Heterogeneity in the determinants of national *versus* foreign population growth at local level

Viñuela, Ana (avinuela@uniovi.es)
Rubiera-Morollon, Fernando
Gutierrez Posada, Diana

Departamento: REGIOlab, Laboratorio de Analisis Economico Regional. Departamento de Economia Aplicada

Universidad: Universidad de Oviedo

Área Temática: *Poblacion y movimientos migratorios*

Resumen: *(máximo 300 palabras)*

Palabras Clave: Formal modeling of local population growth has usually tended to focus on identifying patterns that are presumed to hold universally. Although reliable for long-term dynamics, in some moments or for some places the balance between the different factors can change, giving rise to different specific behaviors (Glaeser *et al.*,2014). Moreover, not only can behavior change across time and space but also between distinct population groups. The aim of this paper is to study the determinants of local population growth in Spain trying to identify the existence of: i) spatial heterogeneity, ii) temporal heterogeneity, and iii) heterogeneity between groups. While in the 1990s local population increases or decreases were almost exclusively due to the natural rate of population growth and/or internal migration flows, the 2000s were characterized by strong inflows of migration and the total population moved from 40 million to more than 45 million inhabitants. This massive arrival of immigrants had consequences at local level as they did not tend to locate evenly across space. Population growth was therefore heterogeneously distributed across space mainly due to the uneven location of immigrants. Why did the immigrants establish in one place and not in another? What are the *local* factors explaining immigrants' location and therefore local population growth? Can some determinants act as a deterrent in some place but promoting settlement and local population growth in another? And can some factors

have a positive effect on the population growth of foreigners but the opposite effect on nationals?
In order to answer these questions, a Geographically Weighted Regression (GWR) estimation procedure is applied for two different decades (1991-2001 and 2001-2011) and for two different groups (nationals *versus* foreigners).The GWR is a non-parametric model that captures spatial variations in the regression coefficients by introducing a weighting matrix in the estimation process and estimating a locally varying sample for each location. A separate group of regression parameters is generated, reflecting the sample heterogeneity by estimating different responses to an explanatory variable across space. Results from the GWR approach will be therefore compared with the global estimators (OLS) for both decades and groups. Highly relevant factors in Urban and Regional Economics such as *size* and *distance* seemingly have different effects on population growth both across space and time, corresponding to the global estimated effects for some areas but not so for others. Also, foreign population growth seems to follow a different pattern to national population growth. Differentiated results at local level support the need for designing policies “a la carte” in order to enhance - or contain, depending on the case- *local* population growth as some determinants could enhance population growth for one place or group but reduce for others.

Clasificación JEL: R23, J11 and C19

1. Introduction

Identifying the determinants of population growth or decline is one of the most relevant issues in the social sciences. Knowledge of which local or regional factors influence population growth can help policymakers to anticipate demographic tendencies for certain area, determine a territory’s future prospects and to conceive specific policies to promote population establishment in another.

Spain has experienced extraordinarily high population growth over the last two decades, from 38,494,033 in 1991 to 40,454,729 in 2001 to 46,639,516 inhabitants in 2011 according to the Censuses released by the National Statistics Institute (INE) every ten years (Table 1). While in the first decade population growth was basically due to the due to internal migration movements and net birth-death adjustments. the 2000s growth was mainly due to the big inflow of immigrants, especially from Southern American and Northern African countries. Thus, according to the Censuses, foreign population multiplied by four in just one decade (2001 to 2011)

Table 1: Population and population growth: 1991, 2001 and 2011

	Population			*Average growth (%)*		
	National	*Foreigner*	*Total*	*National*	*Foreigner*	*Total*
1991	38,152,728	341,305	38,494,033			

2001	38,915,727	1,539,002	40,454,729	0.20	16.25	0.50
2011	40,369,615	6,269,901	46,639,516	0.37	15.08	1.43

Source: 1991, 2001 and 2011 Spanish Census (INE)

It seems to be a stylized fact that migration population tends to be geographically concentrated. For instance, in the year 1990, 63% of the foreign-born population were clustered in the four most populous states (California, New York, Florida and Texas) where only 31% of the overall population lived USA (Zavodny, 1999). In Spain, for the period 2000-2006, 55% of the foreign born population concentrates in three autonomous communities or NUTS 3 level regions: Madrid, Cataluña and Comunidad Valenciana (Conde-Ruiz et al, 2008)

However, as the three bigger metropolitan areas in Spain -namely Madrid city, Barcelona and Valencia city- are located in those regions, it could be argued that it is not the *autonomus community* itself that is the attracting force, but the *city* and its surrounding areas. In other words, the metropolitan areas and the big cities – the core - might be the ones offering new job opportunities, mostly service sector jobs, prompting the arrival of new workers (Bover and Velilla 1999). The ability of cities to attract international migrants is therefore increasingly seen as an important indicator of their growth potential (Glaeser and Resseger, 2010; Glaeser and Saiz, 2003; Moretti, 2012). Evenmore, when differentiating between highly and low skilled workers, factors such as urban amenities and attractions can affect in different ways their location choices. (Levkovicha and Rowendala, 2014). The appeal of an area can be perceived in a different way among different groups.

Thus, it is necessary to descend at more des-aggregated level in order to detect what *local* factors contribute to attract immigrants to certain *city* (not region) or *local area*. At more des-aggregated level in Spain the foreign-born population location seems to follow a more complex pattern. In 2011 the big metropolitan areas of Madrid and Barcelona concentrated 28.33% of the foreign-born population and only 19.42% of the national population lived there (Table 2). However, if considering the next level of metropolitan areas in the urban hierarchy (those ones with less than 2 million inhabitants but more than 500,000), we can realize that the presence of foreign-born population diminishes is in general lower than the equivalent for the national population.

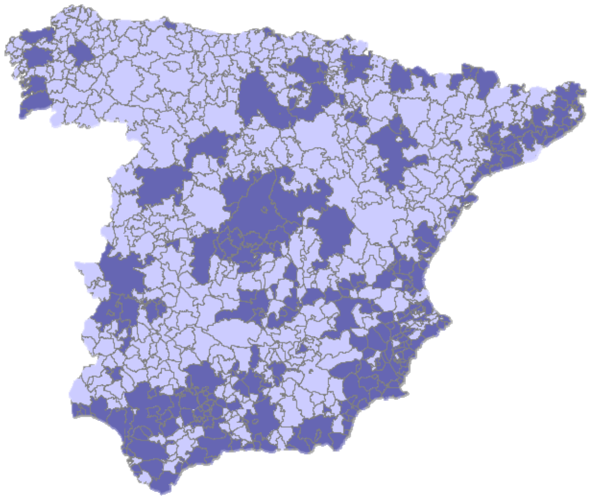
Table 2: Geographical distribution across LLMs (percent): 2011

	National Population (%)	Foreign Population (%)	Total Population (%)
Big Metropolitan Areas (more than 2M inhabitants)			
Madrid	12.64	18.52	13.43
Barcelona	6.77	9.81	7.18
Metropolitan Areas (more than 500,000 inhabitants)			
Valencia	3.49	3.39	3.48
Sevilla	3.21	1.27	2.95
Bilbao	2.40	1.30	2.25
Málaga	1.74	1.60	1.72
Zaragoza	1.67	1.70	1.67
Sabadell	1.49	1.50	1.49
Palmas de Gran Canaria (Las)	1.43	0.89	1.36
Santa Cruz de Tenerife	1.35	1.21	1.33
Granada	1.32	0.70	1.23
Murcia	1.13	1.12	1.13
Palma de Mallorca	1.01	1.77	1.11
Small and Medium Size cities (more than 50,000 inhabitants)			
121 LLMs	39.23	37.12	38.95
Rural Areas (less than 50,000 inhabitants)			
669 LLMs	21.10	18.09	20.70
All	100.00	100.00	100.00
Frequencies	33.24M	2.87M	4.55M

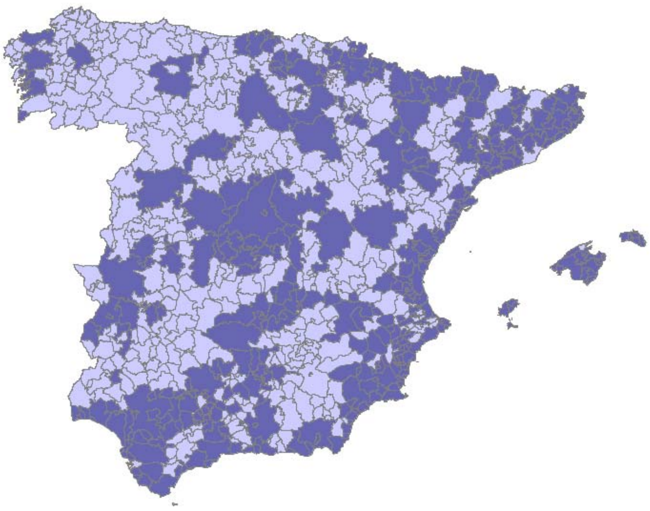
Source: 2011 Spanish Census (INE)

Also national population growth and foreign population growth seem to exhibit a different spatial pattern (Maps 1, 2, 3 and 4). While national population increases only along the mediterranean coast, big cities and Madrid and surrounding areas for both decades, immigrants choose to locate around Madrid Metropolitan area and in rural areas in the 1990s and *everywhere* in the 2000s, but surprisingly the bigger increases are not experienced neither in the Mediterranean coast or the big cities.

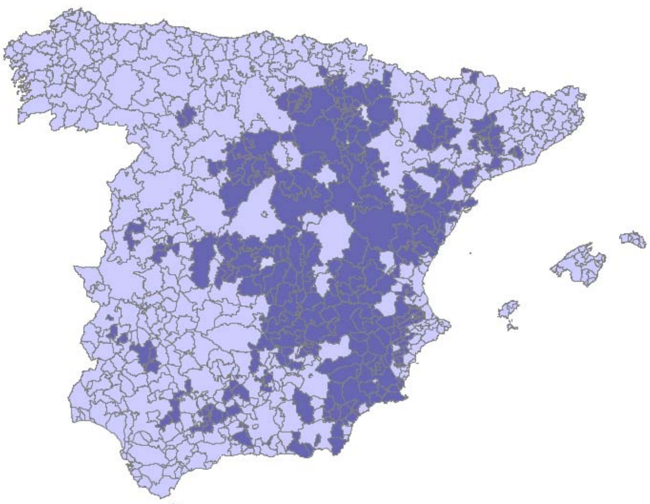
Map 1: Average National Population Growth: 1991-2001



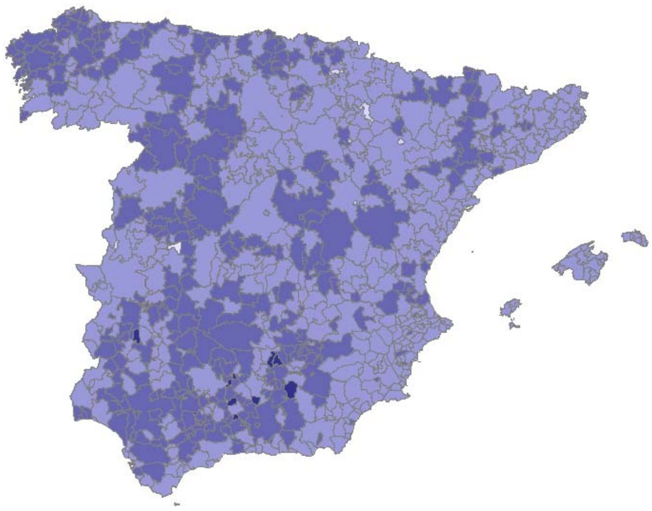
Map 2: Average National Population Growth: 2001-2011



Map 3: Average Foreign-born Population Growth: 1991-2001



Map 4: Average Foreign-born Population Growth: 2001-2011



Legend (for all the maps):

Negative Growth

Positive growth (less than 4%)

Positive growth (more than 4%)

Local-regional factors seem to be exerting different effects over national and foreign population both across space and time. In some cases the centrifugal forces that pushes economic activity and population out from the most congested areas into the less dense ones, prevail over the centripetal forces. By contrast, centripetal forces seem to work in other occasions in the opposite direction, making highly developed and densely populated areas more attractive for economic and population growth. Both forces operate continuously and interact with many other environmental, social, political and geographical factors; making it very difficult to predict which will be the final result.

The objective of this paper is to perform an empirical analysis of the local determinants of population growth in Spain considering i) two sets of population –foreign born population and national population, ii) two different decades, 1991-2001 and 2001-2011; iii) the possible existence of opposite economic forces and environmental, social, political and geographical factors at work; iv) the possible existence of spatial heterogeneity, as population growth and also its determinants (economic growth, employment or education among others) are not homogenously distributed across space.

In order to allow for spatial heterogeneity, we will apply the Geographically Weighted Regressions (GWR) approach to local population growth (nationals and foreigners) for the two decades. The local spatial unit of analysis will be the Local Labour Markets (LLMs), that guarantee homogeneity and reliability across the territory.

This paper is structured as follows. In the next section we review the main the population growth factors acting at local level and immigration location determinants (or foreign-born population growth) and then briefly explain the spatial unit used in this paper. In section three we explain the basics of the GWR, a methodology that captures spatial variations in the regression coefficients. The results obtained regarding the *national* and *foreign-born* population growth determinants applying GWR are then described and compared to the equivalent Ordinary Least Squares estimates. Greater emphasis is placed on those cases where factors seem to have a different or contradictory effect depending on the methodology implemented. This article ends with a section summarizing the main results and the conclusions, and outlining some important economic policy implications derived from the analysis.

2. Population growth determinants at local level: nationals and foreigners

Regional economists are especially interested in linking the economic forces at work with the population dynamics. Formal modelling of local population growth has usually tended to focus on identifying the patterns that are presumed to hold universally. For instance, there are numerous studies aimed at verifying Gilbrat´s or Zipf´s laws (see Gabaix (1999), Gabaix and Ioannides (2004), and Eeckhout (2004), among others); and when the results are not consistent with the pattern predicted, others question the arbitrary administrative limits of cities and explore new definitions based on clustering techniques (Rozenfeld, Rybski, Gabaix and Makse, 2011) or the concept of *natural* cities (Jiang and Jia, 2011).

Evemore, as Glaeser *et al.* (2014) highlighted, population growth laws, although reliable for long-term dynamics, may not be consistent in some moments or for some countries or areas as the balance between *centrifugal* and *centripetal forces* can change and give rise to different specific behaviors.

Most of the empirical analyses of population growth determinants consider a wide set of features such as the initial population and population density, the sectoral employment shares, economic growth or some human capital indicators to evaluate their relevance for the final behavior of population For instance, Beeson et al. (2001) study the long-run population patterns of U.S. counties using the initial levels of population, population density, GDP and human capital, generating a model similar to those of conditional convergence. Also for the U.S., and using the same spatial unit of analysis, Cebula (2002) studies the determinants of local population growth extending the number of factors, finding that public expenditure on education, housing-price inflation, personal income and the quality of life are significant. In another completely different scenario such as China, a link between local population evolution and the massive rural-urban migration, the income gap between both settings and their geographical distance is detected by Zhang and Song (2003).

For the case of Spain, Goerlich and Mas (2008, 2009) describe the changes in the spatial distribution of the Spanish population and find a persistent trend for urban concentration led by regional capital-cities and by regions with certain locational features (coast or valley). However, the study does noes comprise the vast inflow of immigrants that Spain has experienced after 2001. As we can see in maps 1 to 4, only national population seem to respond to that pattern. Foreign-born population growth at local level, or if you prefer, immigrants' location choices, needs further research.

Immigrants' location choice

Migration theory predicts that immigrants are attracted to regions or areas with favorable income prospects. Immigrants usually leave their country of birth in search of a new place to live where they have better work opportunities, higher living standards and/or greater political freedom.

However, migrating to other country is not exempt from costs, both economic, social and psychological. When immigrants decide to move to another country, they have to incur big travel and relocation expenses. Once they arrive into the country their first choice of location tends to be the large cities in which earlier cohorts of co-nationals and other emigrants have settled (Bauer et al, 2001, 2005).

However, assuming that migrating decisions follow an economic rationale, in their adopting country immigrants will settle in those areas where they can maximize the opportunities and minimize the psychological and economic cost of migrating. There are basically two approaches to investigate immigrant's location preferences: one is focusing on the pull factors, i.e., the set of negative or positive social or economic factors in the potential areas of destination which pulls migrants towards them; the other is exploring the push factors, i.e. the set of negative or positive social or economic factors in the area of origin which pushes immigrants away (Lee, 1996). Evidence is far from conclusive. Studies for the USA find contrasting evidence on whether immigrants are sensitive to local differences (Bauer et al 2002, 2005).

One the most cited factors determining location choice is the presence of other foreign-born people in the area. Information about the labour market and many other relevant topics such as housing, schooling, welfare benefit programs etc. is usually accessed through informal channels, i.e. through other family or community or origin members who migrated years before. The presence of foreign population in one country or region, especially if from the same country or area, is a very important determinant of the location choices of immigrants as it eases the relocation process. New immigrants will find cultural and linguistic affinities as well as access to information about the labour market. According to Mincer (1978), while family ties in the destination have an important bearing on people's migration decision, they deter migration in the case of natives

Newly arrived immigrants may not have as much knowledge of regional labour opportunities, regional living standards or welfare policies disparities as natives or immigrants already living in the country (Kaushal, 2005), but it is in their best interest to live in those regions that offer better job opportunities for the parents, educational opportunities for the children and/or higher living standards. Jaeger (2000) finds that immigrants are much more likely to locate in areas with low unemployment and higher wages than other immigrants; Evenmore, they are more likely to locate in areas with growing demands for their skills and areas with increasing real wages.

The initial and subsequent location choice within a country is analysed by Aslund (2005) for the Swedish case. Similarly to the US case, results suggest that immigrants are attracted to regions with large overall immigrant populations and high concentration of co-nationals. However, in this study the hypotheses that information on labour opportunities and or welfare benefits improves over the years cannot be confirmed.

In Europe, the recent study conducted at NUTS I and NUTS II regional level by Rodriguez-Pose and Ketterer (2012) examine the factors that determine the attractiveness of European regions, including economic, socio-demographic and amenity related territorial features. Similarly to results for the USA, economic factors, human capital-related and demographic aspects, as well as the existence of networks and different types of regional amenities, exert an important influence on the relative attractiveness of European regions.

The question is how important are economic and noneconomic local attributes for migration and for nationals' population growth. Given the presence of a number of local characteristics likely to impact on a migrant's location choice and on a national's population growth, the attractiveness of an area seems to be determined by several socioeconomic and geographical attributes.

The aim of this paper is to propose a simple regression model to evaluate the effect of those socioeconomic and geographical determinants on population growth at local level, distinguishing between national and foreign-born population growth.

What spatial unit of analysis? The Local Labour Markets (LLMs)

In Spain there is basically a unique source of information that collects and provides data at highly des-aggregated level (municipality): the Population and Housing Census database. This is the only trustworthy source of information on the socio-economic characteristics of population (age, gender, educational attainment, labor situation, etc.) at municipal level, and is the source commonly used in studies of local population or employment growth.

In our opinion the municipalities are not an appropriate spatial unit for studying population growth –and many other topics- as they fail to comprise economically and socially integrated areas, the more noteworthy cases, though not the only ones, are the metropolitan areas of Madrid and Barcelona, which comprise 151 and 51 municipalities respectively.

Recognizing that labor plays a basic role in people's lives and guides their territorial behavior with regard to the choice of the municipality of residence, and therefore population growth, the so-called *local labor market* (LLM) or travel-to-work area describes a space where the population develops most of its economic and social relationships. An LLM is a place where the common interest of the local population can be identified as a whole, and is therefore the appropriate level for implementing policies at local or regional level (Parr, 2005), something theoretically simple but which for political reasons usually faces strong resistance from the municipal entities involved.

As the construction of the LLMs guarantees that more than 75% of the residents living in the area also work in the area and vice versa, the LLMs have the advantage of ensuring that, regardless of the (unknown) place of work of the workers living in the area, most of the observed *employed population growth* is in fact *employment growth* within the area (Sforzi, 2012). Boix and Galleto (2006) identify 803 LLMS for Spain; LLMs construction guarantees they are self-contained, well-defined and compact functional areas with a high internal homogeneity in terms of labor and income (Rubiera and Viñuela, 2012).

3 . Empirical analysis

The determinants

In order to construct a database using LLMs as the spatial unit, it is crucial to have access to data at a high level of disaggregation.

Our dependent variable (*Y*) is the *Population growth* (nationals and foreign-born) in different periods of time, which we try to explain through a set of possible determinants or explanatory variables (*X*), for *n* spatial units (803 LLMs):

$$Y_i = \beta_0 + \beta_1 X_i + \varepsilon_i$$
$$\varepsilon_i \sim N(0, \sigma^2) \; ; \quad i = 1, \dots, n \qquad [1]$$

Among the growth determinants[1], we include the *Population size* at the beginning of the period considered (*Pop*) as the first explanatory factor. This variable reflects the rank of the place in the urban hierarchy and the presence of agglomeration effects. (Table 3 provides descriptive statistics on the national and foreign-born growth determinants)

Data on population and many other relevant socio-economic variables (employment, employment by sector, educational level, etc.) at such a level of disaggregation is only available in Spain -and many other countries- thanks to the *Population and Housing Census*. Carried out every ten years by the National Statistics Institute (INE), two

[1] For a revision, see Glaeser *et al.* (2014)

decades can be analyzed in the Spanish case: the 1990s (specifically from 1991 to 2001) and the 2000s (from 2001 to 2011).

Table 3. Descriptive statistics of the variables considered

Variable	Min.	Max.	Mean	Std. Dev.
Population Growth 91-01	-0.3130	1.3788	0.0221	0.1607
Population Growth 01-11	-0.3095	1.2703	0.0999	0.1946
National Population Growth 91-01	-0.3895	0.6179	-0.0156	0.1181
National Population Growth 01-11	-0.4054	0.5971	-0.0018	0.1154
Foreign Population Growth 91-01	-1.6094	6.0753	2.0037	1.2185
Foreign Population Growth 01-11	-0.0532	4.9126	1.7141	0.5692
Pop 91 (Population 91)	1,909	4,864,827	48,252.98	225,686.40
Pop 01 (Population 01)	1,951	5,316,487	50,690.66	238,724.20
Nal1991 (National population 1991)	1,907	4,773,164	47512.74	221,895.8
Nal2001 (National population 2001)	1,930	4,929,667	48462.92	224,119.1
Ext1991 (Foreign population 1991)	1	59,994	425.0374	2634.464
Ext2001 (Foreign population 2001)	1	359,050	1916.565	14112.25
DistMA (Distance to nearest MA)	0	614.35	155.90	1.1954
Ed 91 (Educational level 91)	0.0095	0.1510	0.0543	0.02269
Ed 01 (Educational level 01)	0.0317	0.2551	0.1043	0.0391
Emp 91 (Employment rate 91)	0.2270	0.9481	0.4857	0.08531
Emp 01 (Employment rate 01)	0.3552	0.9140	0.5841	0.0795
LQagr 91 (Location quotient agr. 91)	7.227	773.769	262.7603	163.1501
LQagr 01 (Location quotient agr. 01)	7.51	1124.936	284.9863	206.8542
LQind 91 (Location quotient ind. 91)	12.406	284.808	83.12267	53.25639
LQind 01 (Location quotient ind. 01)	8.722	345.393	103.695	65.23617
S 91 (Specialization index 91)	1.3009	7.9924	4.2181	1.0598
S 01 (Specialization index 01)	0.9934	8.405	4.0321	1.0951
Lo (Longitude coordinate)	-18.008	4.263	-3.4872	3.63216
La (Latitude coordinate)	27.757	43.686	39.6759	2.8004
DistCoast (Distance to coast)	0	327.1826	87.4699	0.8711
Rain (Avg. annual rainfall 1987 to 2007)	1	16.412	5.2379	2.7211
$Tmin_{jan}$ (Min. temperature January)	-6.774	12.5	1.2037	3.2716
$Tmax_{jul}$ (Max. temperature July)	20.496	36.7	30.6705	3.5964
Cap (Capital city dummy)	0	1	0.0647	0.2462
Cases	**803**			

In a Regional Economics framework, following *size* comes *location* as a crucial explanatory factor. In this study, location is going to be measured as *distance to size* (*DistMA*)*:* the linear distance to the nearest metropolitan area (an LLM with at least 500,000 inhabitants). Defined this way, this variable contains information about the interrelationships between cities, the power to attract or repel population (national or foreign-born) that a certain area has, and its accessibility to goods, services, amenities and labor markets (Coffey and Polèse, 1988; Polèse and Shearmur, 2004). Thus, for instance the remotest Spanish LLM is located North-West at 614 km from the nearest metropolitan area (Table 3). The LLM's geo-referenced information can be obtained from the Spanish National Centre for Geographical Research (IGN).

The effect of *size* and *distance* should be analyzed hand–in-hand as together they can exert ambiguous effects on population growth: rural and peripheral areas where population does not surpass a certain critical mass might be experiencing a de-population process while rural (small) but central areas might be ideal places of residence for commuters (nationals or foreign-born). On the other hand, urban areas might be experiencing different population growth dynamics depending first on their population size (small-medium city *vs* large metropolitan area) and second on their relative location within the Spanish territory. Moreover, these variables could be exerting a positive effect on population growth in some areas but the opposite in others, on nationals one way and on foreign-born population in another. For instance, one demographic variable that seems to exert a big influence on immigrant's location choice is the presence of immigrants in the area. This factor can act as an attraction for newly arrived immigrants, but as a repellent for nationals or highly skilled immigrants.

To complete the analysis, as controls we include in the local growth model other factors that can be classified as socio-economic, geo-structural and climatic variables. It is worth noting that the only kind of factors that can be affected by means of policy instruments are those belonging to the first class.

As socio-economic variables we consider the *educational level* (*Ed*) measured as the percentage of the potential workforce (population aged 16 to 64) that holds a university degree, with this being an indicator of the initial human capital endowment. This factor is related to several dimensions (population mobility, LLM's economic structure or productivity, income, etc.) that can affect population change and provide interesting hints about its evolution. Over recent decades, tertiary education attainment has increased overall in terms of the mean (from 5% to 10%), the minimum or the maximum values. Other variables considered are the *employment rate (Emp*), representing the labor market conditions at the start of each interval (average of 49% in 1991 and 58% in 2001). Although the debate on whether population follows jobs or jobs follow population is far from settled, in this work we follow the casualty direction found in, among others, Partridge and Rickman (2003), that is, people are more likely to be following jobs. Two *sectoral location quotients* are also included for agricultural

(*LQagr*) and industrial activities (*LQind*), providing information about the concentration of both sectors in a certain area with respect to the national average. Following Shearmur and Polèse (2005), a regional specialization index (*S*) is also considered to reveal the degree of economic specialization/diversification of the place.

The *Longitude* (*Lo*) and *Latitude* (*La*) coordinates of the centroid of the LLM and the *linear distance to the nearest coast* (*DistCoast*) are the geo-structural factors in the model. The former captures the influence of the absolute location of the place in the territory[2] while the latter represents the proximity of the place to the coast, a historically and naturally privileged location for population growth. We also include variables related to the climate: *the average annual rainfall* (*Rain*), *the minimum temperature in January* ($Tmin_{jan}$) and the *maximum temperature in July* ($Tmax_{jul}$). These factors are to some extent indicators of quality of life and natural amenities, and may play a role in retaining and attracting population. This climatological information comes from the historical series (1987-2007) published by the Meteorological State Agency (AEMET).

As capital cities tend to have their own demographic dynamics, a dummy variable will control for those LLMs that correspond to the politico-administrative *capital of a region* (*Cap*). This variable underlines the importance of having been appointed as the administrative center, thereby concentrating a large part of the public sector jobs and offering a larger variety of public services.

The methodology

As stated earlier, it is important to acknowledge that the estimated effects of a variable can vary greatly across countries, and even within the same country depending on the temporal or the spatial framework chosen (Shearmur and Polèse, 2005; Shearmur et al., 2007; Glaeser et al., 2014). Given the existence of spatial heterogeneity, the question is whether a single estimate can properly explain national population growth or foreign-born population growth (immigrants' location choice).

Spatial non-stationarity takes place when the responses to particular variables change across space, and these differences might be caused by the interrelationships between neighboring regions. Adopting a global regression approach such as Ordinary Least Squares (Eq. 1) might lead to deceptive estimates if those are extrapolated to the local

[2] As the Greenwich meridian, reference to measure the Longitude, divides the country in two halves, we proceed to a change in the scale. Longitude zero corresponds to the easternmost LLM (Mahon) and Latitude zero corresponds to the northernmost LLM (Xove).

environment. The conclusions regarding, for instance, the population or economic growth determinants resulting from global estimations can mask significant local variation as a standard overall estimate may point to a certain effect of one factor, while in fact such factor could be stimulating growth in some areas but negatively affecting it in others, showing an average effect which is not representative at the local level due to its high regional variability. This compensation effect is especially dangerous when the average impact is close to zero, as it might be deemed as non-significant and disregarded as an element of the analysis or as a policy instrument.

The simplest approach proposed in the literature to address spatial non-stationarity is the fixed-effects model, where dummy variables are introduced to capture site-specific characteristics (Brunsdon *et al.*, 1998; Greene, 2000). To correct for spatial dependence, Anselin (1988) suggested a spatial error model (SEM) and a spatial lag model (SLM). Both models take into account the problems mentioned above, but parametric heterogeneity is not accomplished, so an important source of regional information is lost.

The Geographically Weighted Regression (GWR hereinafter) is a non-parametric model that represents an alternative to deal with both issues (Brunsdon et al., 1996 and 1998). The GWR approach can be easily implemented, hypothesis testing is akin to that of standard methods and results can reveal interesting spatial regularities undetected by more traditional methods (McMillen and Redfearn, 2010). This methodology captures spatial variations in the regression coefficients by introducing a weighting matrix in the estimation process and estimating a locally varying sample for each location, generating a separate group of regression parameters which reflects the sample heterogeneity by estimating different responses to an explanatory variable across space.

The GWR model, where a regression for each observation is estimated, is specified as:

$$Y_t = \beta_{t0} + \beta_{t1}x_{t1} + \beta_{t2}x_{t2} + \cdots + \beta_{tk}x_{tk} + \varepsilon_t$$

[2]

$$\varepsilon_t \sim N(0, \sigma^2);\ t = 1, 2, \ldots, n$$

giving as a result a separate set of parameters for each of the *n* observations, calculated through the following equation:

$$\hat{\beta}_i = (X^t W_i X)^{-1} X^t W_i Y_i \quad i = 1, 2, \ldots, n \qquad [3]$$

Our dependent variable (Y_i) is the *i*th LLM's *Population growth* (nationals and foreign-born) in different periods of time, and as explanatory variables (X_i) the same ones considered in the previous section.

For each observation (for each LLM) a separate regression is estimated in which the sample is composed of spatial units within a certain distance or *bandwidth*. There are different criteria to specify the distance, such as the minimization of the Akaike Information Criterion (AIC) (information loss indicator) or the minimization of the sum of squared errors (or the cross-validation score, CVS).

The *weights* on the GWR depend on the linear distance between observations and represent the adjacency effects for neighbouring locations within the specified bandwidth (Cleveland and Devlin (1988); McMillen (1996), and Brunsdon *et al.* (1996 and 1998). W_i is a diagonal weighting matrix that selects the observations that intervene in the estimation of the local coefficients, $\hat{\beta}_i$ in point i :

$$W_i = \begin{bmatrix} \alpha_{i1} & 0 & 0 & \ldots & 0 \\ 0 & \alpha_{i2} & 0 & \ldots & 0 \\ 0 & 0 & \alpha_{i3} & \ldots & 0 \\ \ldots & \ldots & \ldots & \ldots & \ldots \\ 0 & 0 & 0 & \ldots & \alpha_{ij} \end{bmatrix} \qquad [4]$$

Following the assumption that more proximate locations are more alike, the weights should decay with distance. Many weighting notations could be used (dichotomous, bi-square or tri-square decay function, etc.). In this work we have chosen a Gaussian kernel weighting function specified as:

$$\alpha_{ij} = e^{-(1/2)\left(d_{ij}/h\right)^2} \qquad [5]$$

where d_{ij} is the distance between observations *i* and *j*, and *h* is the general distance bandwidth adopted. Thus, the weight quickly declines with distance from the geographic observation concerned.

The GWR approach has several advantages over standard methods, but also has flaws. One advantage is that since each area has its own constant term, it accounts for local-fixed effects (Partridge *et al.*, 2008). This approach can reduce spatial error correlation when there is heterogeneity in the GWR coefficients (Fotheringham *et al.*, 2002). One shortcoming is multicollinearity, which can be problematic in individual local regressions, but as the GWR approach produces a considerable amount of regressions, considering a large range of estimates allows us to "average" them, thereby better determining their central tendencies and distribution (Ali *et al.*, 2007; Partridge *et al.,* 2008).

GWR also has some inherent limitations. The fact that each local model does not take into account local spatial dependence may bias local estimates (Shearmur *et al.*, 2007). Some other drawbacks are linked to the local regressions when using a smaller sample size, as the resulting coefficients may be less efficiently estimated than those from global approaches. Apart from that, GWR is computationally intensive and the output can be overwhelming (Ali *et al.,* 2007; Partridge *et al.*, 2008). Other flaws are the robustness of the results, which depends on bandwidth selection, and the existence of possible sample overlaps (Ali *et al.,* 2007). Even more, significant local regression coefficients do not necessarily indicate correlation with certain spatial unit, but that correlation can be observed *across the bandwidth* specified in the GWR process (Shearmur *et al.*, 2007). Probably the strongest criticism for this methodology is the one made by Wheeler and Tiefelsdorf (2005), who found evidence supporting the existence of “false positives” regarding the ability of GWR to distinguish between spatially stationary processes and varying ones, therefore pointing out the unreliability of the estimations. In the light of these statements, Páez *et al.* (2011) conducted several simulations and concluded that, even when deeper examination should be done and caution is recommended, the severity of the mentioned problems decrease with the size of the sample[3].

Results: OLS versus GWR

Using data for the 803 Spanish LLMs, estimations from the OLS approach (Eq. 1) will first be presented and then compared with the ones obtained with the GWR model (Eq. 2) as an attempt to understand the local processes behind population growth (national and foreign-born) and its spatial patterns.

[3] In this paper, the number of observations in our database (803 LLMs) can be considered large enough to produce reasonable results from an econometric perspective.

The OLS estimation provides fourteen global parameters (plus one intercept) for each group that reflect the effect of each explanatory variable on the sample. The result of this estimation may veil the effect of each factor on each region, something particularly troublesome if response heterogeneity exists and the associated variability is large (Ali *et al.* (2007) and Partridge *et al.* (2008)).

The GWR procedure will provide 15 x 803 parameters, one for each factor (plus de intercept) and LLM. We will especially focus on discussing those factors where the GWR estimates add useful and significant information on local population growth determinants. After that, we will represent those factors showing a spatially differentiated effect to deduce spatial or regional policy implications for Spain[4].

Estimations for both models are shown for the 1991-2001 decade (Table 4 for Nationals and Table 5 for foreign-born population) and 2001-2011 decade (Table 6 and 7 for nationals and foreign-born respectively). The second column presents the results for the OLS model, while the following five ones show the quartile intervals for the GWR estimations and for each population growth factor considered[5].

Checking the OLS estimations differences between the two decades analyzed, we find visible temporal heterogeneity in some determinants. The *size* and *distance to size* explanatory variables' behavior is noteworthy. Being a big or medium-small city did not act as a population growth attraction factor in the 1990s, whereas in the first decade of the 21st century size is significant. However, for foreign-born population, size is always a attraction factor in both decades. The sign of the *size* parameter is positive and significant for immigrants, while non significant or close to zero for nationals, which might be giving us the clues of future population growth trends once immigrants are leaving the country. The result for the variable *distance to the nearest metropolitan area* (DistMA) is significant for both samples and periods of time.

The variable employment rate at the beginning of the time interval (Emp) is positive and significant for both groups and periods of time, but the rest of the independent variables remain significant for the whole time interval only for the national population

[4] Only some factors' estimates are represented in this paper. Please contact the authors to request for additional information.

[5] In order to be able to compare results from different decades and different methodologies, the explanatory variables for population growth are included both in OLS and GWR. For the same reason, the average of the optimal bandwidth for each period is used for both decades under the GWR procedure.

growth. There are significant shifts for the immigrant's group behavior in a relatively short time-span, underscoring the relevance of the existence of temporal heterogeneity on the explanatory variables when analyzing a population phenomenon or any other process.

Table 4. Estimation Results for 1991-2001 (National Population)

Variable	Global (OLS)		Min.	1st Qu.	Median	3rd Qu.	Max.	F3 test(a)
Intercept	-0.0383		-1.2100	-0.3955	-0.1148	0.2654	0.7314	+++
NatPop91(National Population 1991)	-0.0003		-0.0538	-0.0178	-0.0016	0.0169	0.0455	+++
ExtPop91(Foreign Population 1991)	0.0044		-0.0146	0.0000	0.0060	0.0114	0.0256	
DistMA (Distance to nearest MA)	-0.0134	***	-0.0796	-0.0377	-0.0216	-0.0054	0.0710	+++
Ed91 (Educational level 1991)	-0.0076	***	-2.2230	-1.0600	-0.6993	-0.2926	1.2120	
Emp91 (Employment rate 1991)	0.0034	***	-0.1173	0.1495	0.3138	0.4435	1.1290	+++
$LQ_{ag}91$ (Location quotient agr. 1991)	-0.0004	***	-0.0007	-0.0005	-0.0004	-0.0001	0.0004	+++
$LQ_{ind}91$ (Location quotient ind. 1991)	-0.0006	***	-0.0013	-0.0008	-0.0005	0.0000	0.0009	+++
S91 (Specialization index 1991)	0.0084		-0.0263	-0.0126	-0.0011	0.0124	0.0665	+++
DistCoast (Distance to coast)	-0.0107	*	-0.1774	-0.0366	-0.0124	-0.0032	0.0805	++
Rain (Avg. annual rainfall 1987 to 2007)	-0.0079	***	-0.0403	-0.0223	-0.0128	-0.0025	0.0094	+++
$Tmin_{jan}$ (Min. temperature January)	0.0113	***	-0.0056	0.0053	0.0098	0.0135	0.0291	+
$Tmax_{jul}$ (Max. temperature July)	0.0024		-0.0166	-0.0024	0.0042	0.0090	0.0401	+++
Cap (Capital city dummy)	0.0178		-0.1376	-0.0055	0.0212	0.0563	0.1666	+++
Adjusted R^2 OLS	0.5082				F-statistic	62.72	***	
Adjusted R^2 GWR	0.7289							
F1 test(b)	0.7164	+++						
F2 test(c)	1.9466	+++						

- */**/*** and +/++/+++ represent global significance or significant variation respectively at 10%/5%/1% level.
- (a), (b) and (c): statistical tests proposed by Leung, Mei and Zhang (2000). F1 and F2 are intended to compare the goodness of fit between OLS and GWR models, while F3 verifies the significance of the variation in the set of coefficients obtained through GWR for each factor.

Table 5. Estimation Results for 1991-2001 (Foreign Population)

Variable	Global (OLS)		Min.	1st Qu.	Median	3rd Qu.	Max.	F3 test(a)
Intercept	2.0094	**	-19.2400	-4.5030	-2.6570	0.1293	6.8550	+++

NatPop91(National Population 1991)	0.4262	***	0.1448	0.5524	0.6603	0.8180	1.2000	+++
ExtPop91(Foreign Population 1991)	-0.5277	***	-1.0460	-0.7644	-0.6534	-0.5704	-0.3670	+++
DistMA (Distance to nearest MA)	-0.0197		-0.7958	-0.2198	-0.0406	0.2386	1.7640	+++
Ed91 (Educational level 1991)	-3.9761	*	-19.8500	-5.1510	-0.8908	3.6490	17.9200	
Emp91 (Employment rate 1991)	4.0115	***	-3.3790	2.0460	3.7970	5.5520	10.3400	+++
$LQ_{ag}91$ (Location quotient agr. 1991)	-0.0004	***	-0.0073	-0.0021	-0.0004	0.0005	0.0026	+++
$LQ_{ind}91$ (Location quotient ind. 1991)	-0.0009		-0.0115	-0.0035	-0.0015	0.0006	0.0128	+
S91 (Specialization index 1991)	-0.0705		-0.4272	-0.0782	0.0076	0.0995	0.8443	+++
DistCoast (Distance to coast)	0.0302	*	-4.2080	-0.3992	-0.0237	0.1709	0.8353	+++
Rain (Avg. annual rainfall 1987 to 2007)	-0.2233	***	-0.4754	-0.1404	-0.0607	-0.0137	0.2260	+++
$Tmin_{jan}$ (Min. temperature January)	0.0149		-0.1196	-0.0274	0.0308	0.0690	0.3195	+
$Tmax_{jul}$ (Max. temperature July)	-0.0615	***	-0.3121	-0.0936	-0.0082	0.0556	0.4263	+++
Cap (Capital city dummy)	-0.058		-1.0920	-0.3621	-0.1242	0.1688	1.5090	
Adjusted R^2 OLS	0.587				F-statistic	88.69	***	
Adjusted R^2 GWR	0.8568							
F1 test(b)	0.5241	+++						
F2 test(c)	1.9769	+++						

- */**/*** and +/++/+++ represent global significance or significant variation respectively at 10%/5%/1% level.

- (a), (b) and (c): statistical tests proposed by Leung, Mei and Zhang (2000). F1 and F2 are intended to compare the goodness of fit between OLS and GWR models, while F3 verifies the significance of the variation in the set of coefficients obtained through GWR for each factor.

Table 6. Estimation Results for 2001-2011 (National Population)

Variable	Global (OLS)		Min.	1st Qu.	Median	3rd Qu.	Max.	F3 test(a)
Intercept	-0.4690	***	-1.0550	-0.6242	-0.3554	0.0949	1.1580	**+++**
NatPop91(National Population 1991)	0.0104	*	-0.0565	-0.0207	0.0039	0.0146	0.0466	**+++**
ExtPop91(Foreign Population 1991)	0.0118	***	-0.0089	0.0114	0.0154	0.0240	0.0456	
DistMA (Distance to nearest MA)	-0.0217	***	-0.0954	-0.0493	-0.0357	-0.0184	0.0520	

Ed91 (Educational level 1991)	-0.3088 *	-1.4820	-0.6062	-0.3824	-0.1404	0.5770	
Emp91 (Employment rate 1991)	0.3755 ***	-0.8422	0.1788	0.2846	0.4341	0.8065	+++
$LQ_{ag}91$ (Location quotient agr. 1991)	-0.0002 ***	-0.0007	-0.0003	-0.0002	-0.0002	0.0001	+++
$LQ_{ind}91$ (Location quotient ind. 1991)	-0.0003 ***	-0.0012	-0.0005	-0.0003	-0.0002	0.0006	+++
S91 (Specialization index 1991)	0.0040	-0.0264	-0.0041	0.0031	0.0129	0.0499	
DistCoast (Distance to coast)	-0.0095	-0.1774	-0.0366	-0.0124	-0.0032	0.0805	++
Rain (Avg. annual rainfall 1987 to 2007)	-0.0010	-0.0535	-0.0204	-0.0104	-0.0010	0.0103	+++
$Tmin_{jan}$ (Min. temperature January)	0.0091 ***	-0.0091	0.0064	0.0107	0.0143	0.0284	+++
$Tmax_{jul}$ (Max. temperature July)	0.0075 ***	-0.0279	0.0015	0.0064	0.0106	0.0261	+++
Cap (Capital city dummy)	-0.0286	-0.1029	-0.0490	-0.0210	0.0151	0.0899	
Adjusted R^2 OLS	0.5101			F-statistic	65.23	***	
Adjusted R^2 GWR	0.7244						
F1 test(b)	0.7322 +++						
F2 test(c)	1.9541 +++						

- */**/*** and +/++/+++ represent global significance or significant variation respectively at 10%/5%/1% level.

- (a), (b) and (c): statistical tests proposed by Leung, Mei and Zhang (2000). F1 and F2 are intended to compare the goodness of fit between OLS and GWR models, while F3 verifies the significance of the variation in the set of coefficients obtained through GWR for each factor.

Table 7. Estimation Results for 2001-2011 (Foreign Population)

Variable	Global (OLS)	Min.	1st Qu.	Median	3rd Qu.	Max.	F3 test(a)
Intercept	0.1078	-3.9630	-0.2466	1.2000	2.6070	5.1290	+++
NatPop91(National Population 1991)	0.4295 ***	0.3218	0.4374	0.4715	0.5420	0.7803	+++
ExtPop91(Foreign Population 1991)	-0.4319 ***	-0.7450	-0.5249	-0.4874	-0.4440	-0.3627	+++
DistMA (Distance to nearest MA)	-0.0408 **	-0.5941	-0.1926	-0.0624	0.0303	0.4630	+++

Ed91 (Educational level 1991)	-0.7585		-6.1220	-2.2460	-0.8875	0.0076	3.6260	+
Emp91 (Employment rate 1991)	1.4441	***	-1.0210	0.5868	0.9652	1.6330	3.5380	+++
$LQ_{ag}91$ (Location quotient agr. 1991)	-0.0004	***	-0.0019	-0.0008	-0.0005	-0.0001	0.0004	+++
$LQ_{ind}91$ (Location quotient ind. 1991)	-0.0020	***	-0.0045	-0.0029	-0.0021	-0.0016	0.0014	
S91 (Specialization index 1991)	-0.0167		-0.1277	-0.0394	-0.0108	0.0257	0.2672	+
DistCoast (Distance to coast)	-0.0859	***	-0.3302	-0.1018	-0.0049	0.0586	0.2897	
Rain (Avg. annual rainfall 1987 to 2007)	-0.0284	***	-0.1634	-0.0668	-0.0451	-0.0270	0.0122	
$Tmin_{jan}$ (Min. temperature January)	0.0047		-0.0615	-0.0019	0.0227	0.0382	0.0761	+++
$Tmax_{jul}$ (Max. temperature July)	-0.0068		-0.1785	-0.0787	-0.0313	-0.0028	0.0650	+++
Cap (Capital city dummy)	-0.0297		-0.3132	-0.0796	-0.0135	0.0457	0.2424	
Adjusted R^2 OLS	0.6137				F-statistic	99.01	***	
Adjusted R^2 GWR	0.7867							
F1 test(b)	0.7139	+++						
F2 test(c)	2.0502	+++						

- */**/*** and +/++/+++ represent global significance or significant variation respectively at 10%/5%/1% level.

- (a), (b) and (c): statistical tests proposed by Leung, Mei and Zhang (2000). F1 and F2 are intended to compare the goodness of fit between OLS and GWR models, while F3 verifies the significance of the variation in the set of coefficients obtained through GWR for each factor.

Estimated using OLS, the model and determinants proposed can explain an important part of Spanish population growth (nationals and foreign-born) at local level in both decades. However, comparing the global results (OLS) with the ones obtained under the GWR approach can enrich the analysis, complementing the time heterogeneity issue discussed above and providing information on the spatially differentiated effects of some explanatory factors.

With this aim, we have to compare de goodness of fit between OLS and GWR estimations and then wonder about the significance of the variation of the 15 x 803 parameters, one for each LLM and factor (plus de intercept), obtained using GWR. Applying both the F1 and F2 statistic tests proposed by Leung *et al.* (2000), the GWR results outperform the OLS approach in both decades. The F1 statistic is defined as the ratio between the squared sum of residuals (SSR) of OLS and GWR, so a value significantly smaller than one (0.71 for the first decade and 0.73 for the second with the national's group; 0.52 and 0.71 with foreign-born group) supports a better fit for the GWR estimation. The F2 test is based on the SSR improvement of GWR over OLS, i.e., the difference between the residuals sums of squares. A large value of this test (1.95 and 1.94 in the first and second decade for nationals: 1.97 and 2.05 in the foreign-born group proves that GWR outperforms the OLS approach. In addition to the F1 and F2 statistical tests, Leung *et al.* (2000) suggested checking the differences between both approaches with the F3 statistic, which tests the significance of the variation in the 15 x 803 parameters estimated using GWR. This F3 statistic, represented in the last column of Tables 4 to 7, needs to be interpreted considering also the significance - or lack of it - of the OLS estimates.

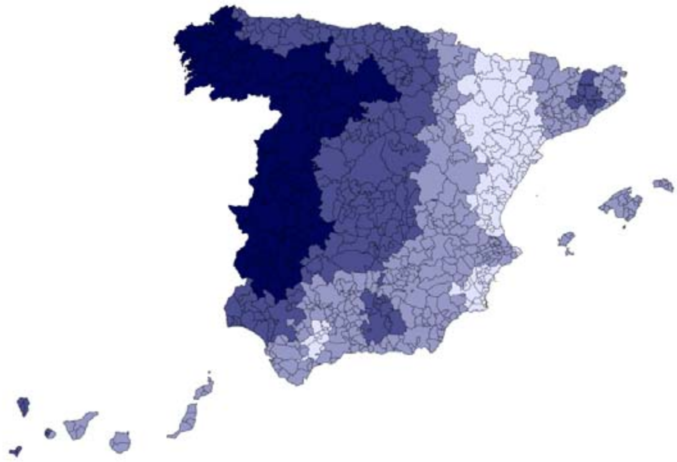
Regarding the comparison between global and variability significance, the first scenario would be having OLS estimates that are significant while the variation in the parameters under GWR is not: in this case the OLS parameter for the whole territory could also be considered representative at local level. In other words, the factor under study does not have a spatially differentiated effect on population growth. The second possibility is having significance of the variation of the coefficients obtained under GWR but no significance of the global OLS coefficient. This situation might be revealing a compensation effect in the OLS estimations, i.e. the existing regional variability leads to an average general effect proximate to zero. The last possibility is when both the OLS estimator and the F3 statistic are significant. In this scenario, although significant, the OLS estimates have failed to capture the existence of spatial non-stationarity revealed by the F3 test. In the second and third scenarios, the use of the GWR approach becomes necessary to understand the spatially differentiated processes at work, and therefore to propose customized policy implications at local level.

As representatives of the first scenario there are the *foreign population (ExtPop*) for the *national* population growth or the *average rainfall* (*Rain*) in the period 1991-2001; and the *employment rate* (*Emp*) or the *distance to the nearest metropolitan area* (distMA) for 2001-2011 (Table 6). These factors affect local population growth in a similar way regardless of the characteristics of the LLM under consideration, which means that these factors have an even impact on the Spanish territory, at least for the decade observed, pointing again to the differences depending on the timeframe chosen.

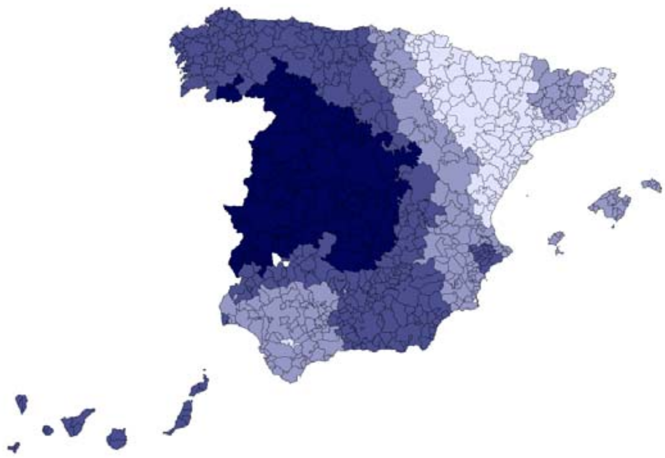
The second scenario is the case of immigrant's location choice depending on the variable *distance to the nearest metropolitan area* or the variable *Specialization index* in the first decade (Table 5). In the OLS estimation, the distance to the nearest metropolitan area show the expected negative sign but estimates are not significant. According to this global result, immigrants´ location choices in the 1990s did not depend on the relative location of the area, but in the immigrants of the 2000s become wiser and gave importance to location. However, when analyzing these same factors allowing for spatial heterogeneity, the GWR estimation shows that in certain LLMs distance to the main urban areas (*DistMA*) in the first decade could be affecting immigrants location choices in the opposite direction to the global effect. In other words, for some LLMs, being apart from the large metropolitan areas could be having a positive influence on foreign-born population growth.

GWR estimates for the variable *Size*, measured by the *initial national population*, seems to be a relevant positive feature for the national's population growth. For both decades, the F3 test for this variable reveals the existence of spatial-stationarity, Representing the GWR estimates on a map (Maps 5 and 6), we can observe that the agglomeration economies-diseconomies processes need to be analysed separately for this group as they are simultaneously taking place in the same geographical area (Spain) for the two decades considered.

Map 5: GWR estimates for *national* population growth at local level: 1991-2001



Map 5: GWR estimates for *national* population growth at local level: 2001-2011

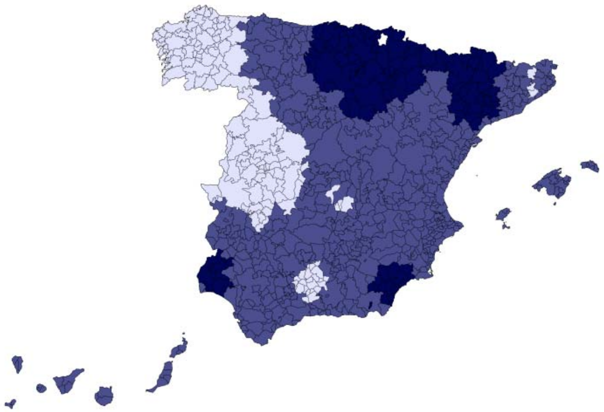


Legend:

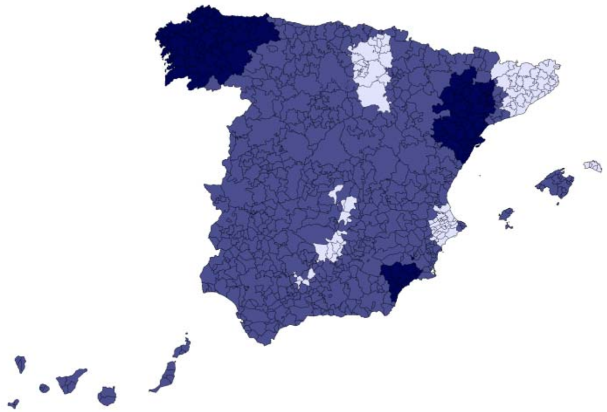
Agglomeration diseconomies processes stand out in the first and second decade in those urban areas located in the North-East (Valle del Ebro Axis), as shown in Maps 4 and 5. This effect is generally associated with negative externalities such as congestion, environmental degradation or higher housing prices in large cities. However, the agglomeration economies still operating for *national* population in the north and north west of the country are not shown when estimating by OLS; the positive effect of size (agglomeration economies) for some LLMS is are offset by the agglomeration diseconomies (negative parameters), and as a result, the parameter for the initial size is not significant.

The other variable we would like to highlight is the *employment rate* at the beginning of the period (Emp). Labour opportunities, proxied by the local employment rate, display a positive and significant effect in the OLS estimates for both periods and groups. However, although significant they fail to capture the existence of spatial non-stationarity, revealed by the significance of the F3 test. The results obtained for this variable both in the OLS and the GWR estimations are however complex to interpret aside from the casualty relationship between employment and population growth. When representing the GWR estimates for this variable regarding national population growth (Maps 6 and 7) we can see how the less dynamic areas in terms of economic activity, and therefore employment, at the beginning of the 1990s were practically stagnant in terms of population growth. Those areas correspond to the western half of the country However having a low employment rate constituted an attraction to national population growth for this area a decade later. The 1990s were characterized by a short but intense economic crisis between 1991 and 1994, with almost one million job losses in three years (Del Valle, 1995; Serrano, 2011). The decade started with a 49.12% employment rate, and by the beginning of the 21st century it reached 59.40%. In between, some Spanish regions and sectors suffered more intensively the effects of the crisis than others; indeed, the recovery was not evenly distributed across space, neither in terms of job creation nor population growth.

Map 6: GWR estimates for initial *employment rate* at local level: 2001-2011 (Nationals)

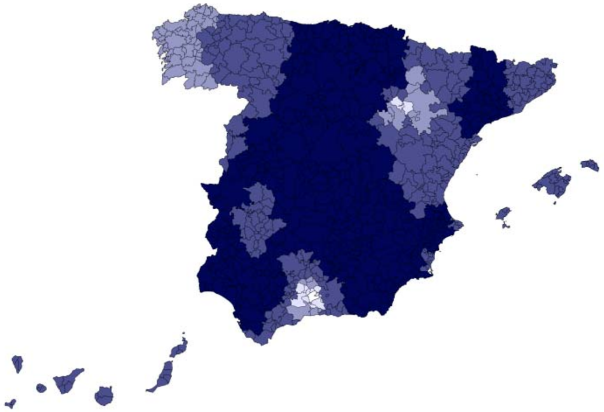


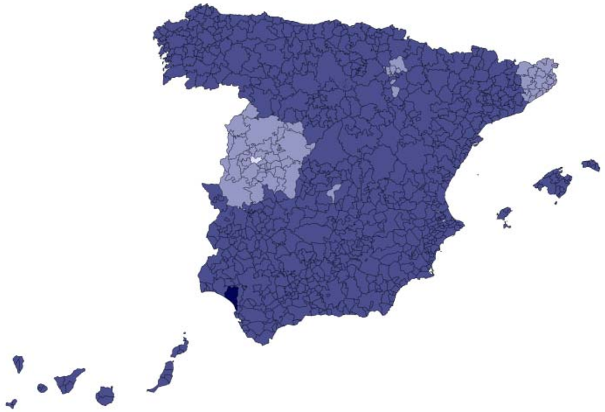
Map 7: GWR estimates for initial *employment rate* at local level: 2001-2011 (Nationals)



Map 8: GWR estimates for initial *employment rate* at local level: 2001-2011 (Foreign-born)

Map 9: GWR estimates for initial *employment rate* at local level: 2001-2011 (Foreign-born)





However, for the foreign-born population, initial employment rates only has a negative relationship over immigration locations in few local labor areas located in and around Zaragoza and Malaga metropolitan areas in the first decade, and around the rural areas of Extremadura and Salamanca in the second one In the rest of the Spanish territory, regardless of their initial level of economic activity or employment, Spanish LLMs underwent a positive foreign-born population growth but with significant local variation.

A spatially differentiated behavior can be observed for other globally significant variables, such as the winter minimum temperature ($Tmin_{jan}$) and summer maximum one (Tmaxij) or the location quotiens for agriculture and industry; thus, a high specialization on the agricultural and industrial sector can work as a deterrent to both national and foreign-born population growth in one area while in others is acting as a population attraction pole.

4. Conclusions

The analysis of population allocation and growth usually searches for long-run universal patterns, but factors explaining the population dynamics and growth determinants can change their effect and significance depending on the group considered (nationals versus foreign-born individuals) an also along time and across space.

The objective of this paper is to analyze the local determinants of population growth in Spain assuming that the same socio-economic, political and geographical factors can affect *national* population growth and *foreign-born* population growth, i.e., immigrants´ location choices.

To capture both spatial and temporal heterogeneity, we adopted the Geographically Weighted Regressions approach - which allow us to observe the possible spatial non stationarity - for two decades (the 1990s and the 2000s) with very different population dynamics in Spain. In both decades national population kept almost steady while more than 4 million immigrants got into the country in the second decade.

Following the literature on population growth and immigrant's location choices, determinants include socio economic variables - such as the initial population level, the previous existence of foreign-born population or the specialisation on the agricultural or industrial sectors-, labour related variables –such as the employment rate at the beginning of the period or the existence of a pool of highly qualified workforce-, and many geographical and amenity related factors -such as the distance to the nearest metropolitan area, the extreme temperatures supported in the area or the distance to the coast-. This exhaustive data at a high level of des-aggregation was obtained from the Spanish Censuses (1991, 2001 and 2011), published by the National Institute of Statistics, the Spanish National Centre for Geographical Research and from the Meteorological State Agency.

In general, our results for the 803 Local Labor Markets in which the Spanish territory can be divided confirm the existence of temporal and spatial heterogeneity in this matter. Even factors as important in Regional Economics as *distance* and *size* change their effects depending on the group under consideration, and also depending on where and when they are considered. Socio-economic factors that can be affected by specific policy action – such as employment - have a heterogeneous effect for both national and foreign-born population growth across space. Thus, trying to promote nationals´ population growth at local level or attract foreign-born individuals through policies aimed at creating employment can be successful in some local areas but have the opposite effect in others.

Our results show the population dynamics for nationals: Agglomeration diseconomies processes stand out in both decades in those urban areas located in the Valle del Ebro Axis, a process generally associated with negative externalities such as congestion, environmental degradation or higher housing prices. For the rest of the country, agglomeration economies are still operating. For foreign-born individuals size is always a positive feature to pursue, with spatial differences between areas but always with a positive relationship, which show the importance of agglomeration economies for this group.

In summary, the estimated spatial heterogeneity in national population growth and immigrants location choices' determinants warns against the general perspective presented by global approaches, given that factors may be significant in opposite ways in different areas, but average to zero or to a certain figure that might not be representative across the whole territory. Different policies should be aimed to attract national versus foreign-born population. In addition, policymakers should be informed of local socioeconomic processes, a need that stresses the importance of statistical methods that can reflect spatial heterogeneity, a common feature in these cases.

These methodology and results give empirical support to the new approaches adopted by the European Union Cohesion Policy on the *smart specialization strategies*. Instead of promoting global regional actions, this new perspective requires an independent regional status and expectations assessment in order to adapt the political intervention to each place. Our conclusion is that, at least for population evolution, a local smart specialization approach should be considered along with the traditional wider territorial scope, and that the analyses should consider, as with the GWR approach, the relevance of the characteristics of neighboring areas.

3. References

Aja, E., Carbonell, F., Ioé, C., Funes, J., and Vila, I. (2000). *La Inmigración Extranjera en España: Los Retos Educativos*. Fundación La Caixa.

Ali, K., Partridge, M.D. and Olfert, M.R. (2007). Can Geographically Weighted Regressions Improve regional Analysis and Policy Making? *International Regional Science Review*, 30 (3), pp. 300-329.

Anselin, L. (1988). *Spatial Econometrics: Methods and Models*. Boston: Kluwer Academic Publishers.

Beeson, P. E., DeJong, D. N. and Troesken, W. (2001). Population Growth in US Counties, 1840–1990. *Regional Science and Urban Economics*, 31 (6), pp. 669-699.

Bailey, A. (2005). *Making Population Geography*. London: Hodder.

Bauer T., Epstein G., Gang IN (2002) Herd effect or migration networks? The location choice of Mexican immigrants in the US IZA discussion paper 551

Bauer T., Epstein G., Gang IN (2005) Enclaves, language and the location choice of immigrants, Journal of Population Economics 18 (4): 649-662.

Boix, R. and Galleto, V. (2006). *Identificación de Sistemas Locales de Trabajo y Distritos Industriales en España.* Dirección General de Política de la Pequeña y Mediana Empresa, Ministerio de Industria, Comercio y Turismo.

Brunsdon, C., Fotheringham, S. and Charlton, M. (1996). Geographically Weighted Regression: a Method for Exploring Spatial Nonstationarity. *Geographical Analysis*, 28 (4), pp. 281–298.

Brunsdon, C., Fotheringham, S. and Charlton, M. (1998). Geographically Weighted Regression — Modeling Spatial Non-stationarity. *The Statistician,* 47 (3), pp. 431-443.

Cebula, R. J. (2002). Net Interstate Population Growth Rates and the Tiebout-Tullock Hypothesis: New Empirical Evidence, 1990–2000. *Atlantic Economic Journal*, 30 (4), pp. 414-421.

Cervero, R. (2003). Road Expansion, Urban Growth, and Induced Travel: A Path Analysis. *Journal of the American Planning Association*, 69 (2), pp. 145–163.

Chi, G. and Ventura, S. J. (2011). An Integrated Framework of Population Change: Influential Factors, Spatial Dynamics, and Temporal Variation. *Growth and Change*, 42 (4), pp. 549-570.

Christaller, W. (1935). *Die Zentralen Orte in Süddeutschland*. Jena: Gustav Fischer.

Cleveland, W.S. and Devlin, S.J. (1988). Locally Weighted Regression: An Approach to Regression Analysis by Local Fitting. *Journal of the American Statistical Association*, 83 (403), pp. 596–610.

CNIG (2001). Mapas Digitales de los Municipios Españoles. Centro Nacional de Información Geográfica, Ministerio de Fomento. http://centrodedescargas.cnig.es/CentroDescargas/equipamiento/BD_Municipios-Entidades.zip. Accessed 13 Dec 2011.

Coffey W. J. and Polèse, M. (1988). Locational Shifts in Canadian Employment, 1971-1981: Decentralization vs. Decongestion. *The Canadian Geographer*, 32 (3), pp. 248-256.

Conde-Ruiz, J. I., García, J. R. and Navarro, M. (2008). Inmigración y Crecimiento Regional en España. *Colección Estudios Económicos, 09-08*, Fundación de Estudios de Economía Aplicada.

Coombes, M. G., Green, A. E., and Openshaw, S. (1986). An Efficient Algorithm to Generate Official Statistical Reporting Areas: The Case of the 1984 Travel-To-

Work Areas Revision in Britain. *Journal of the Operational Research Society*, 37(10), pp. 943-953.

Cowen, D. J. and Jensen, J. R. (1998). Extraction and Modeling of Urban Attributes Using Remote Sensing Technology. In Liverman, D., Moran E. F., Rindfuss, R. R. and Stern, P. C. (eds.). People and Pixels: Linking Remote Sensing and Social Science. Washington: National Academy Press.

Del Valle, R. (1995). Hacia una Nueva División Espacial del Trabajo en España. *Revista de Estudios Regionales*, 42, pp. 131-170.

Duranton, G. and Puga, D. (2004). Micro-Foundations of Urban Agglomeration Economies. In Henderson, J.V. and Thisse J.-F. (eds.). *Handbook of Regional and Urban Economics*, vol. 4. Amsterdam: Elsevier.

Eeckhout, J. (2004). Gibrat's Law for (All) Cities. *The American Economic Review*, 94 (5), pp. 1429-1451.

Fotheringham, S., Brundson, C. and Charlton, M. (2002). *Geographically Weighted Regression: The Analysis of Spatially Varying Relationships*. Chichester: John Wiley.

Fuentes, F. and Callejo, M. (2011). *Inmigración y Estado de bienestar en España*. Fundación la Caixa.

Fujita, M., Krugman, P. R., and Venables, A. J. (1999). *The Spatial Economy: Cities, Regions and International Trade*. Cambridge, MA: MIT press.

Gabaix, X. (1999). Zipf's Law for Cities: An Explanation. *The Quarterly Journal of Economics*, 114 (3), pp. 739-767.

Gabaix, X. and Ioannides, Y. M. (2004). The Evolution of City Size Distributions. In Henderson, J.V. and Thisse J.-F. (eds.). *Handbook of Regional and Urban Economics*, vol. 4. Amsterdam: Elsevier.

Glaeser, E. L., Ponzetto, G. A. M. and Tobio, K. (2014). Cities, Skills and Regional Change. *Regional Studies*, 48 (1), pp. 7-43.

Glaeser, E. L. and Shapiro, J. M. (2003). Urban Growth in the 1990s: Is City Living Back? *Journal of Regional Science*, 43(1), pp. 139-165.

Goerlich, F. J. and Mas, M. (2008). Empirical Evidence of Population Concentration in Spain, 1900-2001. *Population*, 63 (4), pp. 1-16.

Goerlich, F. J. and Mas, M. (2009). Drivers of Agglomeration: Geography vs History. *The Open Urban Studies Journal*, 2, pp. 28-42.

Greene, W. H. (2000). *Econometric Analysis*, 4th ed. Upper Saddle River, NJ: Prentice Hall.

Henderson, J.V. (2010). Cities and Development. *Journal of Regional Science*, 50 (1), pp. 515-540.

Henderson, J.V. and Wang, H.G. (2007). Urbanization and City Growth: The Role of Institutions. *Regional Science and Urban Economics*, 37 (3), pp. 283-313.

INE (2007) Censo de Población, 2001. Instituto Nacional de Estadística. http://www.ine.es/jaxi/menu.do?type=pcaxis&path=%2Ft20%2Fe242&file=inebase&L=0. Accessed 26 Nov 2011.

INE (2012) Censo de Población, 1991, 2001 and 2011. Instituto Nacional de Estadística. http://www.ine.es/censos2011_datos/cen11_datos_res_col.htm. Accessed 4 Feb 2013.

ISTAT-IRPET (1989). *I Mercati Local del Lavoro in Italia*. Milan: Franco-Angeli

Jaeger D., (2000), Local Labor Markets, Admission Categories and Immigrant Location Choice. Manuscript, College of William and Mary, 2000

Jiang, B. and Jia, T. (2011). Zipf's Law for all the Natural Cities in the United States: A Geospatial Perspective. *International Journal of Geographical Information Science*, 25 (8), pp. 1269-1281

Kaushal, Neeraj, 2005. New Immigrant’s Location Choices: Magnets without Welfare, *Journal of Labour Economics*, Vol. 23, No 1 (January), pp: 59-80.

Krugman, P. R. (1991). Increasing Returns and Economic Geography. *Journal of Political Economy*, 99 (3), pp. 483–999.

Lanaspa, L., Pueyo, F. and Sanz, F. (2003). The Evolution of Spanish Urban Structure during the Twentieth Century. *Urban Studies*, 40 (3), pp. 567–580.

Le Gallo, J., and Chasco, C. (2008). Spatial Analysis of Urban Growth in Spain, 1900–2001. *Empirical Economics*, 34(1), pp. 59-80.

Lee, E. (1966) A theory of migration, *Demography* 3 (1): 47-57.

Leung, Y., Mei, C.L. and Zhang, W.X. (2000). Statistical Tests for Spatial Nonstationarity Based on the Geographically Weighted Regression Model. *Environment and Planning A*, 32(1), pp. 9-32.

Levkovicha O., and Rowendala, J., (2014). Location choices of highly educated foreign workers: the importance of urban amenities. Tinbergen Institute Discussion Paper, 093/VIII.

Lösch, A. (1944), *Die Räumliche Ordnung der Wirtschaft*, Jena: Gustav Fischer.

Lucas, R.E. (1988). On the Mechanics of Economic Development. *Journal of Monetary Economics*, 22 (1), pp. 3–42.

McMillen, D.P. (1996). One Hundred Fifty Years of Land Values in Chicago: A Nonparametric Approach. *Journal of Urban Economics*, 40 (1), pp. 100–124.

McMillen, D.P. and Redfearn, C.L. (2010). Estimation and Hypothesis Testing for Nonparametric Hedonic House Price Functions. *Journal of Regional Science*, 50 (3), pp. 712–733.

Mincer, Jacob, 1978. Family migration decisions. *Journal of Political Economy* 86, n 5 (august): 749-773.

Montalvo, J. (2009). La Inserción Laboral de los Universitarios y el Fenómeno de la Sobrecualificación en España. *Papeles de Economía Española*, 119, pp. 172-187.

Páez, A., Farber, S. and Wheeler, D. (2011). A Simulation-Based Study of Geographically Weighted Regression as a Method for Investigating Spatially Varying Relationships. *Environment and Planning A*, 43(12), pp. 2992-3010.

Partridge, M. D., and Rickman, D. S. (2003). The Waxing and Waning of Regional Economies: The Chicken–Egg Question of Jobs versus People. *Journal of Urban Economics*, 53(1), pp. 76-97.

Partridge, M.D., Rickman, D.S., Ali, K. and Olfert, M.R. (2008). The Geographic Diversity of U.S. Nonmetropolitan Growth Dynamics: A Geographically Weighted Regression Approach. *Land Economics*, 84 (2), pp. 241-266.

Parr, J. (2005). Cities and Regions: Problems and Potentials. *Environment and Planning A*, 40, pp. 3,009-3,026.

Pérez, P. and Serrano, L. (1998). Capital Humano, Crecimiento Económico y Desarrollo en España (1964-1997). *Revista Valenciana D'estudis Autonòmics*, 24, pp. 69-86.

Polese, M. and Shearmur, R. (2004). Is Distance Really Dead? Comparing Industrial Patterns over Time in Canada. *International Regional Science Review,* 27 (4), pp. 431-457.

Poston, D. and Frisbie, W. (2006). Ecological Demography. In Poston, D. L. and Micklin, M. (eds.) *Handbook of Population*. New York: Springer.

Pritchett, L. (2001). Where Has All the Education Gone? *The World Bank Economic Review*, 15 (3), pp. 367-391.

Romer, P. M. (1994). The Origins of Endogenous Growth. *The Journal of Economic Perspectives*, 8 (1), pp. 3-22.

Rozenfeld, H. D., Rybski, D., Gabaix, X. and Makse, H. A. (2011). The Area and Population of Cities: New Insights from a Different Perspective on Cities. *The American Economic Review*, 101 (5), pp. 2205-2225.

Rubiera, F. and Viñuela, A. (2012). From Local Units to Economic Regions in Spain. Where the Agglomeration Economies Make Sense. In Fernández, E. and Rubiera, F. (eds) *New Challenges from Data at Local Level. Defining the Spatial Scale in Modern Regional Analysis*. Advances in Spatial Science. Berlin: Springer.

Serrano, C. (2011). Déjà Vu? Crisis de Empleo y Reformas Laborales en España. *Revista de Economía Aplicada*, 56 (19), pp. 149-177.

Serrano, L. (1999). Capital Humano, Estructura Sectorial y Crecimiento en las Regiones Españolas. *Investigaciones Económicas*, 23(2), pp. 225-249.

Sforzi, F. and Lorenzini, F. (2002). I Distretti Industriali. In *L´Esperienza Italiana dei Distretti Industriali*. Istituto per la Promozione Industriale (IPI), Roma.

Sforzi, F. (2012). From Administrative Spatial Units to Local Labour Market Areas. Some Remarks on the Unit of Investigation of Regional Economics with Particular Reference to the Applied Research in Italy. In Fernández, E. and Rubiera, F. (eds) *New Challenges from Data at Local Level. Defining the Spatial*

Scale in Modern Regional Analysis. Advances in Spatial Science. Berlin: Springer.

Shearmur, R., Apparicio, P., Lizion, P., and Polèse, M. (2007). Space, Time, and Local Employment Growth: An Application of Spatial Regression Analysis. *Growth and Change*, 38(4), pp. 696-722.

Shearmur, R. and Polèse, M. (2005). Diversity and Employment Growth in Canada, 1971-2001: Can Diversification Policies Succeed? *The Canadian Geographer,* 49 (3), pp. 272-290.

Vias, A.C. (2012). Micropolitan Areas and Urbanization Processes in the US. *Cities*, 29 (supplement 1), pp. S24-S28.

Vining, D.R. and Pallone, R. (1982): "Migration between core and peripheral regions: a description and tentative explanation of the patterns in 22 countries", *Geoforum*, 13 (4), pp. 339-410.

Wheeler, D. and Tiefelsdorf, M. (2005). Multicollinearity and Correlation among Local Regression Coefficients in Geographically Weighted Regression. *Journal of Geographical Systems*, 7(2), pp. 161-187

Williamson, J. G. (1965). Regional Inequality and the Process of National Development: A Description of the Patterns. *Economic Development and Cultural Change*, 13 (4), pp. 1-84.

Zavodny M (1999) Determinants of recent immigrants' locational choices. International Migration Review 33: pp 1014-1030.

Zhang, K. H. and Song, S. (2003). Rural–Urban Migration and Urbanization in China: Evidence from Time-series and Cross-section Analyses. *China Economic Review*, 14(4), pp. 386-400.

Zipf, G. K. (1949). *Human Behavior and the Principle of Least Effort*. Oxford: Addison-Wesley Press.