problems, we will focus on specifically on *the Modifiable Areal Unit Problem* (MAUP).

According to the seminal work by Openshaw (1983), the object of the analysis in empirical studies conducted at some spatial scaleshould be described before anyone tries to measure its characteristics. However, the situation with real data is hard to achieve: the region exists only after the data are collected. As a consequence, the definition of the object is arbitrary and it can be changed. This author studied the relationship between the percentage of Republican and elderly voters with different territory aggregations. Surprisingly, the range of the possible correlation coefficient goes from -0.99 to 0.99.

The MAUP can be divided in two parts. Firstly, we have the *scale effect*, which is the most important part of the problem. This aggregation bias appears if we aggregate our data into larger units, for example cities to regions. Secondly, there is a *zoning effect*. In this case, we have a problem with the shape of the units because a different form can also change our results.

Figure 5: Illustration of the Modifiable Areal Unit Problem (MAUP).



We have seen that some authors found convergence in their regional databases. The main problem is that they find convergence with information which has been aggregated. As a consequence, the new aggregated variable does not have to maintain the characteristics of the original variable, especially when we do not aggregate following some economic but purely administrative division. If a MAUP issue is present, this would imply that our results of convergence would not be the same if our data were group into a different group of regions or directly observable at a smaller scale.

This problem has been partially explored in the paper by Miller & Genc (2004). In their paper, they argue that convergence processes are movements of economic, not political, units. This is the main motivation for their analysis, which quantifies the convergence between several possible spatial divisions for the US: states, economic units defined by the BEA and counties. Although they do not obtain significant differences (the biggest one is between BEA economic areas and states) this research highlights the possibility of different results using administrative divisions, economic-based regions and data at a more disaggregated spatial scale.

V. The effects of MAUP on convergence analysis: some numerical experiments.

An interesting research question is if the conclusions of convergence studies could be affected by the geographical scale used in the analysis. If, according to NEG, the process of economic growth is generated at local (city) level, dealing with information that is aggregated to some administrative regions could be hiding the real underlying process. The case studied in Miller & Genc (2004) for U.S. sheds some light on the topic, but it limits to a specific case for a particular time period. It would be interesting, in consequence, to extend the analysis to a more general framework.

The purpose of this section of the paper is to do that by means of some numerical simulations. The point of departure is assuming an economy that is divided into different spatial units that are created according to several criteria for geographical aggregation. More specifically, we assume that the economy is divided into i=1,…,N basic spatial units (municipalities or cities) that are aggregated into j=1,…,M ($M < N$) groups (regions). In line with the ideas of NEG and endogenous growth theories, we assume that the income generation process takes place at the basic spatial scale of N units. Then, we study the consequences on the conclusions of convergence analysis depending on the scale at which the outcome data are observable: directly observable at the original scale (N cities) or at the aggregated scale (M regions).

We can wonder if the aggregation of basic N spatial units implies wasting valuable internal information. If the conclusions about the presence of convergence or divergence depend on this level of aggregation, this will be a clear signal that a potential MAUP is somehow ‘contaminating’ our analysis. For example, we could find regional convergence with data at regional level, but we cannot conclude against the hypothesis that the capital cities and/or central areas within each region attract all the activity at expense of the rural areas, due to internal or external scale economies. Note that this is important in terms of regional policy design, since these disparities will be neglected and hidden with the aggregation of data. In other words, the convergence analysis would be relying on observations at a spatial scale that, following NEG, is not correct.

In order to do that, we impose an equation of beta convergence at the local scale with N units that generates the conditions in initial and the final time periods considered (0 and t). We consider two numerical experiments: an introductory one with output or income levels and a more sophisticated case that simulate per capita values under different scenarios of agglomeration economies and population distribution.

Simulating levels

First we simulate the process of output or income (Y) generation in levels. In the initial time period we assume that Y distributes according to an uniform distribution $Y_0 \approx$ U[0,5000] that generates the initial i=1,…,N random values. For time period t, we impose the generating process written below:

$$\frac{\dot{Y}}{Y_i} = \alpha + \beta \ln \mathrm{Y}_{io} + \mathrm{u}_i \qquad (13)$$

The first component α is the constant of the model. In our simulation we use a value of 0.1. In the second component, the β coefficient indicates convergence when negative and divergence when positive. For this coefficient we will impose different values between -0.05 and 0.05 in order to see the result for each case. Finally, the last component is a noise component which we suppose is drawn from the distribution u $\approx$ N[0,0.2]. With this equation we make 100 trials for each case and we estimate the β coefficient distinguishing two possible scenarios: (i) only aggregated information in M regions (M<N) is available and (ii) information of the N locations is observable. Note that the aggregated values for each region j are obtained by the sum $\sum_{i\in j} Y_i$.

For the sake of simplicity, we can use Spain as an example. According to the most recent Population and Housing Census, Spanish territory is divided into N=8,106 basic units (municipalities) that are aggregated into M=50 provinces (NUTS III level), which is the most disaggregated spatial scale at which output and income data are available for the whole country. This NUTS III administrative division does not follow an economic rule but it is just a political decision. It is easy to see that it is arbitrary and in theory modifiable if you compare our administrative division with other countries of Europe, for example, Germany. The aggregation of 8106 municipalities into 50 provinces shows us the scale effect of the MAUP. This effect would be further explored assuming that we depart from N=1,000 that are uniformly grouped into the M=50 provinces. We could also see differences on the estimation between grouping the N=8106 municipalities into the actual Spanish administrative distribution and an alternative uniform division always keeping M=50, which would be produced by the zoning effect. In summary, this first experiment considers three scenarios for the simulation

i. N=1,000 basic units grouped into uniform M=50 regions (i.e., with 20 spatial units in each case),
ii. N=8,106 basic units grouped into uniform M=50 regions (i.e., with $\simeq$ 162 spatial units in each case),
iii. N=8,106 basic units grouped into M=50 regions according to the actual Spanish administrative division.

The results of the first simulations are presented in Table 2. First column shows the real β-convergence parameter imposed in equation (13). Note that the positive values are cases where a divergence process is assumed in the generation of output and the opposite for the negative values of β. In each case, we report the percentage of cases when we conclude that there is an underlying process of convergence because our OLS of β is negative and statistically significant (based on the t-ratio). The percentage of cases where divergence is detected is the amount of trials in the simulations when the OLS estimate is significantly positive. These percentages are presented when (13) is estimated by OLS with the aggregated M and the original N spatial units.

Table 2: Output simulation results with 1,000 locations grouped uniformly.

	Provinces (M=50)		**Spatial units (N=1,000)**	
Real β	% Convergence	% Divergence	% Convergence	% Divergence
0.05	*24*	70	*0*	100
0.04	*28*	66	*0*	100
0.03	*30*	64	*0*	100
0.02	*34*	59	*0*	100
0.01	*39*	52	*5*	88
-0.01	48	*49*	86	*4*
-0.02	50	*42*	100	*0*
-0.03	55	*38*	100	*0*
-0.04	60	*31*	100	*0*
-0.05	68	*26*	100	*0*

The percentage with wrong sign in the diagnostics, (i.e., we detect convergence when the true process is divergent and the other way round) are marked in red italics and the right ones in black. The estimated convergence or divergence with local data has the correct sign, although there are some mistakes with a 0.01 and -0.01 due to the random error which is introduced in the simulation. However, larger errors are found for all the imposed betas with aggregated data. For example, if we assume a beta-convergence of a -0.02, the percentage of cases with convergence and divergence are really similar.

The second case studied in this first simulation bases on the actual number of municipalities of Spain. We assume an aggregation of N=8,106 into M=50 regions following a uniform distribution. Its results are presented in Table 3. As in the previous case, an important error which is introduced by the aggregation causes a lower percentage of correct estimations than a regression with the original N data points. Not surprisingly, in this new scenario the regressions with the original N data points have increased their power of hypothesis testing by the increase in the sample size. In spite of this, the results with aggregated data are almost the same. So, this first change in the simulation indicates us how many information can be wasted due to the aggregation level.

Table 3: Output simulation results with 8,106 locations grouped uniformly.

	Provinces (M=50)		**Locations (N=8,106)**	
Real β	% Convergence	% Divergence	% Convergence	% Divergence
0.05	*26*	65	*0*	100
0.04	*32*	63	*0*	100
0.03	*35*	59	*0*	100
0.02	*38*	54	*0*	100
0.01	*44*	50	*0*	100
-0.01	55	*41*	100	*0*
-0.02	58	*34*	100	*0*
-0.03	64	*28*	100	*0*
-0.04	72	*25*	100	*0*
-0.05	75	*22*	100	*0*

In the final scenario of this first simulation we consider the actual Spanish administrative division instead of a uniform aggregation. As a consequence, this experiment is the closest version to the Spanish case. The results are shown in Table 4.

Note that the aggregation problem is the same as in the previous scenario, i.e., we control for the scale effect. This third scenario, if compared with the second, illustrates how the zoning effect aggravates the problems, making the percentage of correctly identified significant cases even lower than before. Note that, if the aggregated M regions are the only available information, the picture of the economy obtained is extremely confusing even for a real beta of 0.05, because the number of cases with convergence or divergence is really similar.

Table 4: Output simulation results with 8,106 locations grouped by the Spanish administrative division.

	Provinces (M=50)		**Locations (N=8,106)**	
Real β	% Convergence	% Divergence	% Convergence	% Divergence
0.05	*25*	34	*0*	100
0.04	*26*	33	*0*	100
0.03	*26*	33	*0*	100
0.02	*28*	31	*0*	100
0.01	*28*	28	*0*	100
-0.01	31	*26*	100	*0*
-0.02	31	*26*	100	*0*
-0,03	33	*24*	100	*0*
-0.04	35	*24*	100	*0*
-0,05	35	*24*	100	*0*

Simulating per capita values

In this second simulation we try to replicate the more common problem of convergence analysis with per capita values instead with the values in levels. The point of departure is a per capita indicator (y) of output or income obtained as $y = Y/L$, where L stands for population.

Taking again Spain as reference, we depart directly from the case of N=8,106 basic units (municipalities) that are aggregated into M=50 provinces following the real administrative division –i.e., case (iii) in the previous experiment-. We generate population for each location (L_i) and assume that it keeps constant along the time span considered. The population is located into the N=8,106 units following two possible scenarios:

a) According with the actual data in the 2001 Population Census

b) Considering that in each one of the M=50 regions there is a big city with 1,000,000 inhabitants and assuming that the population adjusts to a Zipf law:

$$L_i = 1{,}000{,}000/rank(i) \qquad (14)$$

Additionally, we also assume that the impact of population size on the initial value of the per capita indicator (y_{i0}) adjusts to the following equation:

$$ln(y_{i0}) = \mu + \theta ln(L_i) \qquad (15)$$

Parameter μ represents a random term that determines the initial per capita value on each location that draws from a uniform distribution $U[0{,}20000]$. The parameter θ quantifies the economies of scale, reflecting the effect of municipality size on y_{i0}. We assume two possible cases in (15):

c) Parameter $\theta = 0$, which means that the initial income per capita is completely random.

d) There are some economies of scale in the initial generation of the income per capita. In a recent survey, Mello et al. (2009) found that different empirical estimations for several set of cities in different countries elasticities of scale ranging from 3% to 9%. In order to set a sensible value, we have assumed an intermediate case such $\theta = 0.05$.

Taking all possible combinations of cases (a) to (d) the dynamics of the per capita indicator for the i=1,..,N follows the expression:

$$\frac{\dot{y}}{y_i} = \alpha + \beta \ln \mathrm{y}_{io} + \mathrm{u}_i \tag{16}$$

Being the rest of parameters identical to the first simulation. Now, we have made 200 trials for each combination estimating again the β coefficient with: (i) aggregated information of the M regions (M<N) and (ii) information of the N locations is observable. Now that the aggregated per capita indicators for each region j are obtained by the weighted sum $\sum_{i \in j} [y_i L_i]/L_j$.

The four tables with the estimates, which follow the same structure as in the first simulation, are shown below. Table 5 and 6 report the results using the actual data of the Spanish population by municipalities in 2001.

It is remarkable that in this first and second scenario with aggregated data the percentage of wrong sign in the diagnostic is not symmetric. The proportion of mistakes when a positive real beta is imposed is higher than with a negative beta and, what is more, the errors percentage exceeds the rights one. As a result, there is not only a confusing image of the real process of convergence in this second simulation, but also the results are biased to convergence.

Table 5: Per capita simulation results with 2001 census population ($\theta = 0$).

	Provinces (M=50)		Locations (N=8,106)	
Real β	% Convergence	% Divergence	% Convergence	% Divergence
0.05	*44.43*	42.73	*0.00*	100.00
0.04	*47.27*	42.23	*0.00*	100.00
0.03	*48.27*	40.91	*0.00*	100.00
0.02	*48.77*	40.01	*0.91*	92.73
0.01	*49.09*	38.91	*16.36*	67.27
-0.01	52.72	*36.36*	78.18	*13.64*
-0.02	53.73	*34.55*	90.40	*5.46*
-0.03	56.36	*33.64*	99.09	*0.00*
-0.04	56.55	*33.64*	100.00	*0.00*
-0.05	56.96	*32.72*	100.00	*0.00*

Table 6: Per capita simulation results with 2001 census population ($\theta = 0.05$).

	Provinces (M=50)		Locations (N=8,106)	
Real β	% Convergence	% Divergence	% Convergence	% Divergence
0.05	*48.18*	37.27	*0.00*	100.00
0.04	*49.09*	36.78	*0.00*	100.00
0.03	*49.99*	34.54	*0.00*	100.00
0.02	*50.05*	32.73	*0.91*	93.64
0.01	*51.82*	32.23	*17.27*	68.18
-0.01	56.36	*30.91*	79.09	*15.46*
-0.02	57.27	*30.01*	90.01	*3.64*
-0.03	58.18	*29.09*	99.09	*0.00*
-0.04	59.09	*28.18*	100.00	*0.00*
-0.05	61.82	*26.36*	100.00	*0.00*

Table 6 shows the results when an assumption of agglomeration economies is imposed instead of random distribution of the GDP per capita. In this case, the percentage of significant convergence goes up for every imposed beta convergence. This difference of 2.77 percentage points in mean could indicate us that the agglomeration economies make the MAUP problem even more important. Nevertheless, note that the main conclusion of an important bias to convergence is not affected by the different assumption of agglomeration economies.

Finally, the results assuming that population follows the Zipf law depicted in (14) are reported in Tables 7 and 8.

Table 7: Per capita simulation results with population according to the Zipf law ($\theta = 0$).

	Provinces (M=50)		Locations (N=8,106)	
Real β	% Convergence	% Divergence	% Convergence	% Divergence
0.05	*55.24*	36.19	*0.00*	100.00
0.04	*56.19*	35.19	*0.00*	100.00
0.03	*57.14*	35.24	*0.00*	100.00
0.02	*57.14*	35.24	*0.95*	93.33
0.01	*58.10*	33.33	*15.24*	67.62
-0.01	60.20	*33.33*	78.10	*13.33*
-0.02	60.95	*32.38*	90.48	*4.76*
-0.03	62.86	*31.43*	99.05	*0.00*
-0.04	63.81	*31.43*	100.00	*0.00*
-0.05	63.88	*30.48*	100.00	*0.00*

Table 8: Per capita simulation results with population according to the Zipf law ($\theta = 0.05$).

	Provinces (M=50)		Locations (N=8,106)	
Real β	% Convergence	% Divergence	% Convergence	% Divergence
0.05	55.29	38.82	0.00	100.00
0.04	57.14	36.19	0.00	100.00
0.03	58.10	35.24	0.00	100.00
0.02	60.00	34.29	0.95	94.29
0.01	61.25	33.33	15.24	70.48
-0.01	61.91	33.33	78.10	14.29
-0.02	61.92	32.38	90.48	4.76
-0.03	62.86	31.43	99.10	0.00
-0.04	64.76	30.48	100.00	0.00
-0.05	65.71	29.52	100.00	0.00

The assumption of a population distributed according to the Zipf law gives us other scenario, in order to see how the results change when the different assumptions are modified. In this new case, Table 7 shows that a bias to convergence is also found. Moreover, this bias is even higher when we use the Zipf law assumption. In consequence, the percentage of cases where we conclude convergence with a positive real beta-convergence is always at least 19 percentage points than the correct one.

The simulation results with the agglomeration economies are presented in Table 8. In this last scenario a bias to convergence is also reported. It is increased due to the assumption of agglomeration economies, although it is small (1.34 percentage points in mean). As in the previous case, this bias is bigger when the Zipf law is used. It indicates that our results are strong enough to overcome different changes in the simulation assumptions.

VI. Conclusions of the MAUP effects on convergence analysis

This research has developed a simulation approach in order to discover how the MAUP problem can affect the results in the convergence analysis. Firstly, a simple simulation of the income levels illustrates that there is an important problem of efficiency. It also shows how studies with aggregated data do not take advantage of bigger sample sizes at local level. Moreover, the results are less statistically significant when the aggregation is made with the actual Spanish administrative division.

This lack of precision in the simulated result with aggregated data could be explained by differences in sample size. While the local information provides the researcher with 8,106 observations, the aggregation into Provinces reduces the sample to 50 observations and many researchers use only 17 NUTS II regions. As a result, an inefficiency problem emerges. Trying to explore further this issue, a beta-convergence equation with per capita incomes is evaluated and additional assumptions are used in order to aggregate the result with a weighted mean. In this case more serious problems are detected. There is lack of efficiency and a bias towards convergence when the data are aggregated. Moreover, this bias is amplified by the economies of agglomeration. This effect is generated because this assumption creates intra-regional disparities which are not observable with aggregated data. As a consequence, it is more difficult to detect the real pattern of convergence.

This simulation has also been done with a population according to the Zipf’s law. This change introduces more homogeneity between regions, which are now defined by a very similar population structure, but much higher intra-regional differences that are not considered when aggregated data are observed. As a result, the bias to convergence is even bigger than in the simulation with the Spanish 2001 census.

In this research the MAUP problem in the convergence analysis has been introduced. Our (preliminary) main conclusion is that the beta convergence analysis made with aggregated data should take into account that an important part of the information is missing. This lack of sufficiently disaggregated information can generate an important problem of bias and inefficiency due to relevant intra-regional dynamics, which should be considered in order to provide a better description of the economy.

However, important issues not studied in this research are still pending. For instance, this simulation has studied the MAUP assuming that the municipal division is the ideal spatial unit for the convergence analysis. But other authors can think that this areal choice can be modifiable to different scales such as functional regions or local labor markets. Furthermore, the assumptions made in the simulation can be changed, which, even when they seem robust to the different scenarios simulated, could change the conclusions obtained.

The important consequences that could be derived from this research in terms of regional policy are straightforward. Centrifugal forces can generate an empty periphery with a low level of well-being while a few big cities absorb all their activity. So, the efforts to reduce differences between territories should be careful if they rely in convergence analysis based on aggregated information.

These simulations also open a path to investigate how potential MAUP issues could be having some effects on the conclusion of previous research. More specifically, it could be interesting to replicate previous convergence studies conducted for the Spanish regions if this analysis was based on data observable at a more disaggregated spatial scale. This, in the one hand, raises the problem of the lack of data on output or income at the sufficiently disaggregated scale. But in the other hand, it opens new research opportunities for developing and applying statistical inference techniques that can produce reliable estimates from actual observable information.

References

Abramovitz M. (1986) Catching up, forging ahead, and falling behind, Journal of Economic History, 46, pp. 385-406.

Barro, R.J., Sala i Martin, X. et al (1991): “Convergence across states and regions”, Brookings Papers on Economic Activity, nº 1, pp. 107-182.

Baumol, W.J. (1986): “Productivity Growth, Convergence, and Welfare What the Long-run Data show”, The American Economic Review, nº 76, pp. 1072-1085.

Cuadrado, J.R., García-Greciano, B., Raymond, J.L., (1999): “Regional Convergence in Productivity and Productive Structure: The Spanish Case”, vol 1, nº 22, pp. 35-53.

De la Fuente, A. (1996): “Economía regional desde una perspectiva Neoclásica”, Revista de Economía Aplicada, nº 10, pp. 5-63.

De la Fuente, A. (2002): “On the sources of convergence: A close look”, European Economic review, nº 46, pp. 569-599.

Dolado, J.J., González-Páramo, J.M., Roldán, J.M. (1994): “Convergencia económica entre las provincias españolas: Evidencia empírica (1955-1989)”, Moneda y Crédito, nº 198, pp. 81–131.

Goerlich, F.J., Mas, M., Perez,F. (2002): “Concentración, convergencia y desigualdad regional en España”, Papeles de Economía Española, nº 20, pp. 7-28.

Gotway, C.A., Young, L.J. (2004): “A Spatial View of the Ecological Inference Problem”, Ecological Inference New methodological strategies, ch. 10, pp. 233-244.

Mankiw, N.G., Romer, D., Weil, D.N. (1992): “A contribution to the empirics of economic growth”, The Quarterly Journal of Economics, Vol 107, Nº 2, pp. 407-437.

Martin, R., Sunely, P. (1998): “Slow convergence? “The new endogenous growth theory and regional development”, Economy Geography, nº 74, pp. 201-227.

Masahisa, F., Krugman, P., Venables, A.J. (1999): “The spatial economy: cities, regions and international trade”, ch. 5, pp. 61-76.

Melo, P. C, Graham, D. J. and Noland, R.B. (2009): “A meta-analysis of estimates of urban agglomeration economiest”, Regional Science and Urban Economics, nº 39 (3), pp. 332-342.

Miller, J.R., Genc, I. (2005): “Alternative regional specification and convergence of U.S. regional growth rates”, The annals of Regional Science, nº 39, pp. 241-252.

Openshaw S. (1983): “The modifiable areal unit problem. Concepts and techniques in modern geography”, GeoBooks, Norwich. nº 38