Universidad de Oviedo
Universidá d'Uviéu
University of Oviedo

SPANISH ECONOMIC GEOGRAPHY: A STUDY OF THE LOCAL POPULATION AND EMPLOYMENT DYNAMICS IN SPAIN

Programa de Doctorado en Economía y Sociología de la Globalización

Diana Gutiérrez Posada

2016

Universidad de Oviedo

Departamento de Economía Aplicada

Spanish Economic Geography: A Study of the Local Population and Employment Dynamics in Spain

Memoria que, para la obtención del grado de doctor, presenta

Diana Gutiérrez Posada

Programa de Doctorado

"Economía y Sociología de la Globalización"

(Mención de Calidad)

Programa de Formación del Profesorado Universitario

(Ref. FPU12/00750)

Bajo la dirección de los profesores

Fernando Rubiera Morollón
Ana Viñuela

RESUMEN DEL CONTENIDO DE TESIS DOCTORAL

1.- Título de la Tesis	
Español/Otro Idioma:	Inglés: Spanish Economic Geography: A Study of the Local Population and Employment Dynamics in Spain

2.- Autor	
Nombre: DIANA GUTIERREZ POSADA	DNI/Pasaporte/NIE:
Programa de Doctorado: Economía Aplicada y Sociología de la Globalización	
Órgano responsable: ECONOMIA APLICADA	

RESUMEN (en español)

El principal objetivo de esta tesis es arrojar luz sobre los procesos de crecimiento de población y empleo que han tenido lugar en España en las dos últimas décadas (1991-2001 y 2001-2011). Aunque varios estudios han abordado estas cuestiones antes, aún existe un amplio margen para el análisis de estos temas desde una perspectiva local, teniendo en cuenta los efectos de la estructura espacial del país en su conjunto.

La primera parte de la tesis se focaliza sobre los procesos demográficos, analizando en el primer capítulo en el crecimiento poblacional, con énfasis en la influencia de los factores locales y los efectos de vecindad a lo largo de las décadas 1991-2001 y 2001-2011. El objetivo es identificar las respuestas locales a las características socioeconómicas, geográficas y ambientales, comparando metodologías de estimación globales y locales. En este sentido, la aplicación de Regresiones Geográficamente Ponderadas (GWR en inglés) permite analizar los diferentes efectos sobre los mercados locales de trabajo españoles cuando existe variabilidad significativa en el impacto de un determinado factor. Las implicaciones de este ejercicio van desde el interés académico en la evaluación del grado y las características de la heterogeneidad espacial, hasta la importancia de identificar aspectos particulares susceptibles de intervención en términos de política regional.

Como análisis complementario, esta parte incluye también un estudio análogo sobre el crecimiento de la población envejecida en los municipios españoles entre 2001 y 2014. El objetivo aquí es explorar las circunstancias del marcado proceso de envejecimiento que actualmente vive el país, en términos de convergencia o divergencia en la proporción de población mayor, y cómo los rasgos municipales afectan ese proceso, prestando especial atención a los efectos de la ubicación relativa y la posición dentro de la jerarquía urbana española.

Las dinámicas relacionadas con el ámbito laboral se analizan en la segunda parte de la tesis. El crecimiento del empleo se aborda en el tercer capítulo, con el objetivo de explicar los determinantes de la concentración del empleo en España y la relación entre las características de los mercados locales de trabajo, las características de sus respectivos vecinos, y su desempeño relativo en términos de empleo local. Esta sección presenta un análisis de econometría espacial instrumentado a través de un modelo Durbin que explica la autocorrelación espacial entre regiones, mostrando los vínculos de proximidad entre ellas con respecto a los factores explicativos incluidos. La estructura espacial también se refleja mediante el concepto de distancias incrementales, un conjunto de medidas que corrigen la distancia a un cierto nivel de la jerarquía urbana por la proximidad de otros lugares. Este enfoque empírico proporciona evidencia sobre el funcionamiento del mercado de trabajo español desde una perspectiva regional para las décadas de 1991-2001 y 2001-2011, y puede

FOR-MAT-VOA-010-BIS

servir como guía para la confección de medidas de política laboral en un ámbito local.

El último capítulo trata de contribuir a la escasa literatura sobre salarios locales en España, evaluando la evolución del premium salarial urbano y el impacto que la crisis ha tenido sobre este concepto desde su comienzo en 2008. Los datos sobre salarios agregados indican que su tasa de crecimiento cayó considerablemente tras el inicio de la crisis, presentando incluso cifras negativas en algunos sectores. El principal objetivo de este capítulo es analizar si esta caída de los salarios se traduce en un cambio en la brecha salarial existente entre las ciudades más grandes y las demás áreas del sistema urbano español, comparando el período anterior (2004-2007) y el período de crisis (2009-2013). Este estudio profundiza además en las diferencias locales y en las interrelaciones regionales, dando algunas pistas sobre la resiliencia relativa que ofrecen los municipios.

Las conclusiones resaltan la importancia de la heterogeneidad, tanto espacial como temporal, presente en todos los fenómenos estudiados. Factores que pueden favorecer el crecimiento en una determinada en faceta en un territorio, pueden tener un efecto adverso en otro, o cambiar la dirección de su influencia entre períodos, lo que lleva a la importancia de tener en cuenta las características concretas de cada área, la relación de esta con su entorno espacial, y el carácter dinámico de estas interacciones, para tomar decisiones de política acertadas.

RESUMEN (en Inglés)

The primary aim of this work is to shed some light on the processes of population and employment growth taking place in Spain in the last two decades (1991-2001 and 2001-2011). Although several studies have addressed these issues before, an unfilled gap remains regarding the analysis of these subjects from a local perspective while considering the effects of the spatial structure of the whole country. This dissertation will deal with the topic of local population growth in the first place, and then it will engage in the study of employment growth, to end with an introductory analysis of local differential wages associated to the urban wage premium.

The first part of this dissertation focuses on demographic processes, analyzing in the first chapter population growth with an emphasis on the influence of local factors and neighborhood effects along the decades 1991-2001 and 2001-2011. The objective is to identify local responses to socio-economic, geographical and environmental features, confronting global and local estimation methodologies. In this sense, the application of geographically weighted regressions (GWR) allows for the appreciation of differentiated local effects across the Spanish local labor markets when there exists significant variability in the impact of a certain factor. The implications of this exercise range from the academic interest on evaluating the degree and characteristics of spatial heterogeneity, to the importance of identifying particular aspects susceptible to regional policy intervention.

As a complementary analysis, this section also includes an analogous study on the growth of ageing population in Spanish municipalities between 2001 and 2014. The aim here is to learn about the acute ageing process in terms of the convergence/divergence in the share of older population, and how municipal traits affect that process, with special attention to the effects of the relative location and the position in the urban hierarchy.

Labor dynamics are explored in the second part of the thesis. Employment growth is dealt with in the third chapter, which seeks to explain the determinants behind the employment concentration in Spain and the relationship between the characteristics of the local labor

markets, the characteristics of their neighbors, and their relative performance in terms of local employment. This part presents a spatial econometric analysis by means of a Durbin model that accounts for the spatial autocorrelation among regions, while showing the proximity links between them with respect to the explanatory factors included. The spatial structure is also reflected in reference to the complete urban hierarchy through the concept of incremental distances, a set of measures that correct the distance to a certain urban tier by the proximity of other locations. This empirical approach provides evidence on functioning of the Spanish labor market from a regional perspective during 1991-2001 and 2001-2011, and may help as a guide for labor policy instrumentation.

The last chapter attempts to contribute to the scarce literature on local wages in Spain, by assessing the evolution of the urban wage premium in the country and the impact that the Great Recession in 2008 has on this notion. Correspondingly, data on aggregated wages indicates that their growth rate dropped considerably after the outbreak of the crisis, exhibiting even negative numbers in some sectors. The aim of this chapter is to analyze if the drop in wages translates into a change in the existing gap between the largest urban municipalities and the other areas by comparing the pre-crisis (2004-2008) and during-crisis (2009-2014) scenario, and which of the components of this urban wage premium (sorting, agglomeration, price, learning or coordination effect) are more sensitive to the shock. This study goes deeper into the local differences and regional interrelationships, giving some hints about the relative resilience of municipalities.

Conclusions remark the importance of heterogeneity, both spatial and temporal, involved in almost every aspect of the analyzed phenomena. Factors that enhance growth in any or the regarded dimensions in one region, may have the contrary effect in other, or change the direction of their influence depending on the period observed. Hence, it is crucial to take into account the specific features of the area, its relationship with the urban structure, and the dynamic character of that relationship, in order to develop successful local policy measures.

SR/A. DIRECTOR/A DE DEPARTAMENTO DE ECONOMIA APLICADA/
SR/A. PRESIDENTE/A DE LA COMISIÓN ACADÉMICA DEL PROGRAMA DE DOCTORADO EN Economía Aplicada y Sociología de la Globalización

A mi madre, por su ejemplo.

A mi hermano, por ser mi compañero de odisea.

A mis directores, por su paciencia.

A mis compañeros de la Lewis, por compartir cafés, risas y penas; especialmente Vane, Gabriel y Bego.

Al REGIOlab, por la oportunidad.

To REAL, for being the best mafia ever, led by the hippiest capo dei capi.

A André, por su apoyo incondicional, por hacer mi mundo más grande, por quemar los espacios en blanco, y porque definitivamente no sería lo mismo imaginarte.

Y a todos los que han contribuido a que este barco llegue a buen puerto.

Resumen

El crecimiento económico mundial presenta una distribución claramente desigual a todos los niveles espaciales. De países a comunidades locales, todos los factores socioeconómicos estudiados (salarios, educación, empleo, etc.) parecen tener un reparto muy desigual, y las diferencias crecen cuando nos centramos específicamente en aspectos demográficos. La población urbana ha ido aumentando continuamente en las últimas décadas, dando lugar a grandes metrópolis que además también concentran la actividad económica. Aunque la fuerza de los procesos de aglomeración es fácilmente palpable. En el campo de la concentración económica y de la población aún quedan muchos temas por investigar, ya que el equilibrio entre las fuerzas centrífugas y centrípetas, que da lugar a patrones espaciales, va a depender de las interrelaciones entre las distintas ciudades (en términos de tamaño y especialización) que forman el sistema urbano de un país, y de la evolución a través del tiempo de esos vínculos.

En los últimos años, la preocupación por estas disparidades espaciales persistentes se ha convertido en una cuestión política fundamental, lo que convierte en necesaria la investigación sobre los mecanismos que las generan, con el último propósito de proporcionar herramientas para la instrumentación de políticas. Esta necesidad de ofrecer explicaciones rigurosas de fenómenos de localización y aglomeración promovió el desarrollo de la ciencia regional, un campo de conocimiento que abarca una amplia gama de temas, como los relacionados con la geografía económica y urbana, la planificación o el transporte, entre otros. El objetivo que aúna estos diferentes puntos de vista es la comprensión de los rendimientos crecientes procedentes de la concentración espacial, que a su vez impulsan una mayor aglomeración. Entendiendo cómo reaccionan y cambian estos rendimientos frente a diferentes estímulos sería posible comprender el comportamiento del sistema económico en general y diseñar políticas más eficientes, teniendo en cuenta el componente espacial inherente.

La distribución territorial de la población es un indicador de múltiples dimensiones económicas, ya que revela la preferencia de los individuos respecto a la localización, en un contexto de maximización de su utilidad. En el equilibrio entre las fuerzas centrípetas y

centrífugas que impulsan la aglomeración, la elección sobre la localización debe tener en cuenta aspectos como las condiciones del mercado de trabajo o la calidad de vida, que a su vez se entrelazan con otras características relacionadas con la estructura productiva o el marco institucional de las diferentes alternativas. Todos estos aspectos están en el centro de las disparidades regionales que la ciencia regional pretende comprender, de ahí la importancia del estudio de las dinámicas de la población.

En España, más del 20% de la población vive en municipios de más de 300.000 habitantes desde 1970, un porcentaje que se incrementa hasta alrededor del 40% cuando se consideran municipios de más de 100.000 habitantes. Este proceso de urbanización también es visible en el incremento del número de ciudades por encima de esos valores. En 1950, 23 municipios tenían más de 100.000 habitantes y sólo 3 superaban los 500.000, mientras que en 2011 el primer grupo casi triplica y el segundo duplica su tamaño. Estos cambios en el diseño espacial configuran la estructura urbana del país y afectan las condiciones económicas en las que las diferentes regiones interactúan entre sí. En el caso Español, el sistema de municipios de más de 300.000 habitantes presenta una clara orientación hacia la costa mediterránea, que concentra no sólo la población, sino también la actividad económica.

Las diferencias espaciales pueden encontrarse no sólo en la concentración territorial de la población en su conjunto, sino también en su composición demográfica. En este sentido, un fenómeno que está ganando interés en los últimos años es el del envejecimiento, y las distintas respuestas políticas que deben adoptarse con respecto a la previsión del aumento de la proporción de población en edad avanzada en los países desarrollados. Diferentes tasas de crecimiento de la población pueden tener diversos efectos sobre la estructura por edades y, por consiguiente, las cuestiones relativas a la regeneración demográfica pueden mostrar un comportamiento diferente en un lugar u otro. Una vez más, la perspectiva temporal también es importante dado el carácter dinámico de esta circunstancia.

En el caso de España, el proceso de inversión en el equilibrio demográfico es claro. En 1971 se puede observar la estructura de un país en desarrollo en términos demográficos, con una amplia base de población más joven y no mucha gente mayor en proporción. A medida que pasa el tiempo, la proporción de población joven se hace más pequeña y los niños nacidos durante el baby-boom de los años 60 y 70 avanzan a edades adultas. Hoy en día, la perspectiva del envejecimiento es un desafío: hay una gran parte de la población entre 35 y 49 años de edad, mientras que la población más joven no está creciendo en la misma proporción, y los estudios sobre el tema indican que esta situación no será fácil de atajar. La necesidad de investigación capaz de proporcionar pautas para la confección de políticas se hace evidente en este contexto.

Uno de los ámbitos económicos en los que los hechos expuestos hasta este punto tienen implicaciones de gran relevancia es el mercado de trabajo. El crecimiento y la estructura de la población determinan la mano de obra disponible para llevar a cabo la producción, que se traduce en crecimiento económico. Por otra parte, las condiciones laborales son uno de los factores más importantes al considerar opciones de localización, definiendo en gran medida el potencial de atracción de una región.

El mercado de trabajo español creció significativamente en las dos últimas décadas. La gran mayoría de los inmigrantes que llegaron en la primera mitad de la década vinieron tras las oportunidades laborales que el dinamismo de la economía proporcionaba. Esta tendencia alcista finalizó en 2008 con la Gran Recesión, trayendo consigo un mayor desempleo, por encima de los desequilibrios estructurales existentes, que se ignoraron durante en el período de prosperidad. La tasa de empleo nacional ha disminuido de forma continua desde el tercer trimestre de 2007 hasta el primer trimestre de 2013, lo que representa una caída de más del 10%. Un hallazgo interesante es que la diferencia entre las regiones con mejor y peor desempeño es persistente en el tiempo, independientemente del punto en el ciclo económico, ya que todas siguen una tendencia muy similar. Solamente las islas Canarias, la región más lejana del país, exhibe un comportamiento distinto. Las comunidades autónomas que gozan de una mejor posición en términos de empleo son Madrid y Cataluña, que contienen las ciudades más grandes del país (Madrid y Barcelona), y Navarra, una región con un régimen fiscal diferenciado. En el otro lado de la distribución encontramos Asturias, Andalucía y Extremadura, tres regiones relativamente alejadas de los principales centros urbanos. Estas regiones ilustran la importancia de la posición relativa en la estructura urbana y del marco institucional en la explicación del desempeño económico y las diferencias territoriales.

Por lo tanto, la distribución de la población y el empleo en España hace que el país sea un caso de estudio más que interesante. El sistema urbano, la relación entre las distintas áreas definidas y el análisis de las disparidades regionales que surgen en el curso de la siempre cambiante composición espacial del país ofrecen una gran cantidad de temas de investigación. Al mismo tiempo, las interrelaciones entre los aspectos demográficos y económicos en juego hacen de su análisis una tarea desafiante y fascinante, especialmente si uno se atreve a ir más allá, hasta los niveles regionales más desagregados donde la variedad de escenarios y resultados crece notablemente.

Las regularidades empíricas presentadas establecen un marco sólido para el desarrollo de varias líneas de investigación. Sobre esta base, el objetivo principal de esta tesis es arrojar luz sobre los procesos de crecimiento de población y empleo que han tenido lugar en España en las dos últimas décadas (1991-2001 y 2001-2011). Siguiendo el esquema expuesto en la descripción del caso español, esta tesis abordará el tema de las dinámicas locales de la población en la primera parte, y luego se abordará el estudio de las dinámicas relacionadas con el empleo en la segunda.

La primera parte de la tesis se focaliza en el crecimiento demográfico, con énfasis en la influencia de los factores locales y los efectos de vecindad a lo largo de las décadas 1991-2001 y 2001-2011. El objetivo es identificar las respuestas locales a las características socioeconómicas, geográficas y ambientales, confrontando metodologías de estimación globales y locales. En este sentido, la aplicación de Regresiones Geográficamente Ponderadas (GWR en sus siglas en inglés) permite apreciar los diferentes efectos en los mercados de trabajo locales españoles, en el caso de que exista una variabilidad significativa en el impacto de un determinado factor. Las implicaciones de este ejercicio van desde el interés académico en la evaluación del grado y las características de la heterogeneidad espacial, hasta la importancia de identificar aspectos particulares susceptibles de intervención en términos de política regional. Completando esta parte de la tesis, esta sección también incluye un estudio análogo sobre el crecimiento del envejecimiento de la población en los municipios españoles entre 2001 y 2014. El objetivo aquí es conocer las dinámicas del marcado proceso de envejecimiento en términos de convergencia/divergencia en la proporción de población mayor, y cómo los rasgos municipales afectan a ese proceso, prestando especial atención a los efectos de la ubicación relativa y la posición dentro de la jerarquía urbana española.

La cuestión del crecimiento del empleo se aborda en la segunda parte de la tesis, que pretende explicar los determinantes de la concentración del empleo en España y la relación entre las características de los mercados locales de trabajo, las características de los mercados vecinos, y su desempeño relativo en términos de empleo local. Esta parte presenta un análisis de econometría espacial, realizado mediante la estimación de un modelo Durbin, que explica la autocorrelación espacial entre regiones, mostrando los vínculos de proximidad con respecto a los factores explicativos incluidos. La estructura espacial también se refleja a través del concepto de distancias incrementales, un conjunto de medidas que corrigen la distancia a un cierto nivel urbano por la proximidad de otros lugares. Este enfoque empírico proporciona evidencia sobre el funcionamiento del mercado de trabajo español desde una perspectiva regional para las décadas de 1991-2001 y 2001-2011, y puede servir como guía para las medidas de política laboral local. El último capítulo, que se incluye en esta sección, trata de contribuir a la escasa literatura sobre salarios locales en España, evaluando la evolución del premium salarial urbano en el país, y el impacto que, desde 2008, ha tenido la crisis sobre este concepto. En un principio, los datos respecto a salarios agregados indican que su tasa de crecimiento disminuyó considerablemente tras el inicio de la crisis, mostrando cifras incluso negativas para algunos sectores. El principal objetivo de este capítulo es analizar si esta caída en los salarios se traduce en un cambio en la brecha existente entre los municipios urbanos más grandes y las demás áreas, comparando el período anterior a la crisis (2004-2007) y el período de crisis (2009-2014). Este estudio profundiza además en las diferencias locales y en las interrelaciones regionales, dando algunas pistas sobre la resiliencia relativa que presentan los municipios.

La definición de "local" en los primeros capítulos de cada parte estará representada por los mercados de trabajo locales, dado el interés en comprender cómo la población y el empleo se distribuyen entre las regiones españolas desde una perspectiva amplia pero funcional.

En este sentido, los mercados laborales locales encierran comportamientos residenciales y laborales dentro de una misma área, convirtiéndolos en una definición espacial adecuada para estudiar la concentración. Cada parte incluye también un capítulo adicional centrado en un tema específico relacionado con la población o el empleo. Como se mencionó anteriormente, el capítulo final de la primera parte considera el aumento de la proporción de la población envejecida, mientras que en la segunda parte el último capítulo examina la brecha salarial espacial y su reacción ante la crisis reciente. En ambos casos, el enfoque se dirige al nivel más espacialmente desagregado disponible, en este caso los municipios, para desvelar la dinámica subyacente a estos fenómenos de la manera más precisa posible, en contraste con el ángulo general adoptado en los capítulos de apertura de cada parte.

A pesar de que varios estudios han abordado los temas aquí discutidos, subsiste un nicho de investigación no tan explorado en cuanto al análisis de estos temas desde una perspectiva local, considerando a la vez los efectos de la estructura espacial de todo el país en su conjunto. La principal razón reside en la dificultad de encontrar datos adecuados a un nivel de desagregación elevado. La publicación de los censos poblacionales más recientes y la aparición de nuevas fuentes de datos han posibilitado la realización de propuestas de investigación como la presentada aquí, sin embargo la limitación respecto a los datos sigue planteando un grave obstáculo para el análisis regional. Esta tesis trata de dar una visión general de la geografía económica española y de señalar su relevancia, con la aspiración de que en un futuro próximo la información y los instrumentos que se encuentren al alcance permitan hacer más y, lo que es más importante, hacer mejor investigación.

Table of contents

INTRODUCTION

1. A view through the lens of Regional Science: some subjects and tools

All over the world economic growth shows a clearly uneven distribution at all spatial levels. From countries to local communities, every socioeconomic factor studied (wage, education, employment, etc.) seems to be very unequal, and differences grow wider when the focus is turned on population. Urban population has been continuously increasing in the last decades, giving rise to large metropolitan cities which concentrate economic activity. Although the power of agglomeration can be easily seen, the question about population and economic concentration still has room for complexity, since the balance involving centripetal and centrifugal forces that results in non-arbitrary spatial patterns will depend on the relationships between the diverse cities in terms of size and specialization that form the urban system of a given country, and the evolution through time of those linkages.

In the last years, the concerns about this persistent spatial disparities have become a main policy matter in several countries, which stress the necessity of research aiming to understand the mechanisms behind them, while providing with tools for policy instrumentation. This need to offer rigorous descriptions and explanations of the location, connectivity and agglomeration phenomena fostered the development of regional science, a field of knowledge that currently covers a broad range of regional and urban topics from various perspectives that include regional economics, economic geography, urban planning or transportation science, among others. The idea that brings together the different economic points of view concerned is the understanding of the increasing returns coming from spatial concentration, which boost further agglomeration. By learning how this returns react and change to different stimuli, it would be possible to disentangle the behavior of the entire economic system and design efficient policies.

Despite its relatively recent blooming by the hand of an increased regional interest from the political arena, spatial issues are an old concern in economics. Classical academics as Adam Smith or David Ricardo regarded the role of location and connectivity as important economic characteristics. Salient precursors of regional science, whose contributions have a deep impact on nowadays conceptualization of the field were von Thünen (1826) with his model of land use, Weber (1909) with the study of optimal plant location, or Christaller (1933) and Lösch (1940) with the central-place theory. But it was thanks to the influential contributions of Walter Isard in the 1950s that regional science became a discipline with a sound theoretical basis and methodology, which can give proper answers about the processes at work in the real world. Many challenges posed by Isard remain untackled and serve as motivation for modern regional scientists.

One of the essential subjects of regional science is the concentration of population and economic activity, which is closely related to location choices and agglomeration effects. In turn, these ideas are connected to the concepts of proximity and economies of scale that are key for understanding the spatial behavior of households and firms. In this sense, regional science presents a solid analytical background to examine the drivers behind these aspects as its theory analyzes industrial, residential and facility location, accounting for their spatial concentration and interaction, trying to shed light on the spatial behavior of economic agents from a geographical point of view. This locational perspective also has a marked influence on applied research that can be seen in studies regarding growth poles or industrial districts. In terms of regional economic growth, the emphasis lies on the factors behind the spatial disparities and the influence they exert on existing patterns. The contributions in these regards highlight the value of both theoretical advances and empirical evidence that might conduct to unexplored policy alternatives. Another important lines of research in regional science are the ones focused on labor and housing markets, since these two contexts reflect firm and residential location decisions that might result in different agglomeration designs. Migration, job search, urban sprawl or real estate are some of the topics discussed is these strands of the literature.

Apart from the broad scope of perspectives and topics, another remarkable aspect of regional science is its array of tools specifically developed to deal with spatial data, which enables the performance of analyses of various sorts and purposes. In this context, the instruments used should account for the spatial structure of the data and its changing nature, a dimension that traditional methods fail to capture given the conflict with the independence principle induced by the likely presence of spatial autocorrelation and its lack of uniformity across the territory: spatial data can display different patterns on different parts of a map and at different geographical scales. These particular features yield strong implications for the construction and application of models suitable for inference. To this respect, spatial statistics and spatial econometrics offer a variety of sophisticated techniques that improved significantly the exploratory data analysis and the modeling of spatial autocorrelation processes that are at the core of regional science.

The origin of the statistical advances treating spatial issues can be attributed to the paper of Whittle (1954), who noted the necessity of considering the multidirectional dependence in spatial analysis, and Besag (1974), with the specification of a conditional spatial autoregressive model. Another pioneering work in this field was done by Cliff and Ord (1973), who reviewed earlier attempts to formally detect spatial autocorrelation as the ones devised by Geary (C test) or Moran (I test), and improved them via the inclusion of a spatial weighting structure in their designs. This ground-breaking contribution opened several doors for the construction of regression models and the simulation of spatial spillover effects, which fostered the creation of mechanisms to test the hypotheses raised by spatial economic theory. Paelinck and Klaassen (1979) named this toolbox “spatial econometrics”, a discipline concerned with the specification, fitting and testing of models where the spatial configuration of the data plays a salient role in the explanation of the behavior of a given dependent variable. Anselin (2010) offers a comprehensive review on the developments in this field.

From a technical perspective, spatial econometrics deals with spatial dependence through the introduction of lagged variables reflecting the characteristics of the neighbors and/or the separation between random and spatially dependent error terms. In this kind of models, the spatial structure is represented by a weighting matrix indicating the strength of the relationship between regions, which normally takes the form of contiguity or distance decay measurements, but it may contain other indicators of economic interaction as, for example, trade or population flows. Alongside the matter of spatial autocorrelation, heterogeneity stemming from the different settings considered, or the spatially interconnected response of those settings to a certain shock, is another key element inherent to spatial analysis. If heterogeneity comes from distinctive circumstances that are independent from one region to another, it can be modeled by means of dummy variables as proposed by Anselin (1988), but if the shock studied presents a continuous gradient across space and is related to local dynamics, finer approaches to apply include the expansion method by Jones and Casetti (1992) and the geographically weighted regressions by Fotheringham et al. (2000).

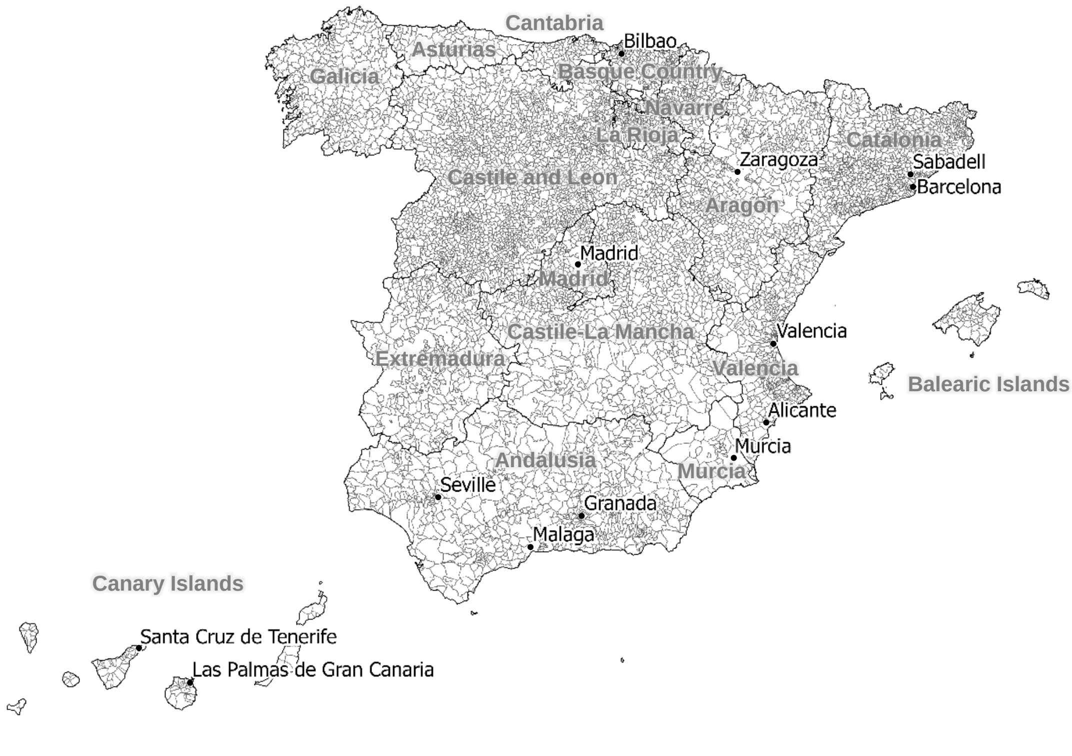
The improvement in the available computational capabilities and the renewed interest from public administrations on high-quality disaggregated data allow for the elaboration of more complex spatial analyses that can capture the economic processes at work better. These advances can be traced from the first stages of data management and preliminary examination, up to the development of more flexible modeling techniques. In this sense, the new methods of the exploratory spatial data analysis (ESDA) can provide with more accurate descriptions of patterns in increasingly large spatial databases (Miller and Han, 2009), while remarkable progress have been done in the field of Bayesian inference in order to deal with the uncertainty associated to different levels of the statistical and econometric procedures (Cressie and Wikle, 2011). Although there are still data and methodological challenges to work on, they are nothing but incentives that will support the progression in the understanding and modeling of spatial processes.

2. Spain on the spotlight: population and employment dynamics

The spatial distribution of the population is an indicator of several economic dimensions, as it reveals the preference of individuals regarding location decisions in a context of utility maximization. In the balance between the centripetal and centrifugal forces that drive agglomeration, individual location choices should take into account issues such as the labor market conditions or the living standards, which are intertwined with other features related to the productive structure or the institutional framework of different alternatives. All these aspects are at the center of the regional disparities that regional science aims to understand, hence the importance of tracing population.

In Spain, more than 20% of the population lives in municipalities of more than 300,000 inhabitants from the 1970s, a share that increases to around 40% when municipalities of more than 100,000 inhabitants are considered. This urbanization process does not only appear in the concentration figures just presented, it is also visible in the increment of the number of cities above those population numbers. In 1950, 23 municipalities had more than 100,000 inhabitants and only 3 went over 500,000, while in 2011 the first group almost triples and the second one doubles in size. These changes in the spatial design shape the urban structure of the country and affect the economic conditions in which the different regions interact with one another. As portrayed in Figure 1, the system of municipalities of more than 300,000 inhabitants in Spain presents a clear orientation towards the Mediterranean Coast, which concentrates not only population but also economic activity.

Figure 1. Spanish municipalities of more than 300,000 inhabitants in autonomous communities in 2011



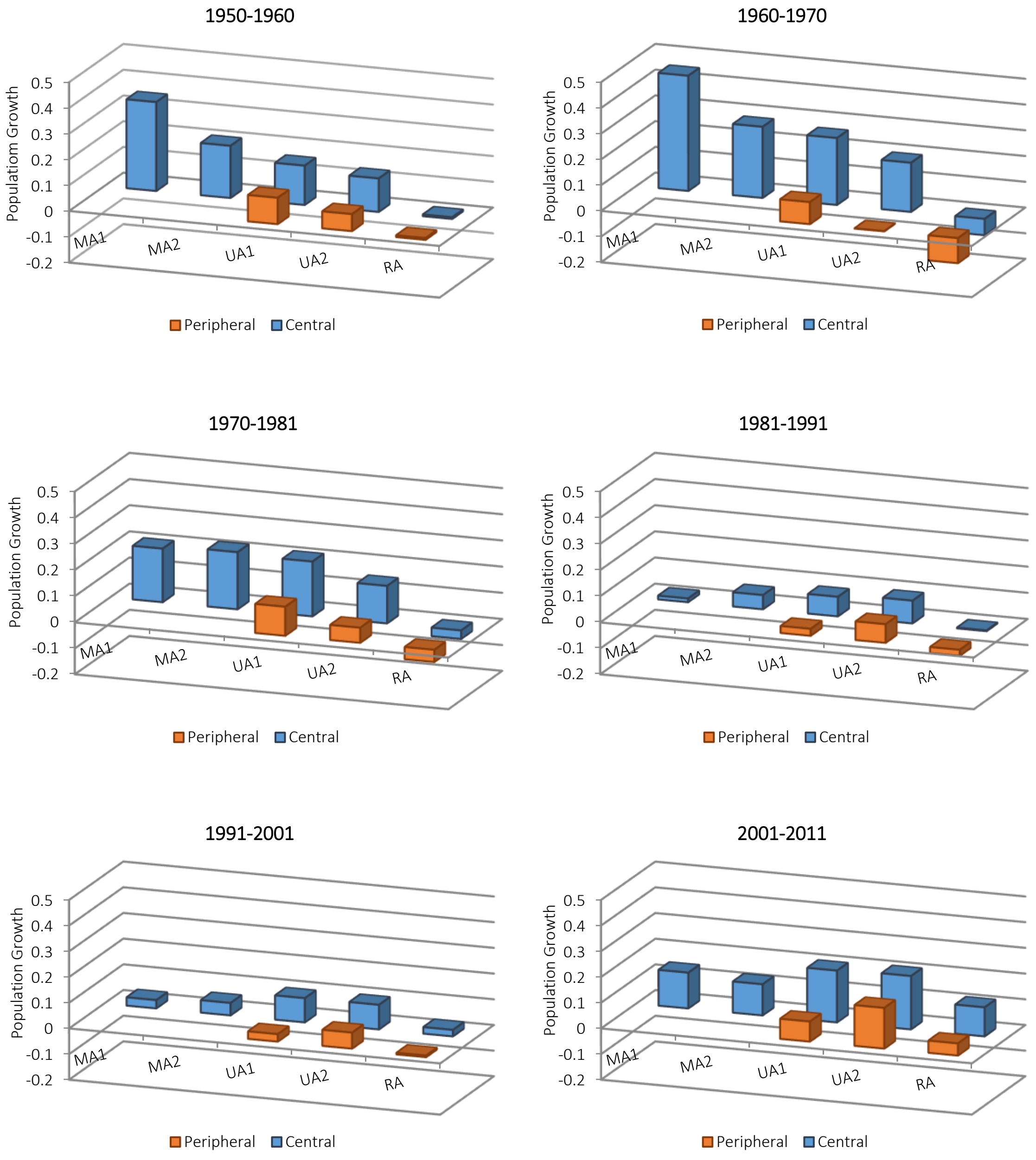
Source: Own elaboration with data from the National Geographical Institute (IGN) and the censuses of the National Statistics Institute (INE).

Administratively defined areas are one of the several ways to configure and analyze space, but functional ones can offer new and interesting insights for research. In Figure 2 it can be seen how the population growth in Spanish local labor markets (LLMs), as defined by Boix and Galleto (2006) considering commuting flows between municipalities, changed along the decades not only in magnitude but also with respect to the urban hierarchy. The vertical axis shows the population growth rate, while in the horizontal axis LLMs rank from the most populated metropolitan areas (Madrid and Barcelona – MA1), followed by smaller metropolises between 1,000,000 and 500,000 inhabitants (MA2), then by urban areas of more or less than 100,000 inhabitants (UA1 and UA2 respectively), and finally rural areas with less than 50,000 inhabitants. Urban and rural areas are also divided as central or peripheral regarding their proximity to the metropolitan areas.

During the 1950s and 1960s the largest metropolitan areas grew at a considerably higher rate than the rest of the country. All the tiers display a clear descending growth pattern as their size gets smaller, and as distance from the urban centers grows longer (peripheral areas). The profile showed in these two decades is the one of a country facing the effects of the rural exodus and the start of an urban concentration process as a consequence of the lower labor requirements in the agricultural sector.

The graphs for the following three decades present a change in the concentration trend, which ends up with a reversal of the relationship between population size and population growth. The differences between the areas shrink in the 1980s, especially in the case of the urban areas, as the main metropolitan LLMs grew at a significantly lower pace, while the rest remained relatively stable. Although peripheral areas lag behind in terms of grow, they perform better in both absolute and comparative terms regarding the previous period, and the depopulation of rural areas is also less pronounced. The change in the classification of population growth is clear throughout the 1980s and 1990s, as can be seen in Figure 2. Despite growth rates are lower in general, urban areas of less than 100,000 inhabitants lead the ranking in these years. This shift is consistent with the notion of agglomeration diseconomies related to the disadvantages that may appear in large urban centers (congestion, higher cost of living, etc.), which in turn may favor the growth of the so called micropolitan areas (Vias, 2012).

Figure 2. Population growth rates in Spanish LLMs classified by size and distance to the main areas between 1950 and 2011 (decades)



Source: Own elaboration using data from the censuses of the National Statistics Institute (INE).

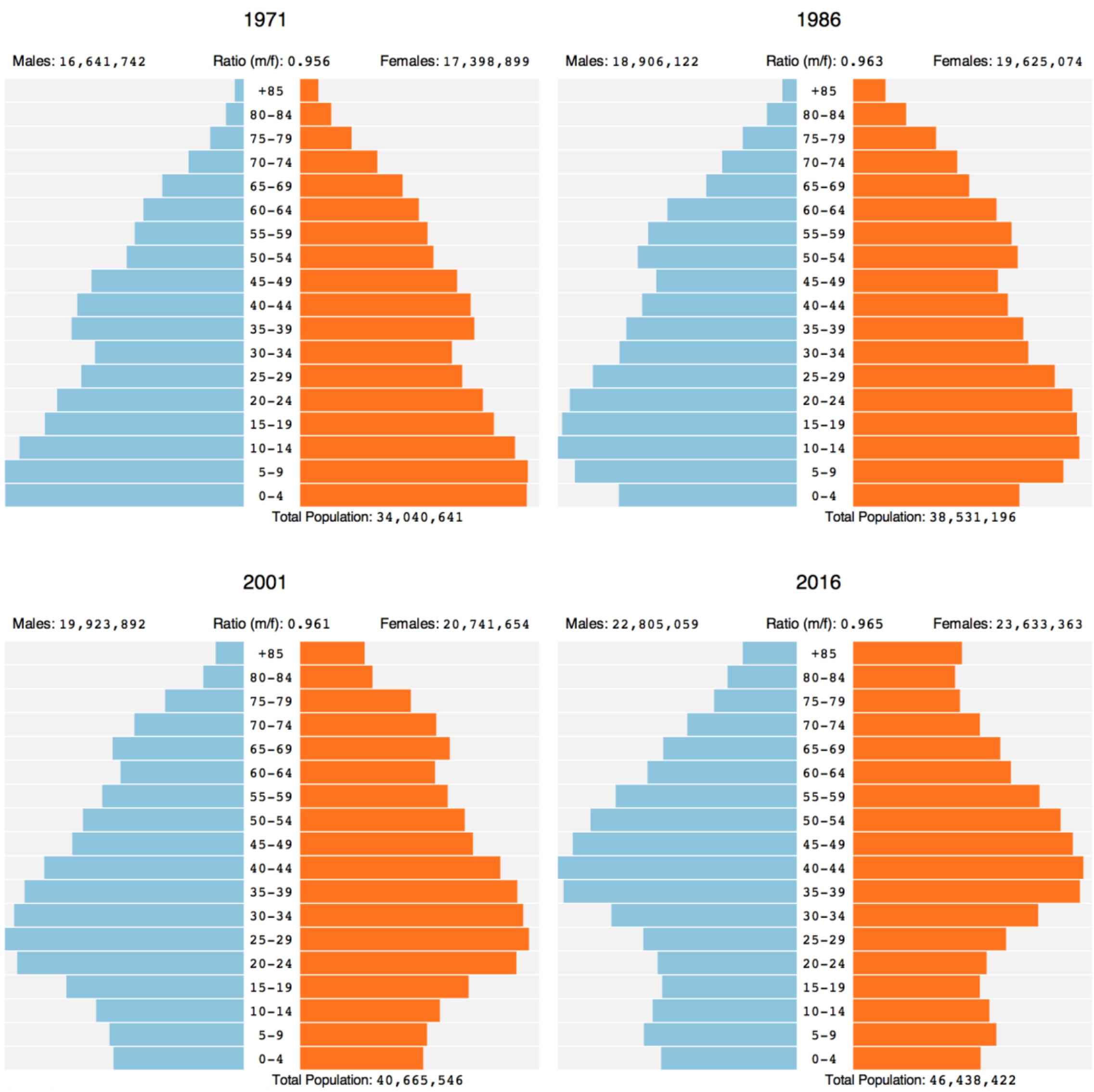
Finally, with the beginning of the new century further changes took place, this time brought by the massive immigration wave that arrived to the country. A consequence of this fact is the evident increase in the population growth rates of every tier of the hierarchy represented. Foreigners in Spain grew from 2% to around 12% of the total population, although this inflow did not distribute evenly across the country. In the 2000s smaller urban areas retain the top position of population growth, however the large metropolises show signs of an incipient resurgence, mainly due to the arrival of migrants.

Differences can be found not only in the spatial concentration of population as a whole, but also in its demographic composition. In this sense, an interesting phenomenon that is gaining interest in the last years is ageing, and the alternative policy responses that should be adopted with regard to the expected increase in the share of older population in developed countries. Different population growth rates can have diverse effects on the age structure, and consequently the issues concerning demographic regeneration might display a different behavior in one place or the other. Again, the time perspective is also important given the dynamic and overlapping character of this circumstance.

In the case of Spain, looking at the changes between the population pyramids in Figure 3, the process of inversion in the demographic balance is clear. In 1971 the pyramid shows the structure of a developing country, with a wide base of younger population and a much narrower tip. The notch in the ages between 25 and 34 reflects the demographic impact of the Spanish Civil War. As time passes, the base of the pyramid becomes smaller, and the children born during the baby-boom of the 1960s and 1970s advance to higher levels. Nowadays, the ageing perspective is a challenging one: there is a very large share of population between 35 and 49 years old, while the younger population is not growing accordingly, and studies on the matter indicate that this situation will not be easily changed provided the current economic situation and its anticipated long term consequences (Adsera, 2005). The need for research able to provide guidelines for policy instrumentation becomes evident in this setting.

One of the economic domains where the facts exposed up to this point have far-reaching implications is the labor market. Population growth and structure determine the available labor force to undertake the production tasks that will translate into economic growth. Conversely, labor conditions are one of the more important factors behind locational choices, largely defining the attraction potential of a region.

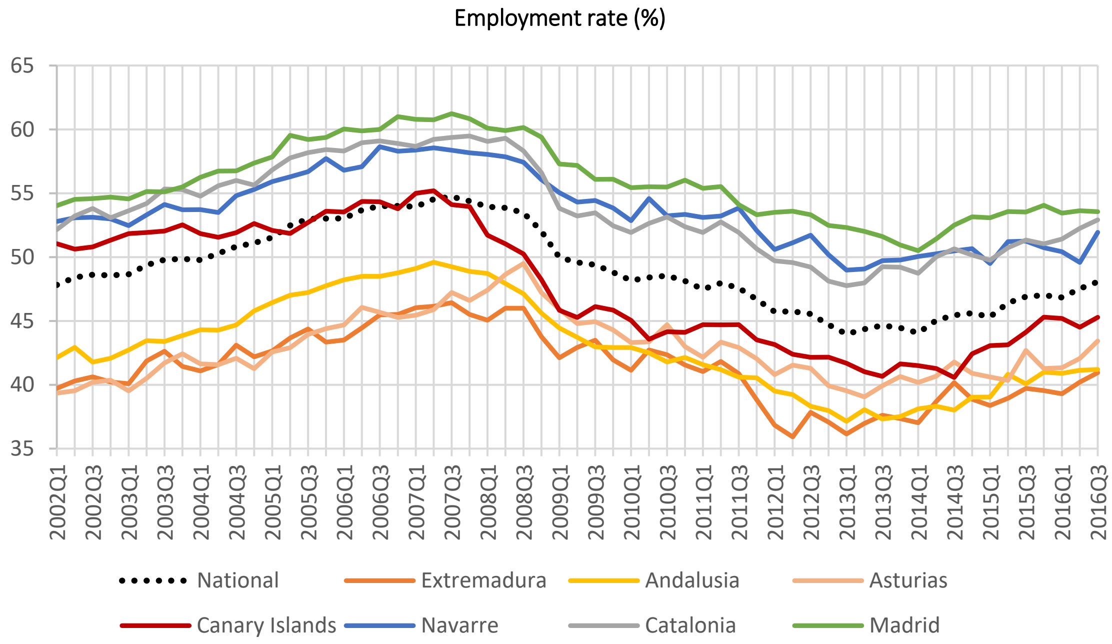
Figure 3. Population pyramids for Spain in 1971, 1986, 2001 and 2016



Source: Own elaboration using data from the censuses of the National Statistics Institute (INE).

Spanish labor market flourished in the beginning of the 21st century. The vast majority of the immigrants that arrived in the first half of the 2000s came following the labor opportunities that the dynamism of the economy provided. This upward trend finished in 2008 with the Great Recession, bringing along higher unemployment on top of the structural imbalances that remained concealed in the prosperity period. In Figure 4 it can be seen how the national employment rate (dotted line) falls steadily from the third quarter of 2007 until the first quarter of 2013, accounting for a drop of more than 10%.

Figure 4. Quarterly employment rate in a selection of Spanish autonomous communities



Source: Own elaboration using data from the Labor Force Survey (EPA) of the National Statistics Institute (INE).

The other lines in the graph represent a selection of the three better (Madrid, Catalonia and Navarre) and three worst (Extremadura, Andalusia and Asturias) autonomous communities regarding employment rate. An interesting finding is that the difference between the top and the bottom regions is persistent in time regardless of the point in the economic cycle, as they follow a very similar trend. Only the Canary Islands (burgundy line), the furthermost region of the country, exhibit a distinct behavior. The gap between regions remains stable around 10 percentage points. The autonomous communities that enjoy a better position in terms of employment are Madrid and Catalonia, which contain the largest cities in the country (Madrid and Barcelona), and Navarre, a region with a special fiscal regime. On the other side of the distribution are find Asturias, Andalusia and Extremadura, three regions relatively far from the main urban centers as can be seen in Figure 1. These regions illustrate the importance of the relative position in the urban structure and the institutional framework at explaining economic performance and territorial differences.

As depicted so far, population and employment distribution in Spain present a broad scope of situations across space and along time that makes the country an interesting case study. The urban system, the relationship between the different areas defined, and the analysis of the regional disparities that emerge in the course of the ever changing spatial composition of the country offer a wealth of subjects for research. At the same time, the interrelationships between the multiple demographic and economic aspects at play make their examination a challenging and fascinating task, especially if one dares to go further into the deeper regional levels where the variety of scenarios and outcomes grows remarkably.

3. Motivation and objectives of this dissertation

The analytical background discussed in the previous sections establishes a solid setting for the development of several research lines. On this basis, the primary aim of this thesis is to shed some light on the processes of population and employment growth taking place in Spain in the last two decades (1991-2001 and 2001-2011).

According to the scheme followed in the description of the Spanish case, this dissertation will deal with the topic of local population in the first part, and then it will engage in the study of employment in the second one. In the development of the argumentation, each part will start with an analysis of the concerned issue taking into account the determinants more commonly used in the related literature, and paying particular attention to the local features that might be fostering concentration on any of these dimensions.

Elaborating further on the contents of this work, the first part of the thesis focuses on population growth, with an emphasis on the influence of local factors and neighborhood effects along the decades 1991-2001 and 2001-2011. The objective here is to identify local responses to socio-economic, geographical and environmental features, confronting global and local estimation methodologies to provide some evidence on the existing variability in the responses to different determinants. In this sense, the application of geographically weighted regressions (GWR) allows for the appreciation of differentiated local effects across the Spanish local labor markets when there exists significant variability in the impact of a certain factor. The implications of this exercise range from the academic interest on evaluating the degree and characteristics of spatial heterogeneity, to the importance of identifying particular aspects susceptible to regional policy intervention. As a complementary analysis, this section also includes an analogous study on the growth of ageing population in Spanish municipalities between 2001 and 2014. The aim here is to learn about the acute ageing process in terms of the convergence/divergence in the share of older population, and how municipal traits affect that process, with special attention to the effects of the relative location and the position in the urban hierarchy.

Employment growth is dealt with in the second part, which seeks to explain the determinants behind the employment concentration in Spain and the relationship between the characteristics of the local labor markets, the characteristics of their neighbors, and their relative performance in terms of local employment. This part presents a spatial econometric analysis by means of a Durbin model that accounts for the spatial autocorrelation among regions, while showing the proximity links between them with respect to the explanatory factors included. The spatial structure is also reflected in reference to the complete urban hierarchy through the concept of incremental distances, a set of measures that correct the distance to a certain urban tier by the proximity of other locations. This empirical approach provides evidence on functioning of the Spanish labor market from a regional perspective during 1991-2001 and 2001-2011, and may help as a guide for labor policy instrumentation. The last chapter included is this section attempts to contribute to the very scarce literature on local wages in Spain, especially on the topic of local differential wages, by assessing the evolution of the urban wage premium in the country and the impact that the Great Recession in 2008 has on this notion. Some allusions to the repercussion of the worldwide economic crisis in the Spanish economy were shown in Figure 4, regarding the effect that it had on the employment rate. Correspondingly, data on aggregated wages indicates that their growth rate dropped considerably after the outbreak of the crisis, exhibiting even negative numbers in some sectors. The aim of this chapter is to analyze if the drop in wages translates into a change in the existing gap between the largest urban municipalities and the other areas by comparing the pre-crisis

(2004-2008) and during-crisis (2009-2014) scenario, and which of the components of this urban wage premium (sorting, agglomeration, price, learning or coordination effect) are more sensitive to the shock. This study goes deeper into the local differences and regional interrelationships, giving some hints about the relative resilience of municipalities.

The definition of "local" in the opening chapters of each part will be represented by local labor markets, given the interest on understanding how population and employment distribute among Spanish regions from a broad but functional perspective. In this sense, local labor markets enclose connected residential and labor behaviors within the same area, making them an acceptable spatial definition for concentration. Each part also includes an additional chapter focusing on a specific subject related to population or employment. As mentioned earlier, the closing chapter of the first part considers the increase in the share of ageing population, while in the second part the last chapter examines the spatial wage gap and its reaction to the recent crisis. In both cases, the focus goes to the most spatially disaggregated level available, in this case municipalities, in order to unveil the precise dynamics behind these phenomena, in contrast to the general angle adopted in the respective preceding chapters.

Although several studies have addressed the issues discussed here, an unfilled gap remains regarding the analysis of these subjects from a local perspective while considering the effects of the spatial structure of the whole country, being the main reason the difficulty to find suitable data at a disaggregated level. The release of the recent censuses and the appearance of new data sources enable for the implementation of research proposals as the one presented, however data limitation still poses a serious handicap for regional exploration. This dissertation is an endeavor to provide an overview of the Spanish economic geography and to remark its relevance, with the aspiration that in the near future the information and tools at hand allow for further, and more importantly, for better research.

PART I. SPATIAL POPULATION DYNAMICS

CHAPTER I. HETEROGENEITY IN THE DETERMINANTS OF POPULATION GROWTH AT THE LOCAL LEVEL

1. Introduction

The determinants of population growth or decline comprise one of the most relevant issues in the Social Sciences. These determinants have been studied from different perspectives and with various methodologies depending on the discipline, for example, regional economists are especially interested in linking the socio-economic forces at work with the population dynamics. Each place produces two sets of opposing forces: centripetal and centrifugal. The set of centrifugal forces pushes economic activity – and population – out from the more congested areas to the less dense ones. Decreasing returns, assumed in all neoclassical economic growth models, are the most relevant centrifugal force. By contrast, centripetal forces act in the opposite direction, making highly-developed and densely populated areas more attractive for economic and population growth, which generates processes of concentration. Agglomeration economies connect the attractiveness of a place with its size in terms of population and business density, and represents the most important centripetal force. In reality, both forces operate continuously and interact with many other environmental, social, political and geographical factors; making it very difficult to predict the final result.

Formal modeling of local population growth has usually tended to focus on identifying 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). When the results are not consistent with the predicted pattern, some studies 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). Moreover, as Glaeser et al. (2014) highlighted, population growth laws which are reliable for long-term dynamics may not be hold in some moments or for some countries-areas as the balance between centrifugal and centripetal forces can change and give rise to different specific behaviors.

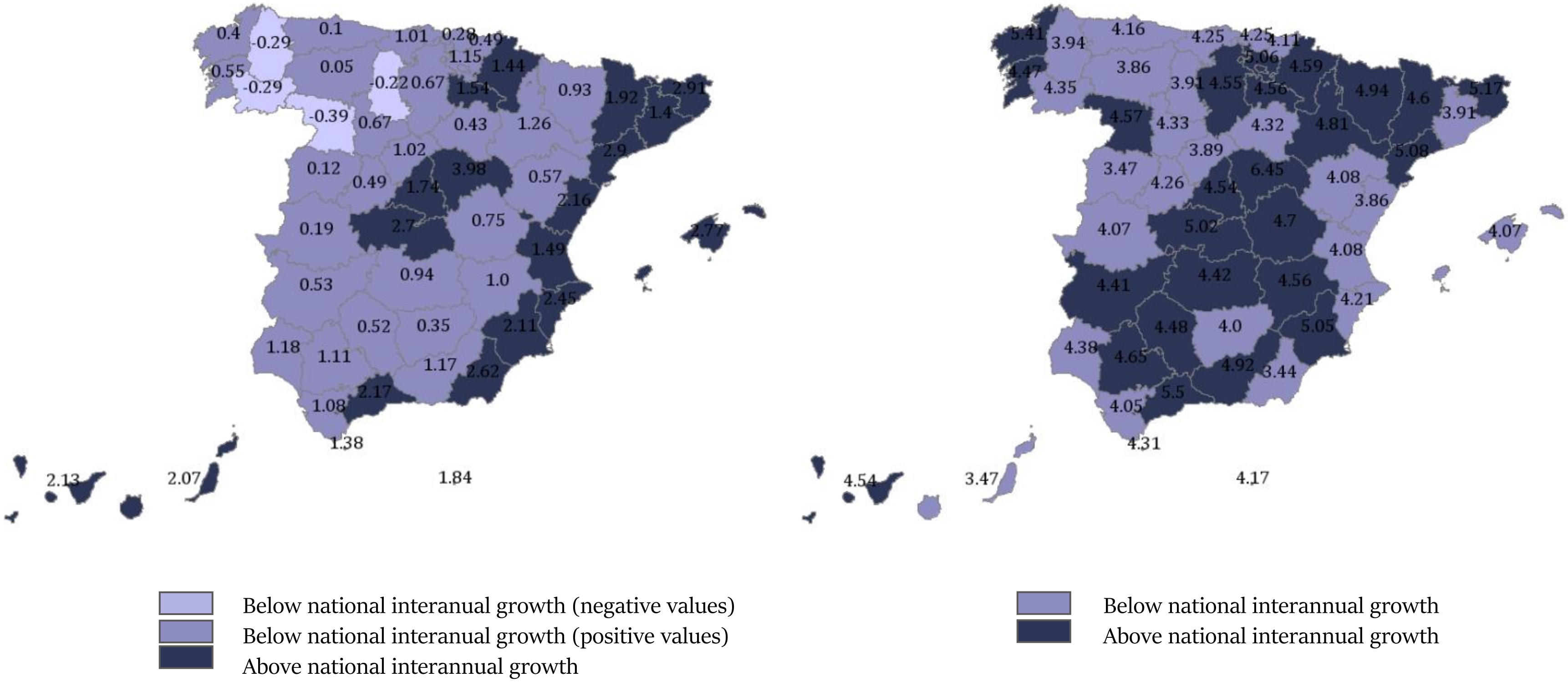
The objective in this chapter is to perform an empirical analysis of the local determinants of population growth in Spain assuming the possibility that different economic forces and environmental, social, political and geographical factors may have different effects across both space and time. In the last two decades and until the onset of the current economic crisis, Spain experienced strong economic growth accompanied by strong population growth (going from 38,872,268 inhabitants in 1991 to 40,847,371 in 2001 and 46,815,916 in 2011). However, population growth, economic growth and many other factors (employment, immigration or education among others) were not homogenously distributed across space. While in the regions of the North both economic and population growth were below the national average (Figure 5), the Madrid province and the Mediterranean regions along with the Ebro corridor (from the province of Barcelona to the Bay of Biscay) concentrated economic and also population growth.

In order to allow for spatial heterogeneity, the Geographical Weighted Regressions (GWR) approach is applied to an empirical population growth model for the last two decades: 1991-2001 and 2001-2011. Furthermore, given the emphasis on the importance of working with a sound definition of the spatial unit of analysis, in this chapter spatial units that guarantee homogeneity and reliability across the territory, namely Local Labor Markets, are chosen.

This chapter is structured as follows. In the next Section the appropriateness of the spatial unit of analysis chosen is justified, not only to set the boundaries of cities but also to ensure a homogeneous partition of the Spanish territory. Once the spatial unit has been explained, the population growth factors considered in the model are specified in Section 2. Section 3 contains an explanation of the basics of the GWR, a methodology that captures spatial variations in the regression coefficients. The results obtained on the local population growth determinants applying GWR are then described and compared to the equivalent Ordinary Least Squares estimates in Section 4. Greater emphasis is placed on those cases where factors seem to have a different or contradictory effect on local population growth depending on the methodology implemented. This article ends with Section 5 summarizing the main results and conclusions, and outlining some important economic policy implications derived from the analysis.

Figure 5. Interannual population and economic growth at provincial level between 2001 and 2011 (%)

(a) Population growth **(b) Economic growth**



Source: Own elaboration with data from the National Geographical Institute (IGN) and the censuses from the National Statistics Institute (INE).

2. Local Population growth: determinants and spatial unit

The modeling of population growth determinants basically depends on the objective and perspective of the discipline[1]. Thus, human ecologists use a holistic human-society-environment framework (Poston and Frisbie, 2006). Geographers are more interested in spatial variations of population distribution, density or composition (Bailey, 2005), while urban planners focus on land use patterns (Cervero, 2003) and environmental planners place their attention on population change and its consequent impacts (Cowen and Jensen, 1998).

Economists are more focused on identifying the connections between population growth and other socio-economic variables. Most of the empirical analyses of population growth determinants consider a broad set of features such as the initial population and population density, sectoral employment shares, economic growth or some human capital indicators and evaluates their impact on the evolution of population. For instance, Beeson et al. (2001) study the long-run patterns of population among the 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 factors such as public expenditure on education, housing-price inflation, personal income and the quality of life are relevant. In the completely different scenario of China, Zhang and Song (2003) detect a link between local population evolution and the massive rural-urban migration, the income gap between both settings and their geographical distance. 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 of urban concentration led by regional capital-cities and by regions with certain locational features (coast or valley). Also for Spain, Conde-Ruiz et al. (2008) found that international immigration had a positive impact on population growth between 2000 and 2006, although this effect is far from homogeneous.

[1] See Chi and Ventura (2011) for an extensive revision of the different perspectives of population growth determinants.

All these papers obtain similar general conclusions: (i) the most populated places in the past are usually the ones that grow most intensely in the present; (ii) changes in the population distributions normally happen in the short-run, with the long-run tendency reverting to the persistence of greater growth in the most populated areas; (iii) the socioeconomic structure and the geographical factors play different roles in different places, or even different roles in the same place in different periods; and (iv) the role of the education level is normally positive but its relevance changes from one place to other and it plays a more salient role in highly-populated areas.

Based on the fundamental insights of New Economic Geography which highlights the importance of agglomeration and location on many socioeconomic phenomena, the empirical model to explain population growth suggested in this chapter is estimated by applying a GWR technique, which allows for spatial variations in the regression coefficients. The model is estimated both for the decades 1991-2001 and 2001-2011, thereby allowing for the detection of variations in the determinants over time. The first step, however, is to justify and define the functional regions called Local Labor Markets (LLMs hereinafter) as the most suitable spatial unit for this local analysis. In the second part of this Section, the choice of the variables used in the empirical model will be detailed.

2.1. Why LLMs as the reference spatial unit

Administratively divided into 17 Autonomous Communities and 50 provinces (NUTS II and III regions respectively, according to the European nomenclature), in Spain there is basically a unique source of information that collects and provides data at local level (municipality): the Population and Housing Census database. This is the only reliable 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 on local population and employment growth.

Despite its generalized use in studies at the local level, municipalities are not the best option to study population growth from a broad perspective as they fail to comprise economically and socially integrated areas (see Rubiera and Viñuela, 2012). The more notable cases, though not the only ones, are the metropolitan areas of Madrid and Barcelona, which comprise 151 and 51 municipalities respectively.

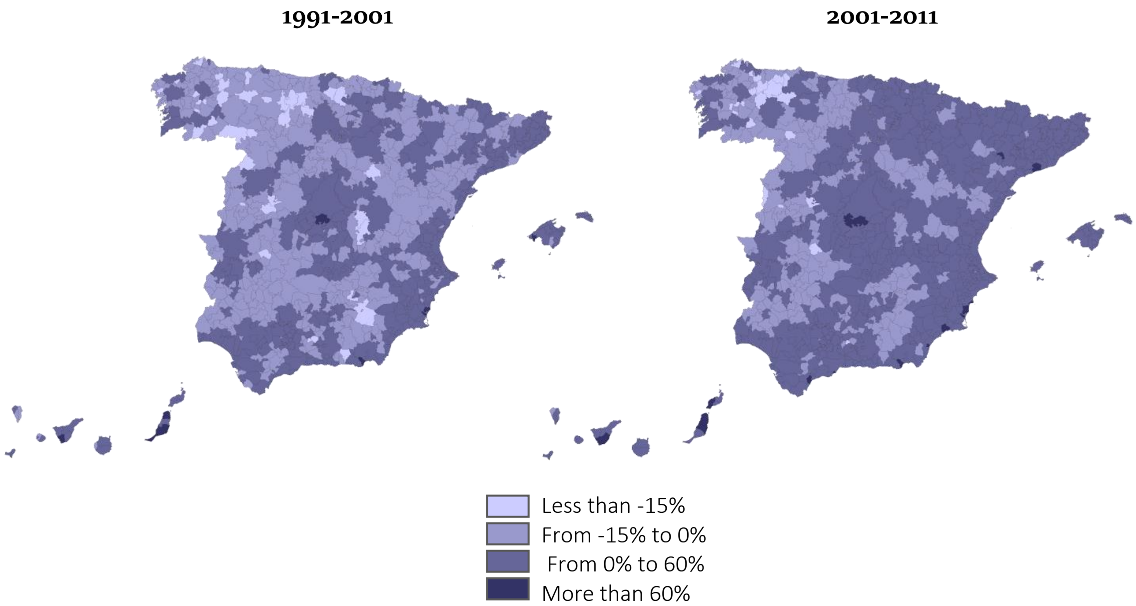
Taking into account that labor plays a basic role in the lives of individuals and conditions their territorial behavior with regard to the choice of the municipality of residence, and therefore population growth, a LLM or travel-to-work area describes a space where the population develops most of its economic and social relationships. The LLM is a place where the common interest of the local population can be identified as a whole. It is therefore the appropriate level for implementing policies at local or regional level (Parr, 2008), something theoretically simple but that, for political reasons, meets with lots of 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 individuals living in the area, most of the observed employed population growth is in fact employment growth within the area (Sforzi, 2012). The regionalization procedure to construct the LLMs was originally developed by Coombes et al. (1986) and applied by, among others, the Department of Employment in Great Britain, the Italian National Institute for Statistics (ISTAT-IRPET, 1989) in Italy and in Spain by Boix and Galletto (2006). A complete explanation of the method used to define these areas can be found in Sforzi and Lorenzini (2002) and Sforzi (2012).

Using the 8,119 Spanish municipalities as building blocks, the five-stage algorithm starts by highlighting candidates to be the center of an LLM. Gradually, other municipalities are added if two principles are fulfilled: a maximum of 25 % of its residents commute to jobs outside the area and a minimum of 75 % of employment corresponds to residents from the area. Therefore, a LLM comprises municipalities among which the vast majority of labor-related movements take place, and therefore where the income generated coincides with the income available, defining a self-contained, well-defined and compact functional area with a high internal homogeneity in terms of labor and income (Rubiera and Viñuela, 2012).

Applied by Boix and Trullén (2010), this methodology identifies 806 LLMs in Spain. Figure 6 shows the population growth in Spanish LLMs over the last two decades. Between 1991 and 2001, population growth was highly concentrated in and around the largest cities, especially those in the Mediterranean coast. The following decade was characterized by a major increase in the foreign population, and high rates of population growth could also be found in the small and medium-sized cities and across most of the eastern and southern parts of the country.

Figure 6. Spanish LLM's population growth rate in 1991-2001 and 2001-2011



Source: Own elaboration with data from the National Geographical Institute (IGN) and the censuses from the National Statistics Institute (INE).

After many years of population concentration, the end of the 20th century was characterized by a contained de-concentration process where large cities started to suffer the effects of agglomeration diseconomies, resulting in a process of counter-urbanization and reversing the divergence trend between the urban and rural areas. This shift in the

Spanish urban structure has been pointed out in several studies using municipal data, showing that individuals have started to consider medium and small-sized cities as a place of residence as they can offer, among other qualities, a well-preserved environment, low levels of crime and a similar range of goods and services to that of the metropolises (Lanaspa et al., 2003; Le Gallo and Chasco, 2008). However, these studies fail to explain what the determinants of this spatial and temporal shift in local population growth are.

2.2. An empirical model with population growth determinants

After defining the spatial unit of analysis, an empirical model to study population growth (Y) across space is proposed. In this chapter the suggestion of Chi and Ventura (2011) is followed. These authors emphasize that it is necessary to consider a wide framework with diverse socioeconomic and geographical factors. This idea was highlighted by Glaeser et al. (2014) who underlined the heterogeneous nature of population behavior both across space and along time, which in turn makes it necessary to consider many different variables that play different roles depending on the place, the moment or the context considered. In accordance with this, a first approximation to explain population growth through a set of possible determinants or explanatory variables (X), for n spatial units (803 LLMs in this case) through a regression model like the following one:

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

The election of the variables is based on previous literature using a similar approach, including Coffey and Polèse (1988), Cebula (2002), Partridge and Rickman (2003), Polèse and Shearmur (2004), Shearmur and Polèse (2005), Chi and Ventura (2011) and Glaeser et al. (2014) among others. The selection of variables is determined by the data availability. As with several other countries, information at the LLM level of disaggregation in Spain is only available through the Population and Housing Census. This Census is carried out every ten years by the National Statistics Institute (INE, 1992, 2007 and 2012), and two decades can be analyzed in detail for the Spanish case: the 1990s (specifically from 1991 to 2001) and the 2000s (from 2001 to 2011).

The first factor includes is the logarithm of the population size (*PopSize*) at the beginning of the period considered, 1991 and 2001 respectively. This variable reflects the rank of the LLM in the urban hierarchy and the presence of agglomeration effects.

In a Regional Economics framework, crucial explanatory factors apart from population size relate to location. In this study, location is going to be measured as *distance to size* (*DistMA*), defined as the linear distance to the nearest metropolitan area, namely a LLM with at least 500,000 inhabitants. Defined as such, this variable contains information about the interrelationships between cities, the power to attract or repel population that a certain area has, and its accessibility to goods, services, amenities and labor markets (Coffey and Polèse, 1988; Polèse and Shearmur, 2004). Thus, the remotest Spanish LLM is located in the North-West at 614 km from the nearest metropolitan area (Table 1). The geo-referenced information of the LLMs can be obtained from the Spanish National Geographical Institute (IGN).

The effect of size and distance should be analyzed hand in hand, as they can exert a joint ambiguous effects on population growth: rural and peripheral areas where populations do not surpass a certain critical mass might be experiencing a depopulation process while rural (small) but central areas might be ideal places of residence for commuters. 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 in Spanish territory. Moreover, these variables could be exerting a positive effect on population growth in some areas but the opposite effect in others.

To complete the analysis, other factors that can be classified as socio-economic, geo-structural and climatic variables, and that will act as controls, are included in the local growth model. It is worth noting that the only kinds of factors that can be affected by through policy instrumentation are the socio-economic variables. Table 1 provides the common descriptive statistics of the variables considered.

Among the socio-economic variables considered is the educational level (*Ed*), measured as the percentage of the potential workforce (population aged 16 to 64) which holds a university degree and which represents an indicator of the initial human capital endowment. This factor is related to several dimensions (population mobility, economic structure, productivity, income, etc.) that may affect population change and provide with interesting insights about its evolution. In recent decades, tertiary education attainment has increased overall either considering the mean (from 5% to 10%) or the minimum or maximum values. Another variable considered is 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, the casualty direction followed in this chapter can be 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 included to capture 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 inserted in the model. The former captures the influence of the location per se 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. There are also 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, 2011).

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

[2] As the Greenwich meridian, the reference for measuring the Longitude, divides the country into two halves, an adjustment was made in the scale. Longitude zero corresponds to the easternmost LLM (Mahon) and Latitude zero corresponds to the northernmost LLM (Xove).

(*Cap*). This variable underlines the importance of being appointed as the administrative center, thereby concentrating a large part of the public sector jobs and offering a larger variety of public services.

Table 1. Descriptive statistics

Variable	Min.	Max.	Mean	Std. Dev.
PopSize91 (population in 1991)	1,909	4,864,827	48,252.98	225,686.40
LnPopSize91 (logarithm of population in 1991)	7.55	15.40	9.64	1.18
PopSize01 (population in 2001)	1,951	5,316,487	50,690.66	238,724.20
LnPopSize01 (logarithm of population in 2001)	7.58	15.49	9.65	1.21
DistMA (distance to nearest metropolitan area)	0	614.35	155.90	119.55
Ed91 (share of work force with a university degree in 1991)	0.01	0.15	0.05	0.02
Ed01 (share of work force with a university degree in 2001)	0.03	0.26	0.10	0.04
PopGrowth9101 (population growth rate from 1991 to 2001)	-0.31	1.38	0.02	0.16
PopGrowth0111 (population growth rate from 1991 to 2001)	-0.31	1.27	0.10	0.19
Emp91 (employment rate in 1991)	0.23	0.95	0.49	0.09
Emp01 (employment rate in 2001)	0.36	0.91	0.58	0.08
LQagr91 (location quotient in agriculture in 1991)	7.23	773.77	262.76	163.15
LQagr01 (location quotient in agriculture in 2001)	7.51	1,124.94	284.99	206.85
LQind91 (location quotient in industrial activities in 1991)	12.41	284.81	83.12	53.26
LQind01 (location quotient in industrial activities in 2001)	8.72	345.39	103.70	65.24
S91 (Specialization index 1991)	1.30	7.99	4.22	1.06
S01 (Specialization index 2001)	0.99	8.41	4.03	1.09
Lo (longitude coordinate)	0.00	22.27	7.75	3.63
La (latitude coordinate)	0.00	15.93	4.01	2.80
DistCoast (distance to coast)	0	422	97.71	90.62
Rain (average annual rainfall from 1987 to 2007)	1.00	16.41	5.24	2.72
$Tmin_{jan}$ (minimum average temperature in January)	-6.77	12.50	1.20	3.27
$Tmax_{jul}$ (maximun average temperature in July)	20.50	36.70	30.67	3.60
Cap (capital city dummy)	0	1	0.06	0.25
Cases	803			

3. Capturing spatial heterogeneity with Geographical Weighted Regressions

As stated earlier, it is important to acknowledge that the estimated effects of a variable can vary greatly across territories 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 population growth.

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 environment. The conclusions regarding, for instance, population or economic growth determinants derived from global estimations can mask significant local variation. Thus, a standard overall estimate may point to a certain effect of one factor whereas this 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 problematic when the average impact is close to zero, as it might be deemed to be 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-effect model, where dummy variables are introduced to capture site-specific characteristics (Anselin, 1998; 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:

$$Y_i = \beta_{i0} + \beta_{i1}x_{i1} + \beta_{i2}x_{i2} + \cdots + \beta_{ik}x_{ik} + \varepsilon_i$$
$$\varepsilon_i \sim N(0, \sigma^2);\ i = 1, 2, \dots, n \quad [2]$$

This yields a separate set of parameters for each of the *n* observations, calculated through the following equation:

$$\hat{\beta}_i = (X'W_iX)^{-1}\, X'W_iY;\ i = 1, 2, \dots, n \quad [3]$$

Our dependent variable (Y_i) is the *i*th LLM's population growth in different periods of time, and the explanatory variables (X_i) are those presented in the previous section.

For each observation (i.e., 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 GWR depend on the linear distance between observations and represent the adjacency effects for neighboring locations within the specified bandwidth (Cleveland and Devlin, 1988; McMillen, 1996; and Brunsdon et al., 1996, 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:

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

Following the assumption that more proximate locations are more alike, the weights should decay with distance. Many weighting schemes fulfilling such a requirement could be used (dichotomous, bi-square or tri-square decay function, etc.), but as Fotheringham et al. (2002) point out, “the results of GWR are relatively insensitive to the choice of weighting function”. When choosing a Gaussian kernel, the weighting function is specified as:

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

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

The GWR approach has several advantages over standard methods, but also its flaws. One advantage is that since each area has its own constant term, it may 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. However, as the GWR approach produces a considerable amount of regressions, considering a large range of estimates allows us to "average" them, and thereby better determine 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 issue of 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, as 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 though further examination should be carried out and caution is recommended, the severity of the mentioned problems decrease with the size of the sample[3].

4. Results: global and local effects

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 and its spatial patterns.

The OLS estimation provides fourteen global parameters (plus one intercept) that reflect the effect of each explanatory variable on the complete sample. The result of this estimation may veil the effect of each factor on each region, which is particularly troublesome if heterogeneity exists, and the associated variability is large (Ali et al., 2007; and Partridge et al., 2008). The GWR procedure provides 15 x 803 parameters, one for each factor (plus the intercept) and LLM. The focus will be set on discussing those factors where the GWR estimates add useful and significant information on local population growth determinants. After that, those factors will be represented, showing a spatially-differentiated comparison in order to deduce spatial or regional policy implications for Spain.

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

Estimations for both models are shown in Table 2 (1991-2001) and Table 3 (2001-2011). 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[4].

Checking the differences in the OLS estimations between the two decades analyzed, temporal heterogeneity is clearly found in some determinants. The impacts of the size and distance to size explanatory factors is particularly worthy of note: being a large or medium-small sized city, and/or being located close to a large metropolitan area did not act as a population growth attraction factor in the 1990s, whereas in the first decade of the 21st century both factors are significant. The sign of the size parameter is positive, which might be revealing the end of the "contained de-concentration" process in Spain – i.e. the expulsion of population from the larger urban areas to smaller ones – as larger cities recovered positions in the last decade. The capital city condition (*Cap*) is only significant (and negative) in the second decade. The result for the variable maximum temperature in July ($Tmax_{jul}$) is anecdotic, being significant in the 1990s but not in the 2000s.

Although the rest of the independent variables remain significant for the whole time interval, it is interesting to note that the shifts for the variables initial population and distance to the metropolises reveal a noteworthy change in population's behavior in a relatively short time-span.

When estimating the model using OLS, the determinants proposed can explain an important part of Spanish population growth 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.

[4] In order to make the results from different decades comparable, the average of the optimal bandwidths for each period is used in the GWR estimations.

Table 2. Estimation Results for 1991-2001

Variable	Global (OLS)	Min.	1st Qu.	Median	3rd Qu.	Max.	F3 test(a)
Intercept	0.102	-10.270	0.074	0.180	0.242	0.357	+++
LnPop91(log of pop. size 1991)	0.0009	-0.222	-0.007	0.002	0.010	0.023	+++
DistMA (Distance to nearest MA)	-0.004	-0.0408	-0.010	-0.005	-0.003	0.101	
Ed91 (Educational level 1991)	-0.615 **	-1.110	-0.985	-0.799	-0.500	7.358	+++
Emp91 (Employment rate 1991)	0.558 ***	-3.019	0.423	0.549	0.669	0.809	+++
$LQ_{ag}91$ (Location quotient agr. 1991)	-0.0006 ***	-0.002	-0.0006	-0.0004	-0.0003	-0.00001	+++
$LQ_{ind}91$ (Location quotient ind. 1991)	-0.001 ***	-0.010	-0.001	-0.0008	-0.0005	0.0003	
S91 (Specialization index 1991)	0.020 ***	-0.012	-0.0009	0.007	0.010	0.353	+++
Lo (Longitude coordinate)	-0.006 **	-0.415	-0.011	-0.008	-0.0003	0.011	+++
La (Latitude coordinate)	0.013 ***	0.014	0.017	0.019	0.021	0.944	+++
DistCoast (Distance to coast)	0.004	-0.042	0.003	0.012	0.016	1.675	++
Rain (Avg. annual rainfall 1987 to 2007)	-0.013 ***	-0.024	-0.017	-0.013	-0.005	0.091	
$Tmin_{jan}$ (Min. temperature January)	0.015 ***	-0.162	0.011	0.015	0.017	0.019	+++
$Tmax_{jul}$ (Max. temperature July)	-0.004 **	-0.011	-0.008	-0.007	-0.005	0.291	+++
Cap (Capital city dummy)	-0.003	-0.085	-0.013	0.017	0.038	0.466	+++
Adjusted R^2 OLS	0.5196		F-statistic	62.96 ***			
Adjusted R^2 GWR	0.6674						
F1 test(b)	0.7500 +++						
F2 test(c)	4.8901 +++						

- Dependent variable: population growth rate from 1991 to 2001.

- */**/*** 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 3. Estimation Results for 2001-2011

Variable	Global (OLS)		Min.	1st Qu.	Median	3rd Qu.	Max.	F3 test(a)
Intercept	-0.102		-0.517	-0.099	0.113	0.409	1.448	
LnPop01 (log of pop. size 2001)	0.027	***	-0.163	0.019	0.031	0.037	0.047	+++
DistMA (Distance to nearest MA)	-0.014	**	-0.033	-0.029	-0.025	-0.020	0.654	+++
Ed01 (Educational level 2001)	-0.794	***	-1.456	-1.281	-1.147	-0.965	5.145	+++
Emp01 (Employment rate 2001)	0.772	***	0.114	0.703	0.721	0.744	2.579	
$LQ_{ag}01$ (Location quotient agr. 2001)	-0.0004	***	-0.001	-0.0005	-0.0005	-0.0004	0.0005	
$LQ_{ind}01$ (Location quotient ind. 2001)	-0.001	***	-0.031	-0.001	-0.001	-0.0009	-0.0004	+++
S01 (Specialization index 2001)	0.020	***	-0.007	0.010	0.013	0.020	0.177	++
Lo (Longitude coordinate)	-0.012	***	-0.596	-0.015	-0.014	-0.009	-0.003	+++
La (Latitude coordinate)	0.007	*	-0.007	0.008	0.011	0.011	0.857	+++
DistCoast (Distance to coast)	0.003		-0.175	0.005	0.017	0.024	3.524	+++
Rain (Avg. annual rainfall 1987 to 2007)	-0.016	***	-0.041	-0.027	-0.019	-0.010	0.353	+++
$Tmin_{jan}$ (Min. temperature January)	0.019	***	-0.125	0.018	0.020	0.022	0.0251	+++
$Tmax_{jul}$ (Max. temperature July)	-0.004		-0.083	-0.012	-0.009	-0.006	0.020	
Cap (Capital city dummy)	-0.045	*	-0.086	-0.041	-0.025	0.010	0.577	+++
Adjusted R^2 OLS	0.5556			F-statistic	72.63 ***			
Adjusted R^2 GWR	0.6562							
F1 test(b)	0.8376	+++						
F2 test(c)	3.5454	+++						

- Dependent variable: population growth rate from 2001 to 2011.

- */**/*** 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.

With this aim, the goodness of fit between the OLS and GWR estimations is compared, and then the significance of the variation in the parameters obtained through GWR is analyzed. 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 (0.75 for the first decade and 0.83 for the second) supports a better fit for the GWR estimation. The F2 test is based on the SSR improvement of GWR over OLS, measured by the difference between their residuals sums of squares. A large value of this test (4.99 for the first decade and 3.54 in the second) implies 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 2 and 3, need to be interpreted considering 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, that is, the factor under study does not have a spatially-differentiated effect on population growth. The second possibility is that the variation of the coefficients obtained under GWR is significant but that the global OLS coefficient is not. 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 the OLS estimates, though significant, fail 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.

Representative of the first scenario are the industrial location quotient (LQ_{ind}) or the average rainfall (*Rain*) in the period 1991-2001, and the employment rate (*Emp*) or the agricultural location quotient (LQ_{agr}) for 2001-2011. The significance of the OLS estimates suggests that these factors are affecting population growth in a similar way regardless of the characteristics of the LLM under consideration. In other words, in all the Spanish territory, the more specialized the area is in the agricultural or industrial sector, the less population growth. The same type of spatially-homogeneous effect applies to the favorable conditions of the area in terms of rainfall (ceteris paribus the rest of variables, population grows more in drier areas)

The factors mentioned have an even impact on the Spanish territory, at least for the decade observed, pointing again to the differences depending on the timeframe chosen. In this regard, attention should be drawn to the particular behavior of the employment rate variable, which proxies the differentials in labor opportunities between areas. Interpreting the OLS estimates for this variable, which are significant for both decades and have the expected positive sign, it can be concluded that for the Spanish territory as a whole one of the main factors driving population growth is the labor market conditions (estimate of 0.558 for the first decade and 0.772 for the second). As anticipated, favorable labor conditions promote population growth everywhere. However, in Spain the population growth experienced in the first decade in some areas did not follow the above economic rationale (i.e., the better the conditions of the labor market, the more the population growth), as in that decade the F3 test shows the existence of spatial non-stationarity and the estimates reveal that in some areas favorable labor conditions are negatively related to population growth[5].

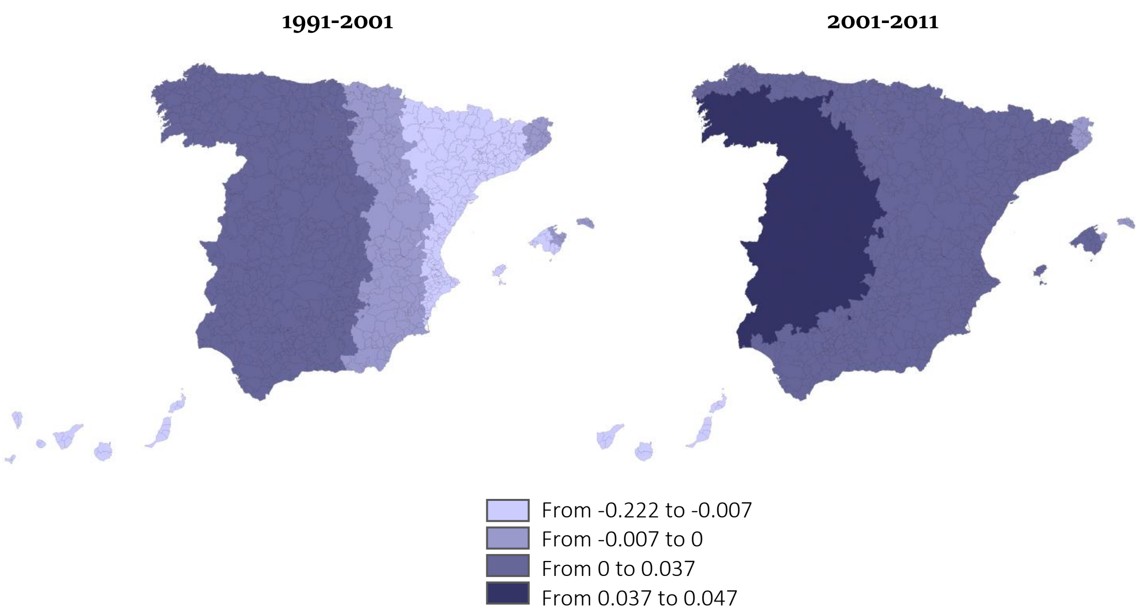
The second scenario (significance of the variation of the GWR coefficients but no significance of the global OLS coefficient) applies in the case of the variables initial population size (*Pop*) in 1991, the distance to the coast (*DistCoast*) in both decades, and the capital city dummy (*Cap*) in the first decade. For the first decade, the OLS estimates for the initial population of the LLM and the distance to the nearest metropolitan area show the expected signs (positive for size and negative for distance), although none of them are significant. Surprisingly, according to these global results and after controlling for the rest

[5] A more detailed discussion on the results for the employment rate factor will be offered after discussing the third possible scenario.

of the determinants, neither size, being a capital city or proximity to the coast had any effect on population growth during the 1990s. However, the Spanish urban areas, and especially those located close to large metropolitan areas, were experiencing higher population growth rates during the 1990s and also 2000s. Hence, this is one of the cases where there is a compensation effect when estimating through OLS, as the existing regional variability (revealed by the significance of the F3 test) leads to an average general effect proximate to zero. If these same factors are analyzed allowing for spatial heterogeneity, the GWR estimators shows that in certain LLMs both size (*Pop*), distance to the coast (*DistCoast*) and also distance to the nearest metropolitan area (*DistMA*) can be affecting population growth in the opposite direction to that expected. In other words, for some LLMs being far away from the large metropolitan areas can have a positive influence on its own population growth, and some LLMs can be positively influenced by their distance to the coast: Similarly, some urban LLMs may be immersed in a process of agglomeration diseconomies (the larger the initial size, the less the population growth) while others are benefitting from agglomeration economies (the larger the initial size, the more population growth). Representing the GWR estimates on a map reveals that the agglomeration economies/diseconomies processes need to be analyzed separately, as they are simultaneously taking place in the same geographical area (Spain) for the two decades considered.

As shown in Figure 7, agglomeration diseconomies processes are prevalent in the first decade in urban areas located in the North-East (and in both decades in the Canary Islands). This effect is generally associated with negative externalities such as congestion, environmental degradation or higher housing prices in large cities. However, using OLS methodologies the agglomeration economies experienced at the start of the 21st century (positive parameter) in the rest of the Spanish geography are offset by the agglomeration diseconomies (negative parameters), and as a result the parameter for the initial size (*Pop91*) is not significant. The next decade (2001-2011) was characterized by a broad agglomeration economies process that covered almost the entire Spanish territory (except the Canary Islands). This effect might be associated with the recent slight weakening of population sprawl (Glaeser and Shapiro, 2003). Something similar occurs for LLMs that contain a capital city: those located at the West have experienced a positive population growth attributable to this condition, while several Mediterranean ones suffered population decay. In the decade 2001-2011, the positive effect is reduced to the north-east of the country.

Figure 7. GWR estimated coefficients for the initial population size factor

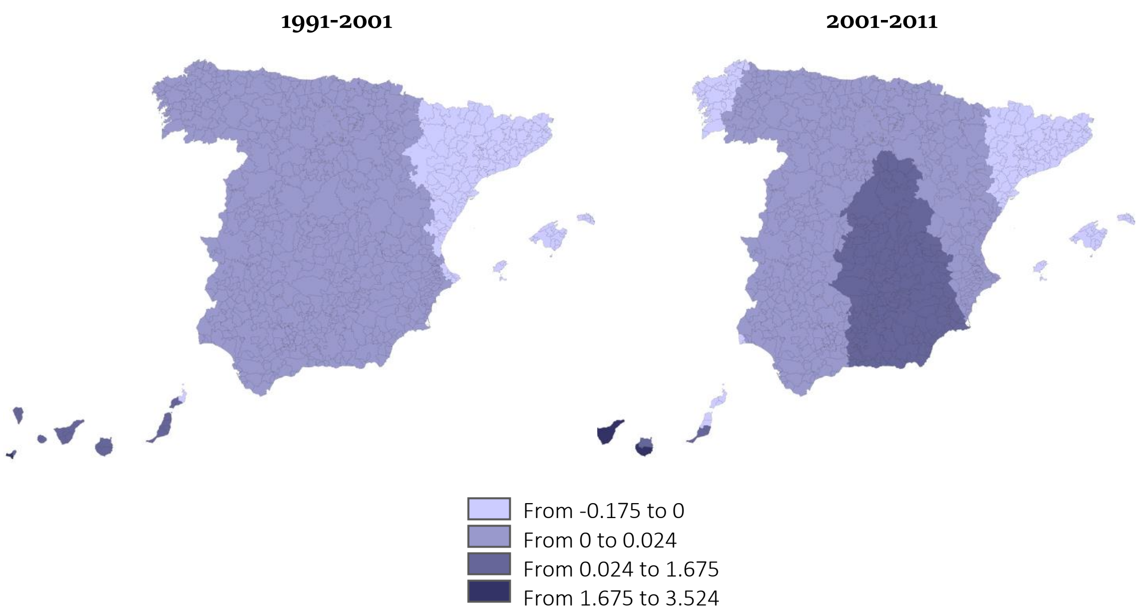


Source: Own elaboration with data from the National Geographical Institute (IGN) and the censuses from the National Statistics Institute (INE).

As for the distance to the nearest Metropolitan area (*DistMA*), the relative location of the LLM turns out to be irrelevant for population growth in the first decade (non-significant global effect, non-significant variability). However, in the following decade *distance* becomes a relevant factor for explaining population growth. The sign of the OLS parameter is significant and has the expected negative sign. Although the variation of this determinant obtained under GWR is also significant, a closer analysis shows that positive values are only observable in the LLMs of the Canary Islands', an ultra-peripheral region in the EU context.

As shown in Figure 8, being located near the coast favors population growth in the north-east corner (negative influence of longer distance). This effect also applies in some regions of the north-west area in the 2001-2011 decade. The impact is the opposite in inner and southern locations, probably due to the saliency of other factors related to the composition of the Spanish urban system. These factors include the geographical dispersion of the metropolitan areas throughout the whole territory and the fact that the powerful influence of Madrid in the center may compensate the weight of the coast, especially compared to other countries where metropolises are more unevenly distributed, such as the U.S. This may also be driving the general non-significance of the distance to the nearest urban area.

Figure 8. GWR estimated coefficients for the distance to the coast factor

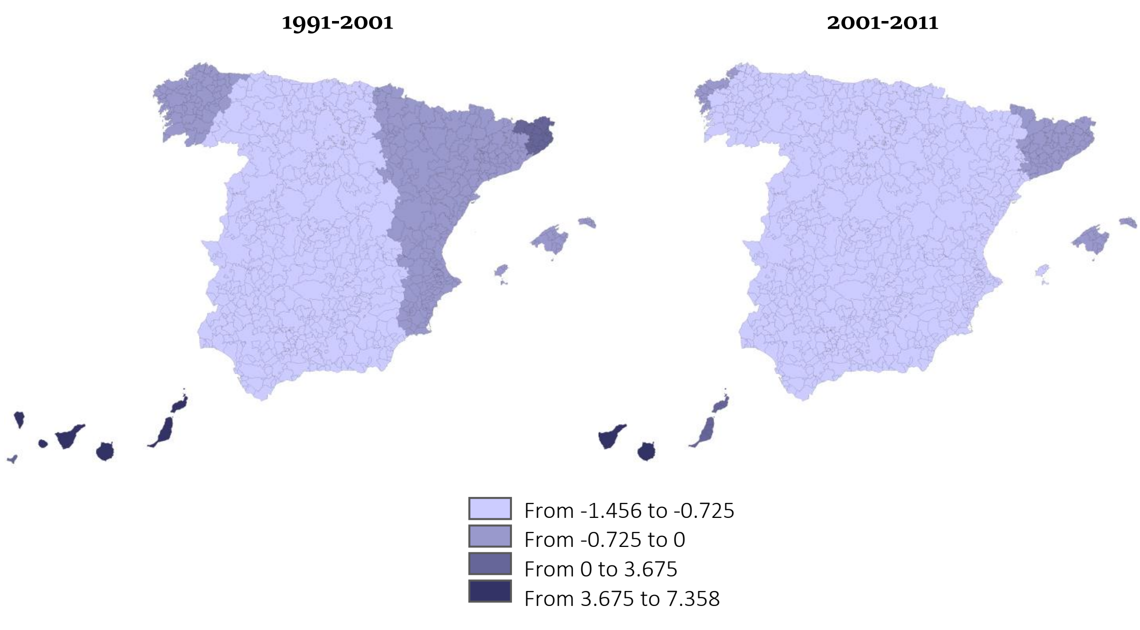


Source: Own elaboration with data from the National Geographical Institute (IGN) and the censuses from the National Statistics Institute (INE).

The third scenario shows cases where the OLS estimates, although significant, have failed to capture the existence of spatial non-stationarity as revealed by the significance of the F3 test. The educational level (*Ed*) in both time intervals and the employment rate (*Emp*) for the first decade fall into these cases.

In both decades the impact of the educational level (percentage of the potential workforce with university studies) on population growth is globally significant and negative*,* i.e., the more potential members of the workforce with a university degree at the beginning of the period, the lower the population growth. In this regard, Pritchett (2001) suggests that the negative and significant effect of higher education may be due to (i) the creation of ill-directed cognitive skills, (ii) stronger growth of the supply than that of the demand of educated workers, or (iii) failure in the creation and transfer of knowledge. However, this initial conclusion requires a deeper examination as GWR estimates demonstrate that the effect of the initial human capital endowment is negative in only some LLMs whereas in others it has a positive effect on population growth (Figure 9).

Figure 9. GWR estimated coefficients for the initial educational level factor



Source: Own elaboration with data from the National Geographical Institute (IGN) and the censuses from the National Statistics Institute (INE).

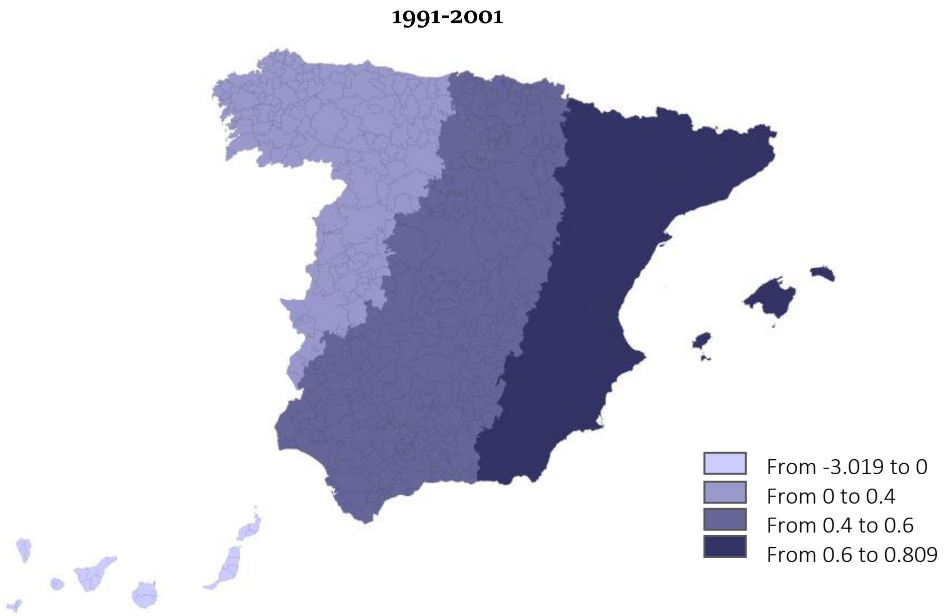
Representing the GWR estimates for the educational level a temporal shift can be seen in the relationship between education and population growth. Apart from its spatially non-homogeneous effect, in the 1990s it is negative in the majority of the LLMs, except for the north-west corner and the Canary Islands. In the 2000s the negative strip gets wider, restricting the positive relationship to the Canary Islands only. This trend shift in the 2000s merits further research, although a possible explanation may be derived from Pritchett (2001) and other authors who underline the Spanish over-qualification problem in certain sectors (Pérez and Serrano, 1998; Serrano, 1999; Montalvo, 2009).

The initial employment rate (*Emp*) is a very relevant factor as it can be affected by both the regional and central governments. Results obtained for this variable both in the OLS and the GWR estimations are, however, complex to interpret. Besides the casualty relationship between employment and population growth mentioned in section 2, spatially-disaggregated results show a different relationship between them depending on the decade under study.

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 the effects of the crisis more intensively than others with the recovery not being evenly distributed across space either in terms of job creation or population growth. OLS estimates yield the expected result, that favorable conditions in terms of employment prompt population growth. However, the GWR estimates and the significance of the F3 for the GWR model also show how some dynamic areas in terms of economic activity (and therefore employment) at the beginning of the 1990s were practically stagnant in terms of population growth or even lost population along the decade, as in those areas favorable economic and labor conditions either do not affect population growth or affect it in a negative way. Those particular areas in the 1990s correspond to the north-western part of the country. On the contrary, having a high employment rate and good labor opportunities constituted an attraction for population growth for inland areas (including Madrid) and especially the Mediterranean coast (Figure 10).

However, in the 2000s the effect of the employment opportunities on population growth was positive for all areas, with no significant variability across space (the F3 test was not significant[6]). That is, regardless of their initial level of economic activity or employment, Spanish LLMs underwent a positive population growth during the 2000s, with no significant local variation. This result is not surprising bearing in mind the sources of population growth in both decades. Although it is not the purpose of this chapter, it is worth pointing out briefly that in the 1990s any population growth experienced at local level was almost exclusively due to internal migration movements and net birth-death adjustments. At the beginning of the decade the share of foreign population living in Spain was 0.90% whereas by 2001 it had grown to 3.81%. The immigrants tended to locate in the islands, the Mediterranean coast and Madrid, working primarily in agriculture, construction, domestic services and tourism-related activities (Aja et al., 2000).

Figure 10. GWR estimated coefficients for the employment rate factor



Source: Own elaboration with data from the National Geographical Institute (IGN) and the censuses from the National Statistics Institute (INE).

---

[6] This is the reason for representing the GWR coefficients for the employment growth factor only in the first decade.

The 21st century was characterized by larger net migratory inflows, and the foreign population increased from 3.81% to 11.22% in only ten years (2001-2011). In this decade the spatial distribution of immigration initially followed the same pattern as that of the previous decade, including all the regions between Madrid and Catalonia, but soon the foreign population was spread over all the Spanish regions. However, the general economic crisis started in 2008 severely hit employment, especially in the sectors that traditionally hires foreign labor (Fuentes and Callejo, 2011). As a result of all these trends, in the second decade spatial non-stationarity cannot be observed for the initial level of employment.

Spatially-differentiated behavior can be observed for other globally significant variables, such as the geo-referenced coordinates (*Lo* and *La*), the winter minimum temperature ($Tmin_{jan}$) or the specialization index (*S*) Thus, a high degree of specialization can work as a deterrent to population growth in one area while in others it acts as a population attraction pole.

5. Summary and main conclusions

The analysis of population allocation and growth usually pursues long-run universal patterns, but factors explaining the population dynamics and growth determinants can change their effect and significance along time and also across space.

This chapter, on the other hand, analyzes the local determinants of population growth in Spain assuming that the different socio-economic, political and geographical factors can affect global population growth in a certain way, but show different effects on local population. Furthermore, the positive effect of one factor for one decade can turn negative for another period of time.

To capture both the spatial and temporal heterogeneity, a GWR approach that allows for the study of the existing spatial non-stationary is adopted for two decades – the 1990s and the 2000s – with very different population dynamics in Spain. Following the literature on population growth and New Economic Geography fundamentals, determinants include the initial population level and the initial employment rate – among other socioeconomic variables –, and several geographic factors such as the distance to the nearest metropolitan area, the average temperatures experienced in the area or the distance to the coast. This exhaustive data was obtained from the Spanish Censuses (1991, 2001 and 2011), released by the National Statistics Institute, the National Geographical Institute and from the Meteorological State Agency.

In general, the results for the 803 Local Labor Markets in which the Spanish territory can be divided confirm the existence of temporal and spatial heterogeneity in the determinants of population growth. Even factors as important in Regional Science as distance or size change their effects depending on where and when they are considered. Socio-economic factors that can be affected by specific policy action – such as employment – have an even more heterogeneous effect on population growth across space. Promoting population growth through policies aimed at creating employment can be successful in some local areas but having the opposite effect in others.

The outcomes of the estimation show that the population dynamics of the local areas close to the Mediterranean coast depend to a lesser extent on the economic and sociological factors, and are increasing population regardless of any unfavorable initial condition. On the other hand, population growth in those local areas located in the south or west of the peninsula shows a positive relationship with their initial population size and a negative one with the distance to large cities. Agglomeration economies seem to be operating in this area but not in the rest of the country.

Regional policies aimed to contain population in rural and/or peripheral local areas should always take into account that efforts to increase the educational level or change the economic structure (specialization vs diversification) can have undesired effects on population growth depending on the characteristics of the local area in question. For instance, promoting economic diversification, investing in active education policies or promoting the arrival of highly educated workers into an area may only have the expected effects (i.e., maintaining or increasing population) in some areas, and cause the opposite in others. In summary, the estimated spatial heterogeneity in population growth 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 figure that might not be representative across the whole territory. In addition, policymakers should be informed of local socioeconomic processes, a need that stresses the importance of statistical methods that can reflect spatial heterogeneity, which is a common feature in these cases.

These results lend empirical support to the new approaches adopted by the European Union Cohesion Policy in terms of smart specialization strategies. Instead of promoting global regional actions, this new perspective requires differentiated assessments of regional status and expectations 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 be careful to consider the relevance of the characteristics of neighboring areas.

CHAPTER II. AGEING PLACES IN AN AGEING COUNTRY: LOCAL DYNAMICS OF OLDER POPULATION IN SPAIN

1. Introduction

As a result of the extended life expectancy and lower birth rates, most developed countries show an ageing population structure. This demographic ageing process will intensify in the upcoming decades; projections from Eurostat (2014) reveal that by 2080 in the EU-28 the share of population aged 65 years or more will be close to the 30%, more than 10 percentage points than at present time. In the case of Spain, the National Statistics Institute (INE) forecasts for that same year are close to 40%.

According to demographic theory, population growth is the balance between natural increase (births minus deaths) and net migration (in- minus out-migration). In the case of Spain, the provincial average crude birth rate lies around 9.7‰, while the global fertility rate is approximately 39.8‰ (per thousand women), being these numbers very stable not only through space but also along the period 2001-2014. The same holds for the crude death rate, around 9.34‰. These figures describe a country with very low natural population growth and few chances to reverse that trend provided its labor market characteristics and how they affect life cycle and fertility decisions, as argued by Adsera (2004 and 2005). The source of demographic growth in this kind of scenario is in-migration, not only because of the inflow of new residents but also for the possibility of higher fertility rates that foreigners may bring. In the Spanish case, the evidence suggests that the substantial immigration wave that took place at the beginning of the present century have not had a large impact on the overall fertility of the country (Vila and Martín, 2007), leaving population movements as the main driver behind population change.

Having into account that the life expectancy at the age of 65 is around 20 years, additional pressure on the public budget (pensions and health care costs) and higher taxes to fund the expenditures are the expected outcomes for national governments, as well as a subsequent reduction on national income and economic growth (Feldstein, 2006). Public finances sustainability is specially a hot topic in a country like Spain, where the political de-centralization experienced in the last decades has turned the regulatory framework into a de facto federal system (Sala, 2014) and 17 territorial welfare regimes (corresponding to the 17 Autonomous Communities) coexists with a common structure of social benefits and rights for the whole country (Gallego and Subirats, 2012). Although in Spain the general healthcare services and the dependent adult care have been transferred to the regional governments, the budget assigned to for the autonomous communities doesn't take into account the regional differences observed in the ageing population process.

Undoubtedly, the age structure of the population poses many critical socio economic challenges (Bloom et al., 2003). Populations with high proportions of children and elderly tend to spend more of their income caring for the young and the old, which may depress economic growth, while populations with increasing proportions of the working-age population may profit from the corresponding increase in labor supply if the labor market functions properly and has the capacity to absorb it. Unless some countermeasures are taken, the demographic burden, i.e., the decline in the shares of the working-age population as a consequence of the increase in the share of ageing population, implies a reduction of the national income and economic growth. Furthermore, the demographic ageing process will also increase the average age of the workforce, which may reduce productivity, innovative capacity, competitiveness and therefore economic growth capabilities.

In addition, if the demographic ageing process is heterogeneously distributed across a territory, internal regional disparities may further increase. The hypothesis is that innovative regions will attract young and skilled population while less idea-driven regions will end up with a high concentration of old population (Kanbur and Rapoport, 2005; Fratesi and Riggi, 2007). If you consider the existence of spatial spillovers and age selective migration, the gap between regions in terms of innovation, ageing population and human capital structure becomes even wider. Using regional data for Germany, Gregory and Patuelly (2015) find a growing age demographic divide between rural and urban areas that corresponds to the innovative/less idea-driven regions divide: those regions with an old (and homogeneous) population structure have a small probability to reverse the trend "due to strong neighboring forces and clusterwise path dependence".

However, the work of Van der Gaag and De Beer (2015) shows that although the increase in the share of ageing population will be smaller in urban than in rural regions, in terms of GDP per capita both types of regions will suffer to more or less the same extent the demographic burden. The assumption behind their research is the increase of immigration in those areas where the size of the working-age population decreases (i.e. the demographic burden increases). In other words, in the long term there will be a demographic indicators convergence process as the socio-economic and cultural differences between the EU member states/regions fade.

The objective of this study is to check the existence of a demographic burden convergence or divergence process after controlling for local factors that partially explain the differences on the distribution and growth of ageing population. Using municipalities instead of provinces – NUTS 3 level – as the basic spatial unit of analysis, local specific characteristics that may be attracting or deterring the ageing population concentration (such as the degree of urbanization, the geographical location or the existence of amenities) are included in order to anticipate the demographic burden, and to providing guidelines to the regional governments to design local-specific policies. Certainly, the ageing process also have an impact in the internal structure of the municipalities since the demographic evolution of the different districts is very heterogeneous, as shown in García-Ballesteros and Sanz-Berzal (2002) for Madrid or in Del Valle-Ramos (2007) for Seville, creating a mosaic of neighborhoods with contrasting profiles closely related to age-specific needs and conditions. In this sense, it would be of great interest for this research to descend to a district scale to account for the different characteristics and their relationship with ageing in terms of living standards (accessibility, environmental quality or housing features for example). Unfortunately, data constraints limit severely the local scope that this study can reach, preventing the undertaking of a deeper and richer analysis.

After setting the spatial range of the study, the following queries emerge: Are those municipalities with higher demographic burden at present time converging with those more dynamic both in terms of population and economic trends? Or is the gap between the ageing municipalities belonging to an ageing country widening with respect to those municipalities enjoying different population and economic dynamics? Shall we observe a convergence process and therefore similar ageing population shares across the Spanish territory in future? Do we need local (instead of regional or national) specific policies dealing with ageing population, health services or dependency policies?

Many other questions arise through the development of this chapter. Why some areas have experienced sharp increases of their senior population in recent years while in some others the share of ageing population has diminished? How important are the geo-economic factors, the demographic composition, the economic dynamism or the degree of urbanization (size) of the municipality to the concentration of ageing population? Can some local factor be an attraction pole for the ageing population in some region but a repulsive in another? Does the ageing population react in the same way towards the local characteristics of the area throughout the country?

Collecting disaggregated data for the 8,122 municipalities in 2014, an analysis of spatially heterogeneous growth of the demographic burden in Spain is conducted. Local demographic, socio-economic, geographic and natural features of every municipality are included as local determinants to explain the growth in the share of the ageing population. To allow for different responses across space to the local determinants (spatial non-stationarity), the methodology chosen is the Geographically Weighted Regression approach (GWR hereinafter), which also deals with the spatial autocorrelation of the data.

The chapter is organized as follows. The next Section briefly describe some basic ageing population spatial patterns, comparing Spanish data both at regional and municipal level. The relevance of using the maximum spatial level of disaggregation available is conveniently remarked as the NUTS 3 regional level data (provinces) masks a high degree of internal heterogeneity. Section 3 presents the main features of the empirical convergence model proposed, which will be estimated through GWR methodology, along with the local characteristics that might explain (at least partially) the increase or reduction of the growth in ageing population concentration. The results obtained using are described in Section 4 and compared to the analogous Ordinary Least Squares estimates. This article ends with a summary of the main results, as well as the main conclusions and policy implications drawn from the analysis, in Section 5.

2. Spatial patterns of ageing in Spain: references, data and hypothesis

There is a global tendency to ageing (see Wilson, 2001), which is specially remarked in the OECD countries in general, and in the European Union in particular (see Wilson and Pison, 2004). According to the aggregate figures from EUROSTAT for the EU-28, the share of population older than 64 has grown from an average of 15.8% in 2001, to 18.5% in 2014. Spanish figures are similar to the European average: population older than 64 represented 16.8% in 2001, and this percentage moved to 18.1% in 2014, even after taking into account the massive immigration flows experienced since the early 2000s (Reig, 2007).

Going deeper into the analysis of ageing in Spain, a first striking phenomenon is the strong spatial differences in the distribution of the ageing population, which becomes more relevant when local information is used. For instance, while in 2014 the share of ageing population reach 4% in some local areas, in others the percentage increases almost to 80%. Some studies have tried to understand the causes of this spatial disequilibrium in the localization of elder population. Based on migration as a source of change in the local concentration of ageing population, Lardiés (2009) conducts a qualitative analysis showing that the Spanish ageing population complies with the traditional migration motivations linked to the different stages in lifetime: amenity-related migration, return or family-related migration, and professional care-related migration. Other authors try to give quantitative support to those statements, as Pujadas et al. (2014) and Parreño et al. (2014), confronting the idea of negligible ageing population mobility and identifying the migration intensity and patterns of the older age groups. Using data on Spanish Autonomous Communities (NUTS 2), Blasco (2008) describes a situation where the north and north-western regions present higher ageing rates, the eastern part of the country shows intermediate values, and those areas located in the south, the islands and Madrid region bear the lower ageing population figures. This study also considers several factors that might be related to the ageing process, namely the rural-urban category given to the NUTS 2 region, its poverty levels, health indicators and migration flows.

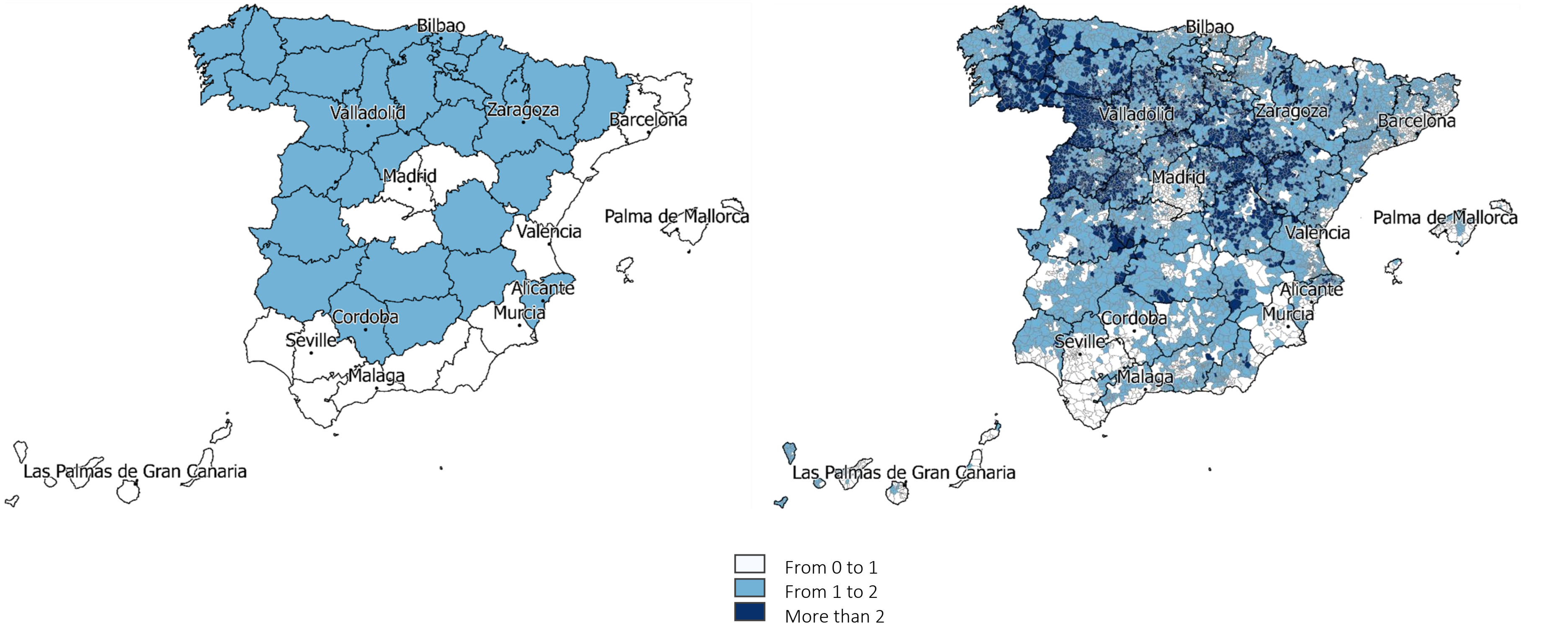
Indeed, the share of elderly population seems to be related to the degree of urbanization of the region, as the low birth and mortality rates that cities exhibit are acting as centripetal forces and therefore enhancing concentration dynamics, as Grafeneder-Weissteiner and Prettner (2013) pointed out. This statement is coherent with what can expected according to Urban Economics or to the literature on population location decision. See Baley (2005), Beeson et al. (2001), Chi and Ventura (2011), Glaeser and Shapiro (2003) or Rozenfeld et al. (2011) among others. For the Spanish case see Gutiérrez et al. (2016), for example, Goerlich and Mas (2008).

Figure 11 represents an adaptation of the traditional location quotient (LQ) to show the spatial concentration of ageing population. This ageing LQ[7] compares the share of population aged 65 or more with the national average, first using provincial (NUTS 3) data and then municipal (NUTS 5) data. The darker the color, the higher concentration of ageing population. Areas in white correspond to "lower specialization" on ageing population regarding the national level. Although there seems to be spatial heterogeneity regardless of the level of spatial disaggregation, it is important to point out the marked local differences that arise within provinces. In other words, there is a high degree of internal heterogeneity within the country and within the provinces as well. Checking LQs for provinces, the demographic burden seems to be especially worrying in those provinces away from the most economically dynamic areas (the Mediterranean corridor, the Ebro corridor and Madrid province and surrounding areas). However, the municipal LQs suggest that ageing population concentration might be associated to the urban-rural dichotomy, as for many medium and large population-sized areas, especially those around the main cities, the ageing LQ is below the national average (i.e. it has a value between

[7] $LQ_{aging} = \left(\frac{Older\ pop_{\ local}}{Total\ pop_{\ local}}\right) / \left(\frac{Older\ pop_{\ national}}{Total\ pop_{\ national}}\right)$

Figure 11. Ageing population concentration (ageing LQ) in Spain at the provincial and municipal levels in 2014

Provinces (NUTS 3 level) **Municipalities (NUTS 5 level)**



Source: Own elaboration with data from the National Geographical Institute (IGN) and the Municipal Register of Inhabitants (Padrón) from the National Statistics Institute (INE).

zero and one). Apart from that, it can be seen that several municipalities lie far above the average, located mainly in the northern half of the country, a fact that is masked when looking at the provincial distribution of the relative share of ageing population.

The descriptive maps of Figure 11 allow us to observe that: (i) the spatial behavior of ageing population indeed fits well with what can be expected from the Urban Economics and population location decision theories, but (ii) in order to observe the patterns more clearly it is necessary to work at the maximum level of spatial disaggregation, otherwise the aggregation of information at provincial level could hide relevant local behaviors. Consequently, in this chapter municipalities are used as basic spatial units of analysis. Municipalities are the smallest administratively defined area in Spain for which there is available data. At the present time there are. In 2014 there were 8,122 municipalities[8].

In order to structure the empirical analysis and model formulation, there should be a combination of the suggestions emerged from the descriptive analysis and of the findings on the previous literature and theoretical approaches, with the aim of proposing a set of hypothesis.

First, the representation of the degree of concentration of ageing, both at provincial and municipal level, provides a comprehensive picture of the present demographic burden by regions in Spain. However, the increase or decrease of the share of ageing population from 2001 to 2014 has to be calculated at both spatial levels (regional and local) as an initial assessment of the existence of a convergence/divergence process in ageing. If places with initial lower ageing levels present a higher increase of ageing population, there is a signal of a spatial convergence pattern on ageing population shares (as Wilson, 2001; and Wilson and Pison, 2004 find for European Union countries). This hypothesis can be formulated as a classic beta convergence model:

[8] In 2001 there were 8,108 municipalities in Spain. The required transformations in the data for the 2014 municipalities were performed in order to have a consistent data set. The data preparation is explained in Section 3.3.

H. 1: There is an overall process of spatial convergence at municipal level to the same share of ageing population: the relationship between the initial percentage of population over 64 years of age and the growth of this ratio is negative and significant.

On the contrary (null hypothesis), if those provinces or municipalities with high initial LQs also show high ageing population growth, or the low-ageing places are able to attract younger population, a spatial divergence process regarding the demographic burden can be foreseen, which might translate into an increase in regional economic disparities.

In Figure 12 the ageing population growth rates for the period 2001 and 2014 are presented, both at provincial level and municipal level. Once again, there exists spatial heterogeneity at both levels. Using provincial data, it might give the impression that there are only a few provinces where the ageing population issue should be addressed. However, when looking at municipal data the convenience of customized local policies is evident in essentially all autonomous communities.

The representation of these demographic age structure variables leads to other interesting questions: why in the last decade some areas have increased their concentration of ageing population while in others the percentage diminished? Is there any local characteristic that explains the appeal or repeal of certain areas to ageing population? Are those characteristics related to the economic and labor market situation (employment/unemployment rate, human capital stock), to the relative or absolute location of the municipality (central or peripheral area), to the population size of the municipality or to the amenities available in the area[9]? What is the role played by size, i.e., the degree of urbanization, of the area towards ageing population concentration?

According to Figure 12, the general hypothesis to consider is that after controlling for the initial ageing share, the convergence/divergence process on ageing population should be influenced by the geographic and socio-economic characteristics of each territory. The

[9] Bearing in mind the low Spanish birth rates and the close to zero natural increase, any significant change on the ageing population share must be caused either by internal mobility of residents or by immigration flows. Although not the focus of this study, results from this chapter could also shed some light on what are the local factors explaining the arrival of native or foreign population to a certain location.

expected behavior of elder people location could be analyzed using what is known about general population patterns (see Baley, 2005; Beeson et al., 2001; Glaeser and Shapiro, 2003 or Gutiérrez et al., 2016), in combination with the studies of Lardiés (2009) and Grafeneder-Weissteiner and Prettner (2013). With this basis, the three additional three hypotheses are suggested:

H. 2. Large municipalities (cities) are more innovative and therefore will attract young (and educated) workers, enhancing an ageing population concentration divergence process. Cities will experience a decrease in the demographic burden while in rural areas the share of ageing population will increase.

H. 3. Peripheral areas not benefiting from the proximity to economically dynamic areas and spillovers will experience an increase in their demographic burden.

H. 4. Those areas with dynamic labor markets, where new workers can be easily pulled, will attract young workers and therefore will have lower ageing population share in future.

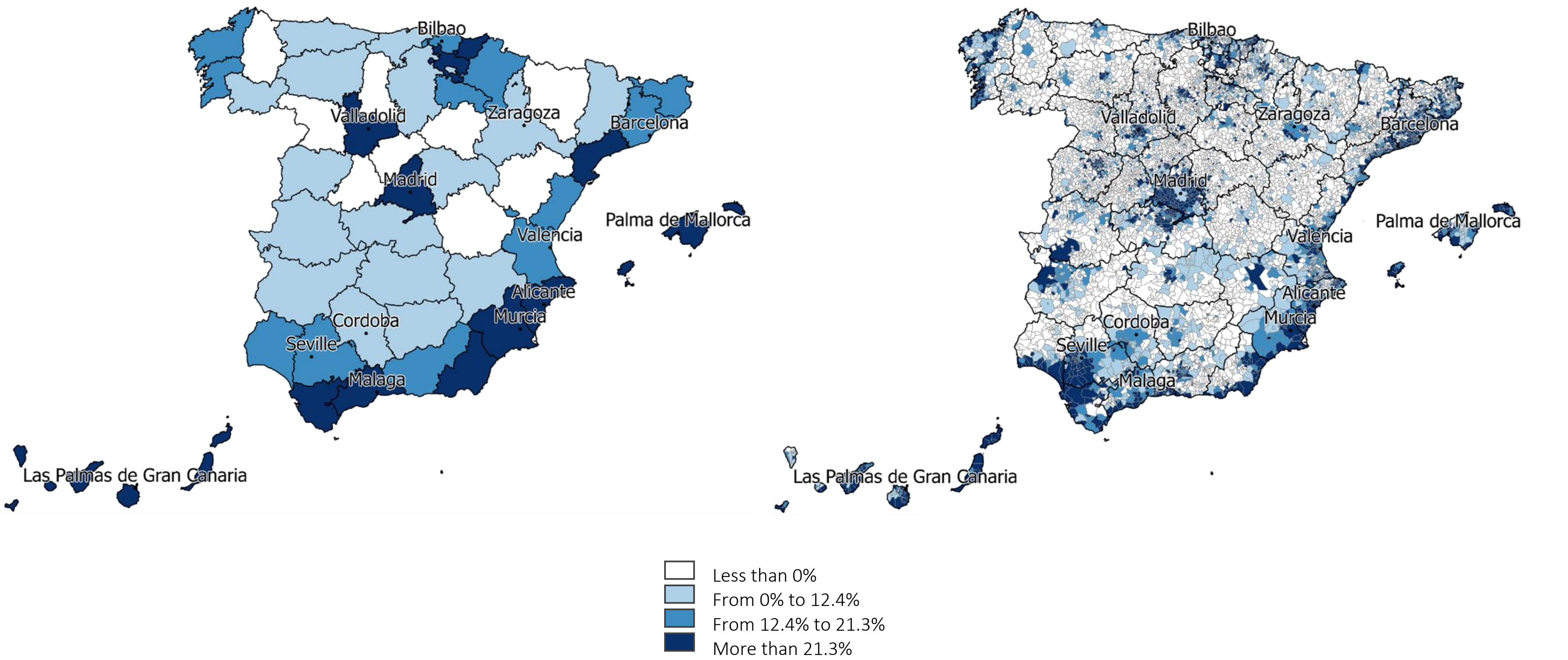
H. 5. Those places offering better climate conditions or closer to the coast will attract both ageing population and workforce, being the final outcome unclear over the demographic burden.

However, as is supported in Coffey and Polèse (1988), Patridge et al. (2008), Glaeser et al. (2014), and in line with the previous Chapter, the population location dynamics are subject to certain factors that affect in the location decision and that can change and give rise to different specific behaviors depending on the local characteristics. In order to analyze this source of spatial heterogeneity the following two hypotheses appear:

H. 6. Although a general process of ageing population concentration convergence is expected (Hip. 1), given their specific characteristics some places will be able to reverse the trend and follow their own path (either as places with ageing concentration higher than expected or vice versa). Allowing for spatial non-stationarity, local areas where the beta parameter suggests the existence of a convergence or, conversely, a divergence process will be exposed.

Figure 12. Growth rate of aging population concentration at the provincial and municipal levels between 2001 and 2014

Provinces (NUTS 3) **Municipalities (NUTS 5)**



Source: Own elaboration with data from the National Geographical Institute (IGN) and the Municipal Register of Inhabitants (Padrón) from the National Statistics Institute (INE).

H. 7. Local factors determining the ageing population concentration process might be contributing to a convergence process in some areas while having the opposite effect in others.

Regarding to the spatial heterogeneity, the focus is not only on testing the previous hypothesis, but also on identifying the existence of spatial patterns across the Spanish geography.

3. Empirical approach: convergence model and non-parametric estimation

3.1. *β-convergence estimated through Geographically Weighted Regressions*

The β-convergence analysis, originally proposed by Baumol (1986), Barro et al. (1991), Barro and Sala-i-Marti (1992) and Sala-i-Marti (1996), was born in the context of the Neoclassical Economic Growth theoretical framework, which is mainly based on the role of the decreasing returns for the different production factors. Provided the information on the relevant inputs or determinants and the existence of decreasing returns in the production process, Solow's economic growth (1956) concludes that, in the long run, all territories will converge to the same level of per capita income. The β-convergence model takes shape in a simple empirical equation where one variable's growth rate for a period of time depends on the initial level of that variable. So, the β-convergence approach is based in the following equation:

$$log\left({y_i^T}/{y_i^0}\right) = \alpha_0 + \beta_0 log(y_i^0) + \varepsilon_i \qquad [6]$$

where $log(y_i^T / y_i^0)$ is the logarithm of growth of the variable under study (usually per capita income), of the i spatial unit from the initial year (y_i^0) to the final year (y_i^T). Traditionally, this model shows if initially poorer countries tend to grow faster or not than the initially wealthier ones, as the neoclassic approach predict. If the β estimated parameter has a negative sign, regions would experience a convergence process in the long run, as initial lower levels of income per capita are associated with higher growth rates.

A number of studies have applied this approach to study convergence patterns in different demographic issues. The most usual application is to study convergence in average life-expectancy (see Becker et al., 2005; Bourguignon and Morrisson, 2002; Easterlin, 2000; Edwars, 2011; Goesling and Firebaugh, 2004; Neumayer, 2003 and 2004, Pradhan et al., 2003 or Ram, 2006 among others), but it has been used also to analyze fertility (Dorious, 2008) or mortality (Goli and Arokisamy, 2014).

In line with this previous literature, this chapter offers an application of the β-convergence approach to analyse convergence patterns in ageing, linking the share of ageing population and its growth over time. It means that in the original equation [Eq.6], the per capita income variable is replaced by the ageing population share, testing the existence of a spatial convergence process towards similar levels of ageing population shares in the long term, as Van der Gaag and De Beer (2015) suggest (Hypothesis 1), or the existence of a spatial divergence process as Kanbur and Rapoport (2005), Fratesi and Riggi (2007) and Gregory and Patuelly (2015) find. Using non-parametric methodologies it is possible to estimate a β parameter for each municipality, which allows to differentiate those areas participating in an age population convergence process from the ones following their own divergent path. Thus, the relationship between the initial share of population older than 64 years old and the growth of that population share can be negative and significant for some areas (convergence) while positive and significant for others (divergence).

The method chosen to estimate the conditional convergence model proposed above is again GWR, as the main purpose here is to explain the existence of an ageing concentration convergence or divergence process at the municipal level in Spain, as well as the different effects that each control variable might exert over the ageing process across space due to the interrelationships between neighboring areas. In this sense, the focus on spatial heterogeneity is also present in this study. The technical details related to this methodological approach can be found in Section 3 of Chapter I of this dissertation.

3.2. Control variables in the conditional model

When the only independent variable is the initial level of per capita income, as it is the case in Equation 6, the parameter estimated is considered an unconditional β-convergence parameter. However, it is common to include a set of explanatory variables reflecting the influence of specific characteristics of each territory. When these so-called control variables $\sum_{z=1}^{Z} \gamma^z x_i^z$ are included in mentioned Equation, the framework changes to a conditional β-convergence model, where:

$$log\left({}^{y_i^T}\Big/_{y_i^0}\right) = \alpha_0 + \beta_0 log(y_i^0) + \sum_{z=1}^{Z} \gamma^z x_i^z + \varepsilon_i \qquad [7]$$

Besides working as control variables, these local characteristics might play a relevant role in attracting or deterring ageing population concentration. A close analysis of the estimates and significance of the different γ^z parameters associated to those descriptive variables will allow to test the hypotheses related to the initial size (Hip.2), the importance of location, i.e. centrality or peripherality of the municipality (Hip. 3), the local labor market conditions (Hip.4) or the amenity related features (Hip. 5).

Given the scarcity of data at municipal level, the selection of these local characteristics is commonly determined by data availability. There is access to data on the two main factors in Urban Economic theory: size (proxied by population) and relative location (measured as the linear distance to the nearest metropolitan area, named *DistSize*). These urban factors (*UF*) combined with some socioeconomic factors (*SEF*) and amenity related variables (*EA*) bring along the following equation [3]:

$$\begin{aligned} log\left({}^{y_i^T}\Big/_{y_i^0}\right) = {} & \alpha_0 + \beta_0 log(y_i^0) + \gamma_i^{UF} Size + \gamma_i^{UF} DistSize \\ & + \sum_{SEF=1}^{3} \gamma^{SEF} x_i^{SEF} + \sum_{EA=1}^{5} \gamma^{EA} x_i^{EA} + \varepsilon_i \end{aligned} \qquad [8]$$

The urban factors (size and distance) should be analyzed jointly as they are used to classify areas as urban/rural or center/periphery. For instance, those municipalities that do not surpass certain critical mass (i.e., rural areas), and also considered peripheral due to their location, might be experiencing an acute ageing process (mainly related to depopulation), meanwhile municipalities of the same size but central might be attractive places of residence for young population (lower housing costs, wide social amenities supply, etc.), and therefore, ageing population concentration would decrease. On the other hand, urban areas might experience different population growth dynamics depending first on their population size (small-medium city vs. large metropolitan area), and second, on their relative location in the Spanish territory (Coffey and Polèse, 1988; Polèse and Shearmur, 2004; Polèse et al., 2007).

As well as the urban factors, a wide array of socioeconomic ($\sum_{SEF=1}^{3} \gamma^{SEF} x_i^{SEF}$) and amenity related factors ($\sum_{EA=1}^{5} \gamma^{EA} x_i^{EA}$) should be considered when analysing population dynamics (Chi and Ventura, 2011). Underlining the heterogeneous character of the population structure across space and time, Glaeser et al. (2014) also suggests the need to consider the possibility of different variables having different impacts depending on the area, the moment and the context regarded. The effects of a given variable might vary considerably across regions depending on the spatial framework considered (Shearmur and Polèse, 2005; Shearmur et al., 2007; Glaeser et al., 2014), especially if there is clear spatial heterogeneity, calling into question whether a single global estimate is able to explain spatial ageing processes accurately.

3.3. Data: Spanish census and municipal register of inhabitants

Table 1 shows basic information on the variables (local specific factors) and data sources used in each case. There are two sources of information that collect and provide data at the municipal level: the Population and Housing Census and the Municipal Register of Inhabitants (Padrón). The Census is the main source of information on the socio-economic characteristics of population (age, gender, educational attainment, labor situation, etc.), and is commonly used in studies of local population or employment growth. The information from this source can be complemented with the Padrón, an official record containing information about the residents of a certain municipality. National population figures coming from this record are released on a yearly basis while the Census is only carried out every ten years. The data required to build the distances from every municipality to the nearest metropolitan area and to the coast was retrieved from the National Geographic Institute (IGN), and amenity related figures come from the State Meteorological Agency (AEMET).

The analysis covers the time span between 2001 and 2014, which accounts for the massive migratory inflow that arrived to Spain in the first years of the 2000s, and enables the use of the data contained in the Census of 2001 that is fundamental for this chapter since it is the only source of municipal information on local sectoral employment, educational level and other municipal features, essential for the construction of the agricultural LQ, the unemployment rate and the socio-economic index. 2014 was chosen as the last year because it was the last one available in the Padrón to the date this study was conducted.

In order to have a set of comparable areas between the beginning and the end of the period considered, the changes in the municipal structure observed in 2014 were numerically reversed, so it could match the one of 2001. The variables regarded were added up in the case of a separation, or proportionally distributed according to their original share in the case of a union. Following this procedure, the final database contains 8,108 spatial units.

As a measure of the degree of centrality/peripherality of the municipalities, a distance to the main urban areas is included. It this case a population threshold of 500,000 inhabitants was considered in order to distinguish the main Spanish cities that possess a clear urban character (namely Madrid, Barcelona, Valencia, Seville, Zaragoza and Malaga). In spite of this consideration, since the methodology applied takes into account a certain neighborhood for each unit, other important cities of the country are also indicated in the maps to make the interpretation of the results clearer.

Meteorological indicators, as well as cultural and health facilities, are added as a mean to reflect the attractiveness of a municipality regarding its natural and manmade amenities, which are considered as important factors for population and economic growth since the seminal work of Graves (1980). The climate related variables were built with all the information available to elude abnormal cyclic behavior that might conceal the actual weather profile of an area.

Table 4. Control variables included in the empirical model of aging population concentration convergence for Spain

$$log\left(\frac{y_i^T}{y_i^0}\right) = \alpha_0 + \beta_0 \log(y_i^0) + \gamma_i^{UF} Size + \gamma_i^{UF} DistSize + \sum_{SEF=1}^{3} \gamma^{SEF} x_i^{SEF} + \sum_{EA=1}^{5} \gamma^{EA} x_i^{EA} + \varepsilon_i$$

<table>
<tr><th colspan="3">Variable description</th><th>Source</th></tr>
<tr><td>Dependent variable</td><td>$\boldsymbol{log\left(\frac{y_i^T}{y_i^0}\right)}$</td><td>- Increment of the share of population of more than 64 years old from 2001 to 2014</td><td rowspan="2">Municipal Register of Inhabitants (Padrón) from the National Statistics Institute (INE)</td></tr>
<tr><td rowspan="7">Independent variables</td><td>$\boldsymbol{log(y_i^0)}$</td><td>- Share of population of more than 64 years old in 2001</td></tr>
<tr><td rowspan="2">Urban factors: size and distance to size
$\gamma_i^{UF} Size + \gamma_i^{UF} DistSize$</td><td rowspan="2">- Size: logarithm of population size in 2001
- Distance to nearest metropolitan area of more than 500.000 inhabitants in 2001</td><td>Population and Housing Census (2001) from the National Statistics Institute (INE)</td></tr>
<tr><td>Digital maps of the National Geographic Institute (IGN)</td></tr>
<tr><td>Socioeconomic factors
$\sum_{SEF=1}^{3} \gamma^{SEF} x_i^{SEF}$</td><td>- Socio-economic index: indicator of the standard of living according to several municipal features in 2001
- Location quotient in agricultural activities in 2001
- Educational level: percentage of population with university degree in 2001
- Unemployment rate in 2001
- Share of foreign population in 2001</td><td>Population and Housing Census (2001) from the National Statistics Institute (INE)</td></tr>
<tr><td rowspan="3">Natural and cultural environmental amenities
$\sum_{EA=1}^{5} \gamma^{EA} x_i^{EA}$</td><td>- Distance to the nearest coast</td><td>Digital maps of the National Geographic Institute (IGN)</td></tr>
<tr><td>- Average annual rainfall from 1987 to 2007
- Minimum temperature January (average from 1987 to 2007)
- Maximum temperature July (average from 1987 to 2007)</td><td>State Meteorological Agency (AEMET)</td></tr>
<tr><td>- Cultural facilities per capita in 2001
- Health facilities per capita in 2001</td><td>Population and Housing Census (2001) from the National Statistics Institute (INE)</td></tr>
</table>

As a final remark on the data and the estimation configuration, it is necessary to mention that most of the explanatory variables are taken for 2001 in order to avoid the reverse causality problem that might arise when the independent factors are measured within the time range under study. The underlying rationale is that the independent factors measured at the beginning of the period cannot be affected by the dependent variable. This procedure is common in the literature that use models where the left- and right-hand-side variables share a bidirectional causality relationship, as in this case.

4. Results: a spatially heterogeneous ageing process

Trying to understand the territorial processes of ageing population considering the specific characteristics of the area, allowing for spatial heterogeneity and using the municipalities as the basic spatial unit is a complex task. While the OLS estimation will provide fifteen global parameters (including one intercept), the GWR approach will provide 15 x 8,108 parameters, one for each factor and each one of the municipalities which reflect the effect of each explanatory variable on the complete sample.

Analogously to the presentation of the results followed in the previous Chapter, in the first place global estimations (OLS approach) will be commented and compared with the ones obtained by through the non-parametric approach (GWR). The OLS estimates may veil the effect of each factor on each region, something particularly troublesome if the response of one variable varies cross space and the associated variability is large (Ali et al., 2007; and Partridge et al., 2008).

The second column on Table 5 presents the results for the OLS model. The following five columns show the quartile intervals for the GWR estimations for each factor affecting the ageing population process. The last column shows the F3 test, indicating the presence of significant variability in the local coefficients, in contrast to the global estimation obtained through OLS.

After presenting the results and discussing the existence of a convergence/divergence ageing process at local level, the factors where the GWR estimates add useful and significant information on territorial ageing determinants will be analyzed in detail. Only those factors showing a spatially differentiated effect will be represented, to finally deduce spatial or regional policy implications for Spain.

When a global OLS approach is chosen, the parameter for the initial share of ageing population – which, as explained before, can be interpreted as the β-convergence parameter – has a negative and significant value, suggesting the existence of an overall process of ageing convergence (Hip 1). Thus, the convergence speed would be 2.46%, a result that nonetheless has to be interpreted alongside with the rest of the estimations obtained. OLS estimations in Table 5 show that almost all the control variables are significant, which means that the process of convergence is taking place for different steady states.

The OLS estimations for the two variables associated to the urban-rural, central-peripheral character of a municipality (i.e., the initial population size and the distance to the nearest relevant metropolitan area) support hypothesis 2 and 3. Size is negatively associated to the ageing process and distance has a positive relationship with the local ageing. In other words, the larger the city or municipality, the lower the growth of the share of ageing population; while the more peripheral the municipality, the stronger the ageing process.

Almost all the estimators for the socioeconomic conditions are significant and have the expected sign (hypothesis 4). Economically dynamic areas present a lower tendency towards ageing population processes. The parameter for the socioeconomic index is significant and negative (higher living standards implies lower ageing population growth) and the unemployment rate suggests that ageing areas are basically losing workforce and population due to the lack of labor opportunities. Also, municipalities specialized in agricultural activities show higher increases in the percentage of older population. Nevertheless, the educational level or the share of foreign population are not significant.

Table 5. Ageing population process determinants at municipal level between 2001 and 2014

Variable	Global (OLS)		Min.	1st Qu.	Median	3rd Qu.	Max.	F3 test[(a)]
Intercept	0.122	***	-0.682	0.000	0.126	0.238	1.049	+++
Initial share of population > 64 years old	-0.274	***	-0.661	-0.415	-0.330	-0.221	0.115	+++
Log. of population size	-0.001	**	-0.018	-0.007	-0.003	0.0003	0.014	+++
Distance to nearest metropolitan area	0.005	***	-0.212	-0.006	0.010	0.031	0.176	+++
Socio economic index	-0.032	***	-0.161	-0.058	-0.025	0.015	0.121	+++
LQ agricultural activities	0.001	***	-0.001	0.001	0.001	0.001	0.001	+++
Educational level	0.009		-0.273	-0.034	0.010	0.071	0.265	+++
Unemployment rate	0.001	***	-0.002	-0.001	0.001	0.001	0.003	+++
Share of foreign population	-0.005		-1.060	-0.311	-0.181	-0.056	0.541	+++
Distance to the nearest coast	0.006	***	-0.223	-0.005	0.010	0.031	0.212	+++
Avg. annual rainfall from 1987 to 2007	0.001	***	-0.030	-0.002	0.001	0.005	0.021	+++
Min. temperature January	-0.001	***	-0.051	-0.005	-0.001	0.001	0.030	+++
Max. temperature July	-0.002	***	-0.030	-0.005	-0.002	0.002	0.025	+++
Cultural facilities pc	-0.087		-2.727	-0.785	-0.132	0.319	1.826	
Health facilities pc	0.129		-5.564	-0.441	0.096	0.480	2.073	
Adjusted R^2 OLS	0.5556			F-statistic	72.63 ***			
Adjusted R^2 GWR	0.6562							
F1 test[(b)]	0.8376	+++						
F2 test[(c)]	3.5454	+++						

- Dependent variable: increment of the share of population older than 64 years old from 2001 to 2014

- */**/*** 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.

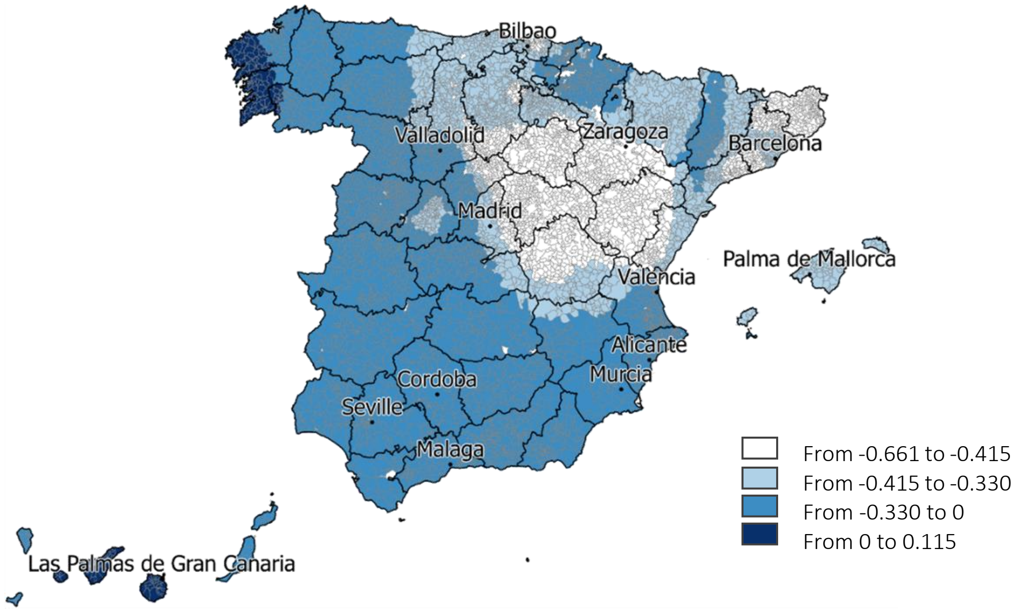
Regarding the natural amenities, such as climate and geographical conditions, OLS results show that places with "gentle" weather conditions seem to attract ageing population. On the other hand, cultural amenities and access to health facilities do not show significant parameters. These results only partially support hypotheses 5.

Although reasonable, the global OLS estimates can be masking spatial heterogeneity and non-stationarity. Adopting a GWR approach it can be observed how the previous conclusions might change across the space. The F3 test points out the spatially heterogeneous distribution of the impacts of all the globally significant variables mentioned above (including those ones that could not have been significant in the global OLS estimation, such as the share of foreign population or the educational level, but be significant only in some areas). The F3 test for all the variables confirms that the Spanish ageing process is very heterogeneous across the country.

Thus, with regard to "the ageing β-convergence parameter", it can be seen that parameters spatially shift from negative to positive sign (hypothesis 6). When the estimation of those significant parameters for each area are represented (Figure 13), a β-convergence trend can be identified across the country, in almost all the Spanish territory (as the global OLS result suggests), but in some places the parameter is zero or even positive.

This map shows that from the center (Madrid) towards the north-east (Ebro-axis) and Catalonia, there is a very fast ageing convergence process. On the contrary, the south and west of the country exhibit low levels of convergence (close to cero) and even an ageing divergence process (positive and significant ageing β-convergence parameter) in some coastal municipalities. This is also the case in the northwest coast and in some specific points of the Basque country (large metropolitan area). So, apart from the existence of different steady states according to the significance of the control variables included in the OLS estimation, across the Spanish geography there are different speeds of convergence and even a divergence ageing process, what calls the attention to a spatially heterogeneous ageing profile.

Figure 13. GWR coefficients for the initial share of older population



Source: Own elaboration with data from the National Geographical Institute (IGN) and the Municipal Register of Inhabitants (Padrón) from the National Statistics Institute (INE).

Spatial heterogeneity is also found in the control variables (hypothesis 7), since most of them are significant and vary their sign across the country. In Figure 14 the GWR parameters for the urban factors are shown: initial population size to the left (map 4a) and distance to main metropolitan areas to the right (map 4b). After representing these parameters it is clear how, in some areas, size do not have the aforementioned restraining effect on ageing, and it even triggers the growth of the ageing population share. The same happens with the central or peripheral condition of the area: its effect changes across space. The effect of being in a coastal location also differs depending on the specific coastal point concerned (see Figure 15). Thus, ageing population share grows faster in coastal municipalities, an increase that is even more intense if it is a coastal and also urban municipality.

5. Summary and main conclusions

The location and migration decisions of the individuals are usually studied from a labor market perspective, focusing on the movements of the active labor force and its economic effects. However, from some time ago the focus is starting to shift towards the consequences of an ageing and ageing scenarios – mainly due to the loss of young and active population –, especially regarding regional financing, wealth, innovation activity, productivity and economic dynamism.

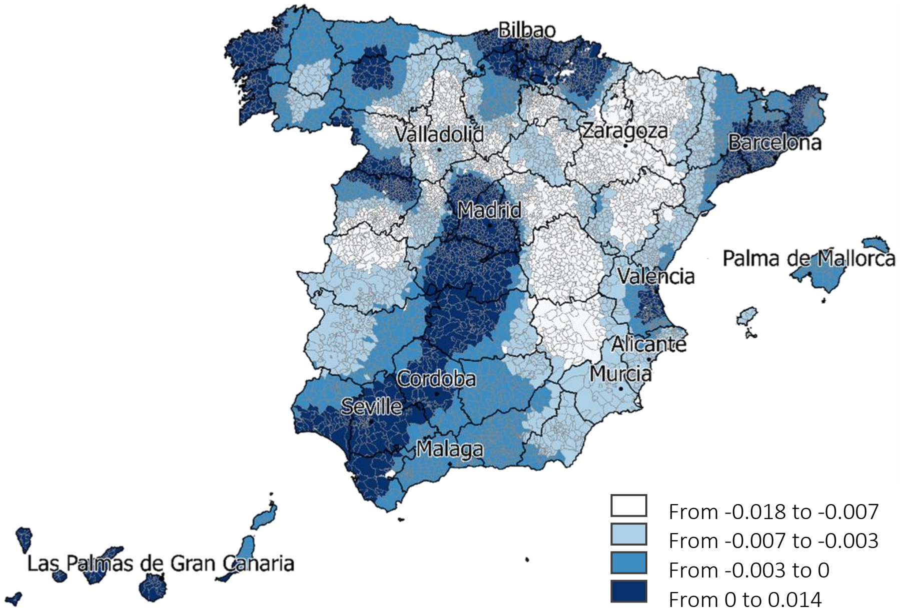
In this chapter the interest lies in analyzing those local factors that determine or explain the high concentration and growth of ageing population in Spain. This country poses an appealing case for the matter taking into account the confluence of two circumstances: the migratory inflow at the start of this century, which influenced significantly the demographic composition and growth of the Spanish population; and the progressive internal ageing process, which is unevenly distributed across space.

Collecting data on several socioeconomic characteristics of the population of the 8,122 Spanish municipalities between 2001 and 2014, an ageing conditional β-convergence model is proposed and estimated both with OLS and GWR. The GWR approach allows us to identify the different effects exerted by the local factor on the ageing process.

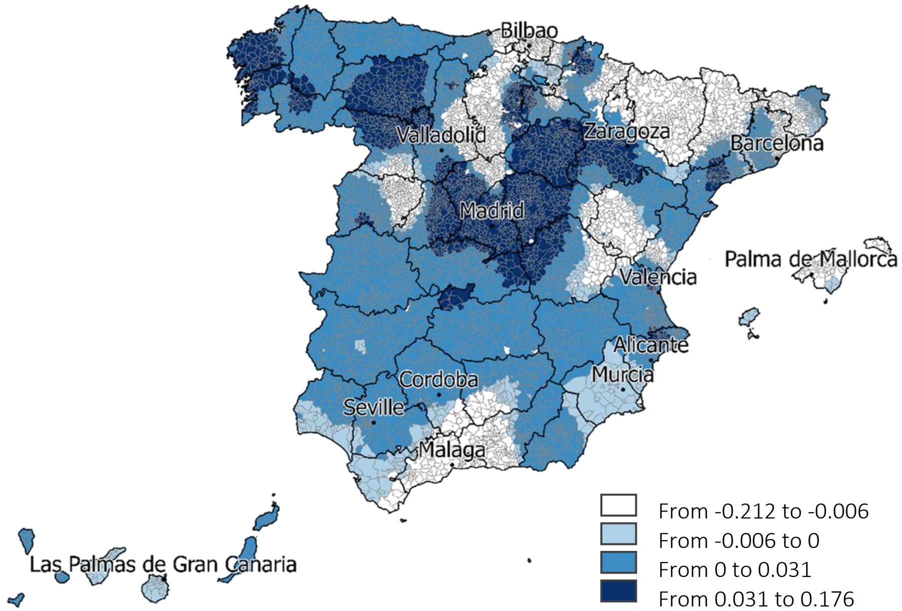
OLS results suggest an inverse relationship between the initial share of ageing population and the growth of that share between 2001 and 2011, which means the existence of an overall ageing convergence process, being the speed of convergence 2.46%. However, the urban-rural and central-peripheral character of the municipality -depicted by the initial population size and the distance to the nearest metropolitan area-, exhibit a significant effect and the expected sign: negative for size and positive for distance. In other words, the larger the municipality, the lower the ageing population growth; and the more remote the municipality, the faster the speed growth of ageing population.

Figure 14. GWR coefficients for urban factors

(14a) Log. of population size

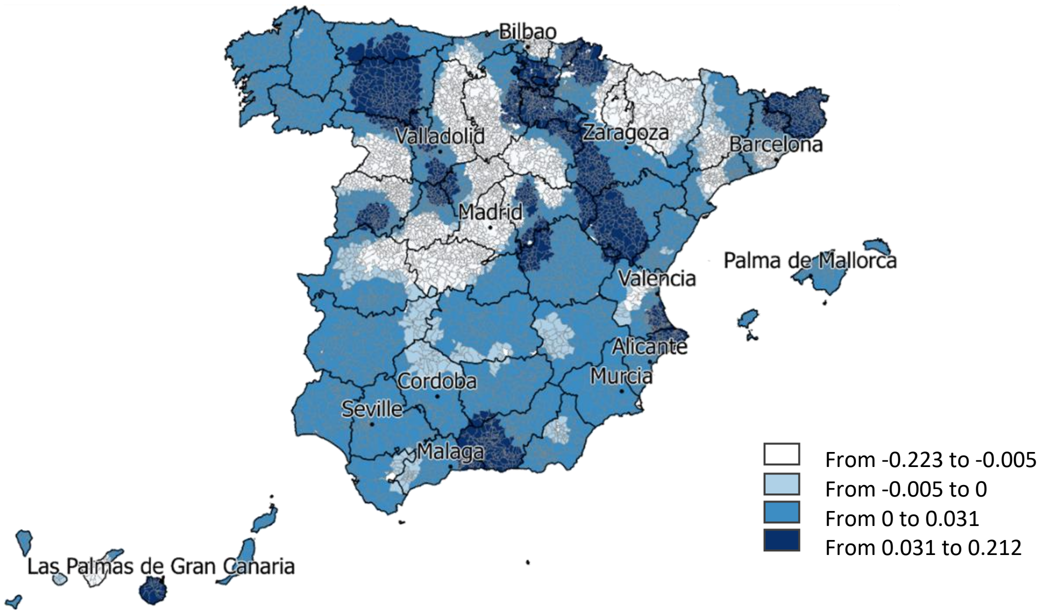


(14b) Distance to nearest metropolitan area (>500,000 inhab.)



Source: Own elaboration with data from the National Geographical Institute (IGN) and the Municipal Register of Inhabitants (Padrón) from the National Statistics Institute (INE).

Figure 15. GWR coefficients for the distance to the coast



Source: Own elaboration with data from the National Geographical Institute (IGN) and the Municipal Register of Inhabitants (Padrón) from the National Statistics Institute (INE).

However, estimating the same model adopting a non-parametric approach allows for addressing spatial non-stationarity, so the varying effect of one variable over the ageing population growth across the space might be observed. In this case, results show the co-existence of ageing convergence areas (from Madrid towards the north-east) with others clearly diverging and following their own growth path (north-west).

Spatially differentiated results show how the ageing population process is not necessarily connected to the birth-mortality rates but with the population migration flows and therefore the socioeconomic characteristics of the local area. Ageing population growth is faster in the north and northwest area while in the east and coastal municipalities it is less severe. After controlling for many other local characteristics of the municipality, results show that ageing population will be converging at different speeds and even diverging in some areas. Even the effects of the control variables such as size and distance to large metropolitan areas vary and change their sign across space. Overall, cities are ageing at a slower pace than rural (smaller) municipalities; but there are some cities showing the opposite effect over ageing population growth, such as those located across the Ebro-axis, one of the most economically dynamic areas in Spain. Relative distance or distance to the nearest metropolitan area exerts a negative and significant effect over ageing population growth almost everywhere in different degrees.

The spatial heterogeneity observed when working at a high level of disaggregation adds complexity to any policy measures oriented towards the reduction of the demographic burden, either through birth promotion or, better again, the attraction of young workforce. However, treating the whole territory as a homogeneous space and applying one-size-fits-all policies would condemn such policies to failure even before being implemented. The spatial heterogeneity observed in the population ageing process requires local tailored actions considering the range of determinants under policy control.

PART II. SPATIAL EMPLOYMENT DYNAMICS

CHAPTER III. WHY SOME AREAS DO BETTER THAN OTHERS IN TERMS OF EMPLOYMENT?

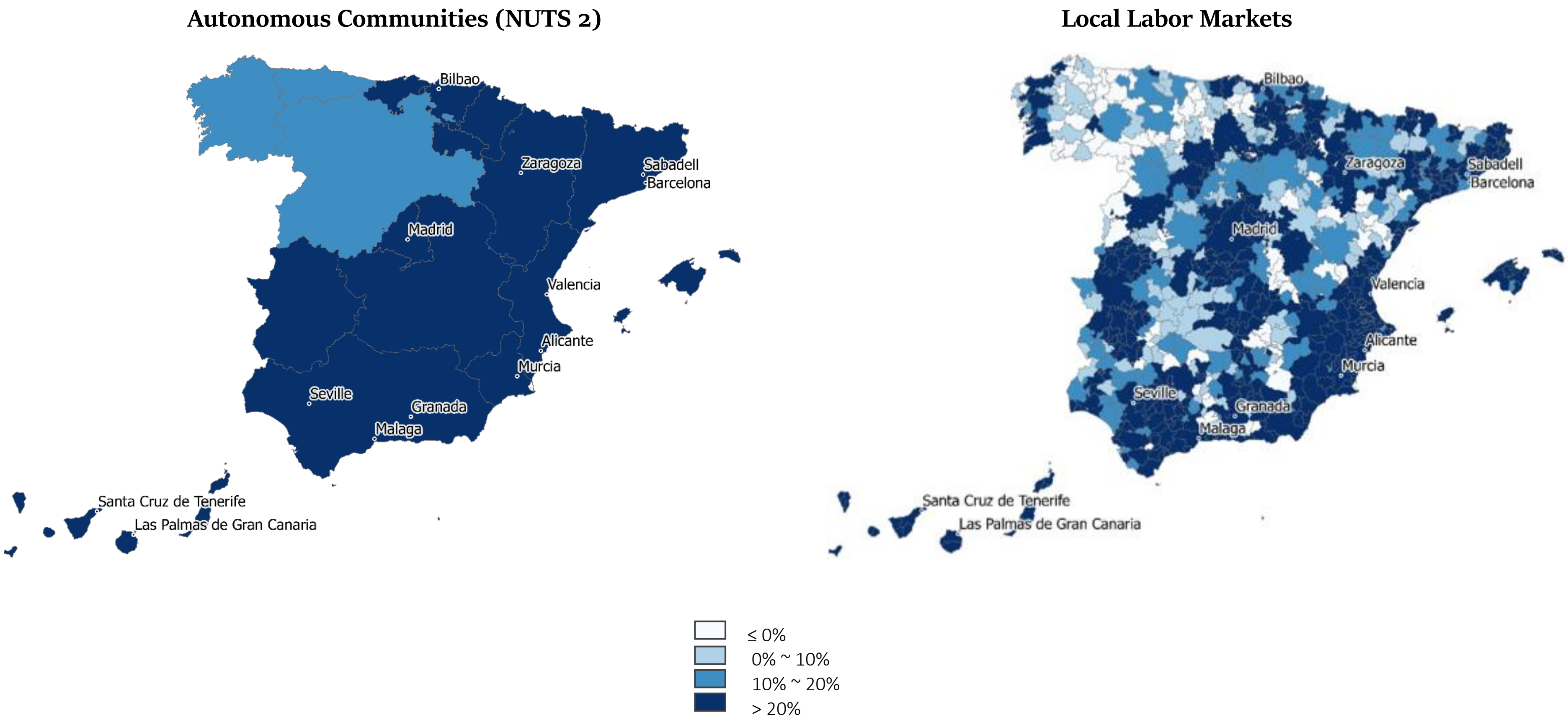
1. Introduction

What places grow faster? What are the main factors explaining higher economic growth at local level? How important is the size of the locality or its location a country? How relevant is the region to which they belong? Is there any room for policies to promote local growth or local growth might be deterministic and basically determined by socio-economic and-or structural factors out of any policy scope? Although there is a large body of empirical and theoretical literature about growth, when talking about growth at the local level, the empirical conclusions are still limited and unclear.

Several reasons explain the limitations of the empirical analysis of local growth. First, for most countries there are strong data limitations. In Europe, any economic growth analysis based on domestic production becomes a challenging task if requiring highly spatially disaggregated data. A suitable alternative is to consider employment growth instead, which figures are generally available at a higher degree of spatial disaggregation. A second reason explaining the scarcity of studies on local growth is that, when they exist, its results are sometimes inconsistent with economic growth theories. Exceptions to the expected pattern tend to occur during periods of fast growth in resource dependent regions (Gunton, 2003), or during certain periods of time (Massey, 1995; Blien et al., 2006 and Shearmur and Polèse, 2007), or in certain areas within a region (Strambach, 2001).

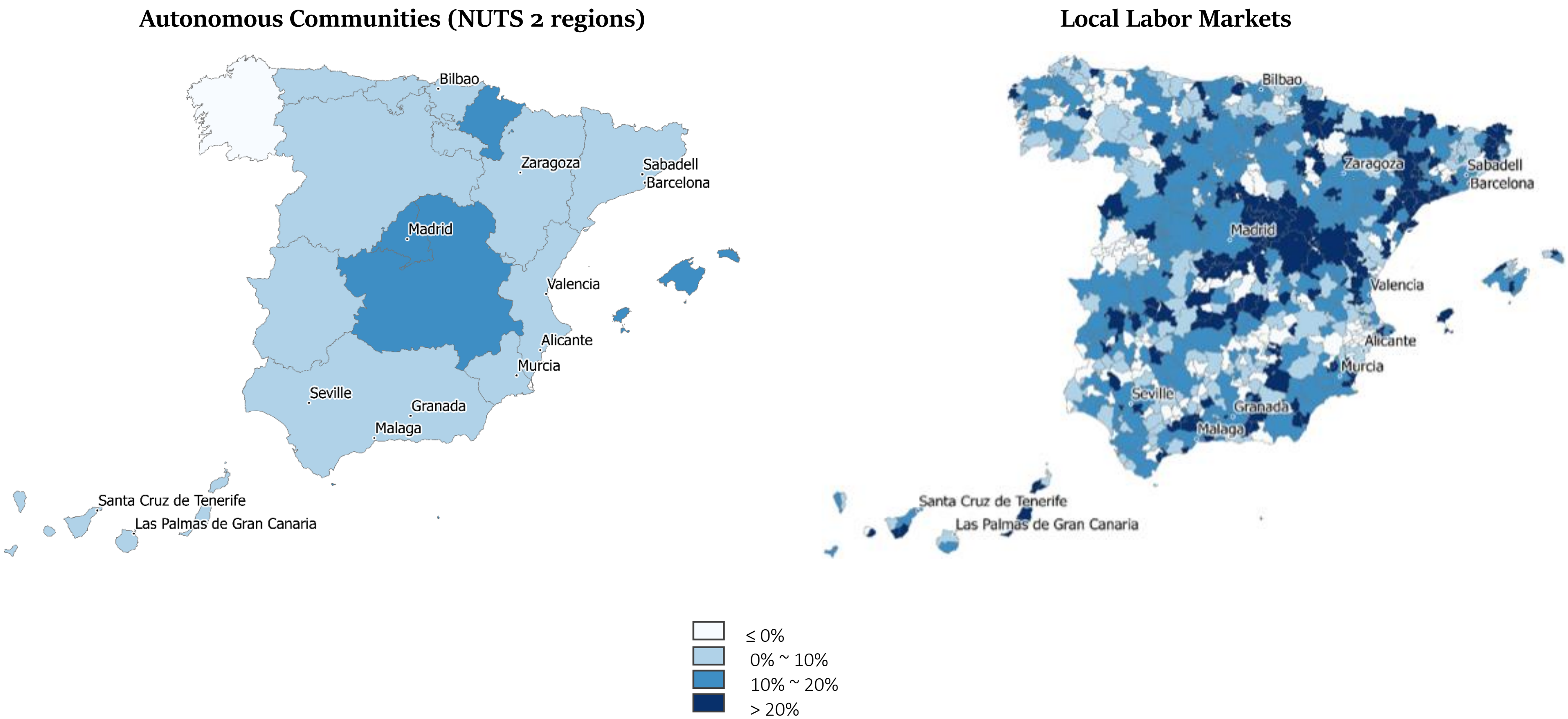
Bearing in mind the inconsistencies of the existing literature on local growth, and data limitations for domestic production beyond NUTS level 3 regions, in this chapter an empirical analysis of local employment growth is proposed for the whole Spanish territory, covering twenty years of growth. Spain poses an interesting case for several reasons. In the last two decades, the overall Spanish economy has experienced double-digit growth figures (33.52% from 1991 to 2001 and 19.01% from 2001 and 2011) as well as strong employment growth (from 12.42 million employees in 1991 to 16.33 in 2001 to 17.51 in 2011), but growth has not been evenly distributed, either across space or time. Moreover, the interest of the Spanish case rises when considering data at the local level, given the high degree of spatial heterogeneity on employment growth within the regions. Figures 16 and 17 show employment growth rates at the autonomous community level (NUTS 2) and at local level using Local Labor Markets as the basic spatial unit for both decades. In the Spanish case, not only the overall employment growth picture differs depending on the spatial unit chosen (with local clusters of high employment growth in regions experiencing low growth rates and vice versa) but also over time, with deprived areas along the first decade performing relatively well in the second one (and vice versa). Among these remarkable differences in time and space, spatial patterns that, in principle, are in line with the Urban Economics and New Economic Geography literature can be identified, consistently with the idea that the faster growing regions and localities (both in terms of employment and GDP) are those classified as large urban areas, containing the main cities (labeled in the maps) as well as their neighbors, which suggests the importance of agglomeration and locational effects. The spatial distribution of the most dynamic regions and localities might not be therefore random. The Mediterranean coast comprises several areas growing above the national average in the first decade; as so are Madrid, Bilbao and their surroundings. However, in the second decade only a few regions on the Mediterranean coast remain on the top positions, while several higher growth areas are located in the North-East quarter enclosed by Bilbao, Madrid, Valencia and Barcelona.

Figure 16. Employment growth rates for Spanish Autonomous Communities and Local Labor Markets between 1991 and 2001



Source: Own elaboration with data from the National Geographical Institute (IGN) and the National Statistics Institute (INE).

Figure 17. Employment growth rates for Spanish Autonomous Communities and Local Labor Markets between 1991 and 2001



Source: Own elaboration with data from the National Geographical Institute (IGN) and the National Statistics Institute (INE).

Given the spatial and temporal heterogeneity observed in local employment growth figures, the aim of this chapter is to identify the existence of any empirical regularities with regards to employment growth, paying special attention to the influence of structural factors (location, size and geography), as well as to the presence of spatial interactions. The objective is to uncover the relationship between the characteristics of the area and its employment growth, assessing which attributes could be intensely contributing to local employment growth. In order to do so, the variables included in the empirical model will play a crucial role, but also the spatial units and the econometric approach chosen. Eventually, the success of an area does not only depend on its own characteristics but also on the ones of its neighbors. Moreover, it is important to bear in mind that some of the characteristics might be subject to regional policy, while others cannot or are only modifiable in the long term.

This chapter is organized as follows. In the next Section the relevant literature on employment growth is summarized, trying to focus on the main contributions over the factors determining local employment growth. The review of existing empirical exercises will be used as a guide in the selection of the variables and the spatial unit chosen, which will be the object of discussion of Section 3. In Section 4, the use of a spatially autocorrelated model in the analysis of local employment growth in Spain is justified. Section 5 presents the empirical results, and finally Section 6 summarizes the main conclusions and policy implications derived from the analysis.

2. Factors of employment growth: literature review

Economic growth is one of the main topics in Economics, and consequently, the amount of analysis previously undertaken is vast. If only focusing on local growth, the yet extensive number of relevant contributions could be easily summarized from an empirical perspective by classifying the type of factors explaining growth in three main groups.

The first group consist of factors considered difficult to measure and that include substantial qualitative components, making their effects also difficult to quantify and evaluate using a statistical approach (Doloreux et al., 2001). These factors encompass elements such as the local institutional context (Cooke et al., 2004), the presence of specific actors such as groups, organizations and communities (Galaway and Hudson, 1994), or the existence of inter-firm dynamics and knowledge spillover (Malecki and Oinas, 1999; Porter, 1990). In particular, researchers in the field of innovation studies describe how certain regions have managed to develop local systems of innovation by combining these factors in particular ways (Cooke et al., 2004). Also, numerous case studies have described how such factors can induce employment growth at the local level. Although the nature of such studies does not enable extrapolation to general conclusions (Markusen, 1999), they have been used to determine best practices that can serve as a basis for implementing policies in other regions.

The second group of determinants is comprised by factors that are also local in nature, but that can be quantitatively measured and are based upon general economic growth theory. The area's endowment of human capital (Florida, 2002; Romer, 1990), its industrial structure (Porter, 1998), its relative costs (Weber, 1929), or its level of industrial diversity (Jacobs, 1984; Porter, 1990) have all been theorized to be solid growth determinants. The effect of such factors upon employment growth, either at national and subnational level, can be verified by statistical analyses of various sorts. Usually a large group of regions or cities is considered, and either the effect of each factor upon the city or region's growth (Beckstead and Brown, 2003; Florida, 2002; Shearmur and Polèse, 2007) or upon the entire nations (Barro and Sala-i-Martin, 1995) is determined. From this type of analysis, some broad conclusions about *local* growth are typically drawn. Florida (2002), for instance, suggests that a highly educated local workforce is conducive to local growth of employment and income. Henderson (2003) shows that local specialization in certain industries tends to lead to employment growth in those industries, while Quigley (1998) shows that a diverse economy tends to be associated with higher growth in the region.

However, analysis on growth factors are not always conclusive and can even be contradictory. As Blien et al. (2006) points out, there is an ongoing debate between researchers who defend that a diverse economy leads to growth (Jacobs, 1984; Markusen, 1996; Quigley, 1998) and those who defend that specialization is conducive to growth

(Porter, 1996). Despite being to some extent compatible with human capital theory, Florida's (2002) results on the relationship between a regional educated workforce and regional growth is not supported by evidence – after controlling for other growth factors – in the Canadian case (Shearmur and Polèse, 2007). Thus, despite the theoretical possibility of deriving conclusions upon local growth from analysis that use regional data, empirical results applied to regions has proven elusive so far.

The third group is comprised of factors that are structural as they can be considered geographical and historical traits. Geographical location, in particular, the proximity to markets (Krugman, 1995; Partridge and Rickman, 2006), the center-periphery and urban-rural divide (Parr, 2001), and historical events or tracks, such as the US technological superiority (Krugman, 1995; Davis and Weinstein, 2002), have been suggested as having an effect upon employment growth outcomes.

Despite of the greater emphasis upon local economic development over the past twenty years (Martin and Sunley, 1998; Parr, 2001), policy makers have showed little interest in geographical and historical factors, which generally operate at greater scale than the local dimension (Eisinger, 1988; Keating, 1993). The failure in the 1960s and 1970s of top-down policies, which were aimed at balancing growth across wide geographical areas, explains only partially this temporary obscurity. Other factors include the fact that statistical models used to analyze geo-structural effects fell out of fashion over this period as they were claimed to be too simplistic – unable to deal with the qualitative factors that are also important for understanding economic development – and to make claims about development factors that did not bear out in practice (Philo et al., 1998). Moreover, these structural factors can be considered fixed in the short and medium run and are therefore unable to be substantially changed or affected by economic policies.

Ignoring these structural factors does not necessarily mean, however, that they do not affect local employment growth. One of those structural factors, the size of a region – commonly quantified in regional economics in terms of population –, classifies an area as rural or urban, and assists in determining the existence of agglomeration economies. Once regions are classified as rural or urban, a second dimension comes into play: the hierarchy within the urban system, as economic flows tend to favor larger cities (Parr, 2002). Regarding these two dimensions (center-periphery and agglomeration), several empirical studies show that employment growth tends to be spatially distributed following a non-random pattern. Coffey and Polèse (1988), Coffey and Shearmur (1996) and Polèse and Shearmur (2004) studied the distribution of employment growth across the urban system, and between central and peripheral areas for Canada, while the UK case was analyzed with a similar approach in Mackay (2003). The Spanish case was addressed in Polèse et al. (2006). These studies show how employment growth – particularly in strategic economic sectors such as high-order services – concentrate in and around cities, and more specifically in and around large metropolitan areas.

Regarding location, the proximity to areas with important interregional and international trade flows is deemed relevant, as well as the proximity to the coast, given that it facilitates access to ports, reducing transportation costs (Hummels, 1999). This idea is reinforced by the fact that airports in coastal cities are normally the most important gateways to international connections. The coast also includes spaces with a greater propensity to develop a standard type of tourism. Rappaport and Sachs (2003) studied the relevance of the coast in the US economy, finding clear correlations not only with population density, but also with productivity and growth.

As Polèse (2009) points out, not only absolute location matters since industries (and therefore economic activity and employment) are drawn to places best suited for commerce and interaction with markets. In this sense, proximity to size – relative location – also matters. Thus, spaces can be classified according to their size into rural and urban areas (regarding an urban hierarchy), and into center or periphery according to their proximity to major metropolitan areas. This classification is consistent with the classic landscape described in Christaller (1935), Lösch (1938) and Von Thünen (1826), with a large marketplace at the center (the metropolis) surrounded by some areas of smaller population size which, depending upon their respective distances to the metropolis, would be categorized as either central or peripheral.

The problem lies in how these spaces can be defined in terms of being located close or far from the system of metropolises. A way of solving this issue was proposed by Partridge et al. (2008 and 2009) based upon Christaller's ideas regarding the hierarchy of places, and the connection suggested by Zipf (1949) between urban size and the hierarchal position of the cities. If it is assumed that only large cities are able to offer a full range of goods and services, and only the distance to the closest large metropolitan area (the one ranking highest in Christaller's hierarchy) is considered, the fact that certain goods and services are also offered in smaller and nearer cities is neglected. One way to address this idea is to define a set of incremental distances to each urban tier in order to measure the relative location of an area regarding larger agglomerations. Individuals and businesses need access to the higher-order services, urban amenities, high-qualified employment/labor and low-cost products that, due to the presence of strong agglomeration economies, are only present in highly populated places.

3. The Spanish case: spatial units and data

Like in the previous part of this dissertation, a first important decision is the selection of the spatial unit of study. Many times this choice arises from the lack of alternatives or the impossibility to find a spatial unit consistent with the objective of the research, but the main concern here is that underestimating the link between the purposes of the research and the spatial unit chosen could end in meaningless results.

Provided that the aim is understanding what local factors determine employment growth, the notion of "local" should be carefully defined. According to Fox and Kumar (1965), Smart (1974) and Becattinni (1962, 1975) the concept of a local area could be associated with a space that internalize the home-to-work journeys of its residents. Several techniques have been developed to set the boundaries of the local areas based on administrative regional data from municipalities or districts. One of the most accepted methodologies was proposed by Sforzi et al. (1997), already presented in section 2.1 of Chapter 1 as a justification for the election of local labor markets (LLMs). The Spanish municipalities finally translate into 677 LLMs after some adjustments in the original algorithm (hence the difference between the number of LLMs in this chapter). It is worth recalling that a LLM encompasses within the same unit the vast majority of labor and income-related movements, being areas with high internal homogeneity and, at the same time, high external heterogeneity (Rubiera and Viñuela, 2012).

These travel-to-work areas have interesting characteristics for the purpose of our analysis. Recognizing that labor plays a basic role in the lives of individuals and guides their territorial behavior with regard to the election of the municipality where they live, LLMs have the advantage of including both locations (work and residence) within the same spatial unit. This is an important attribute when using databases at a very high level of disaggregation, as some information regarding the place of work can be censored for small municipalities due to confidentiality concerns. However, at municipal level the common figure provided by the Statistic National Institutes reflects the "municipality of residence of the employed population", but not the "municipality of work of the population residing in the municipality". Given these data restrictions, the use of LLMs as the spatial unit of analysis guarantees that the "employed population residing in the municipality" growth figures at local level will not greatly differ from the "employment growth" (Sforzi 2012). Moreover, a LLM describes a place that corresponds to the area where the local population develops most of its economic and social relationships, it is a place where the common interest of the local population can be identified as a whole, and therefore it could be considered an appropriate level for implementing policies at local or regional level (Parr 2008). The implementation of such policies requires some degree of municipal cooperation or coordination, something theoretically simple that, however, for political reasons usually confronts lots of resistance.

LLMs can be classified first according to size and then to distance to size. The first column of Table 6 presents the distribution of Spanish LLMs by population size in four tiers. The first two tiers, LLM_1 and LLM_2, correspond to the largest metropolitan centers. Given the substantial size gap between Madrid and Barcelona metropolitan areas and those classified as LLM_2 (with more than 500,000 but less than 2,500,000 inhabitants), it is appropriate to distinguish between the two levels. The next level (LLM_3) includes cities of more than

50,000 inhabitants but less than 250,000. Finally, those local labor markets with less than 50.000 inhabitants are considered rural areas (LLM_4).

Table 6. Distribution of LLMs by population size

Tiers	% population 1991	% population 2001	Category
LLM_1 > 2,500,000 inhabitants	20.59 %	20.48 %	**METROPOLITAN AREAS (MAs)**
2,500,000 inhabitants ≥ **LLM_2** > 500,000 inhabitants	16.76 %	20.68 %	
500,000 inhabitants ≥ **LLM_3** > 50,000 inhabitants	41.12 %	39.01 %	**URBAN AREAS (UAs)**
LLM_4 ≤ 50,000 inhabitants	21.53 %	19.83 %	**RURAL AREAS (RAs)**
TOTAL	38,871,359 inhabitants	40,847,385 inhabitants	

Source: Authors' own elaboration using Spanish Census data (INE 2007).

Despite the severe limitations on statistical information at high levels of spatial disaggregation, the Spanish Censuses, administered by the National Statistics Institute of Spain (INE, 1992, 2007 and 2012), provide information on several relevant economic variables that have been previously identified as local growth factors. Although there are partial updates every three years, a comprehensive database only becomes available every ten years.

Employment and population figures can be calculated for each of the 677 Spanish LLMs using the 1991, 2001 and 2011 Censuses. In this empirical proposal the dependent variable will be the logarithmic employment growth rate (G_{emp}) in each LLM, calculated as the difference between the logarithm of the employment in 2001 and the logarithm of the employment in 1991. The logarithm of the population size will be used in order to preserve consistency, as many other factors included are rates. The LLMs initial population size (*Pop*) and employment rate (*Emp*) are included as explanatory factors. Certainly, population and employment in a local area can have a bivariate causal relationship (Freeman 2001), but in this case the objective is to explore the effect of initial population on employment growth.

Another interesting variable provided by the Census is the municipal socioeconomic condition (*SC*) index, which combines information from the labor market status (employed, unemployed and inactive) with the professional classification to build a measure of the average social position, that can be interpreted as an indicator of the living standards of the area. The censuses also offer data regarding the highest level of education reached by the resident population at the local level. Thus, in order to capture the effect of human capital endowment in the LLM, the percentage of the population living in a particular LLM in 1991 and 2001 who holds a secondary and a tertiary degree (*SecEd* and *TerEd*) is accounted for. Other demographic information on the area that can be valuable for this analysis is any aspect associated to the economic dynamism, such as the initial percentages of foreign population (*Fpop*), the proportion of young people (less than 16 years of age)(*Ypop*) and old people (over 65 years of age) (*Opop*).

The economic structure of the area could also be another factor explaining employment growth. Data regarding employment by sector is also available from the censuses and location quotients (*LQ*) can be calculated for each LLM and sector. The location quotients compare the employment share in one sector in a LLM with the equivalent share at national level, and therefore provide information about the specific sectors in which a LLMs is specialized. In this case, the agricultural, industrial and public services location quotients can also provide information about the rural or urban nature of the region. Moreover, following Shearmur and Polèse (2005) the degree of specialization/diversity of the LLM in the base year (*S*) is also included and calculated as the following specialization index:

$$S_r = Ln\left[\frac{1}{\sum_{i=1}^{n} Emp_i}\sqrt{\sum_{i=1}^{n}\left(Emp_i(LQ_i - 100)\right)^2}\right] \qquad [9]$$

where S_r is the specialization index for area r; LQ_i is the location quotient of sector i for area r; and E_i is the employment in sector i for area r. Accordingly, values tend to $-\infty$ when the degree of specialization in the LLM is identical to that of the whole national economy; the value tends to $+\infty$ as the specialization profile of the LLM diverges from the overall Spanish specialization (i.e., when the LLM is more specialized in one or more of the n sectors analyzed).

The linear distance to the nearest coast (*DistCoast*) and some climate related variables are also included as geo-structural factors. The distance to the coast reflects the proximity to a location that historically fosters population and employment growth, and is available from the Spanish National Geographical Institute (IGN, 2001). Some climatic features are also taken into account: the average annual rainfall (*Rain*), the minimum temperature in January ($Tmin_{jan}$) and the maximum temperature in July ($Tmax_{jul}$). These factors depict the impact of natural amenities, and can influence the attractiveness to certain economic activities. The climatological information introduced comes from the series (1987-2007) published by the Meteorological State Agency (AEMET, 2011)

As a political factor, a dummy variable *Cap* is also included, taking the value 1 if the LLM contains an administrative capital of a province or an autonomous community, and the value 0 otherwise. This variable captures the influence of having been identified as the administrative center of the province, thereby concentrating a large proportion of public sector jobs and offering a large range of public services.

All the variables available at LLM level that will be used in the empirical model are summarized in Table 7.

Table 7. Variables and data

VARIABLE	TERM	VARIABLE DEFINITION	DATA SOURCE
LOCAL EMPLOYMENT GROWTH	G_{emp}	Logarithmic Employment growth between 1991-2001 and 2001-2011)	1991, 2001 and 2011 Spanish Censuses. National Statistics Institute (INE).
DEMO-ECONOMIC FACTORS			
Initial employment	Emp	Employment rate	1991 and 2001 Spanish Censuses. National Statistics Institute (INE).
Initial population	Pop	Population size (ln)	
Socioeconomic condition	SC	Index of the average labor market status of the population	
Human Capital Endowment	SecEd TerEd	percentage of resident population with a secondary/tertiary degree	
Specialization/ Diversification	S S^2	Specialization index	
Potential economic performance	Fpop Ypop Opop	Percentage of foreign population	
		Percentage of young population	
		Percentage of old population	
Type of Specialization	LQ_{agr} LQ_{ind} LQ_{PS}	Location quotients for agricultural and industrial activities, and public services	
GEO-STRUCTURAL FACTORS			
Distance to the Coast	DistCoast	Dummy variable. 1 if the LLM contains a coastal municipality	National Geographical Institute (IGN).
Natural amenities	Rain	Average rainfall between 1987 and 2001	Meteorological State Agency (AEMET)
	$Tmax_{jul}$	Average higher temperature reached in July between 1987 and 2001	
	$Tmin_{jan}$	Average lower temperature reached in January between 1987 and 2001	
Relative Location	ID_{LLM1} ID_{LLM2} ID_{LLM3}	Incremental Distances from a LLM to the nearest LLM in each higher tier	National Geographical Institute (IGN)
POLITICAL FACTORS			
Capital city	Cap	Dummy variable. 1 if the LLM contains a provincial or Autonomous Community capital city	

4. Econometric approach: spatial Durbin model

The concept of local-labor-markets applied to the Spanish case results on 677 LLMs, that can be classified in terms of size and location (using the concept of incremental distances) as described in Table 6. Once the variables outlined in the literature as growth determinants are calculated al LLM level from municipal data, the equation to be estimated would take the following form:

$$\begin{aligned} G_{emp} = \alpha &+ \beta_1 Emp + \beta_2 Pop + \beta_3 SC + \beta_4 SecEd + \beta_5 TerEd \\ &+ \beta_6 S + \beta_7 S^2 + \beta_8 Fpop + \beta_9 Ypop + \beta_{10} Opop \\ &+ \beta_{10} LQ_{agr} + \beta_{11} LQ_{ind} + \beta_{12} LQ_{PS} + \gamma_1 DistCoast \\ &+ \gamma_2 Rain + \gamma_3 Tmax_{jul} + \gamma_4 Tmin_{jan} + \gamma_5 ID_{LLM1} \\ &+ \gamma_6 ID_{LLM2} + \gamma_7 ID_{LLM3} + \zeta_1 Cap + \varepsilon \end{aligned} \quad [10]$$

However, there is increasing evidence on the fact that economic growth tend to be spatially correlated, meaning also that employment growth is not randomly distributed among regions. Figures 16 and 17 suggest that employment growth in Spain is not different in this respect, since certain clustering can be identified at a glance for both of the decades considered.

In order to confirm this conjecture and formally test for the existence of spatial association in the data, the Moran scatterplot is shown (Figure 18) to check for the presence of a global spatial pattern, taking into account also the local magnitude of the spatial autocorrelation through a local indicator of spatial association (LISA) (Anselin, 1994). In this analysis, the definition of neighborhood relies on a second order queen contiguity matrix (geographical neighbors plus their neighbors), which is comparatively the best fit to the distribution of employment growth in Spain in the decades regarded. Figure 19 shows the link between LLMs' employment growth in each period and employment growth of the neighbor LLMs The slope indicates a positive relationship, being less pronounced in the second decade, which jointly with the lower value of the Moran's I statistic indicates a decrease in the level of global association. The LISA maps of Figure 19 give support to the spatial clustering suggested earlier, showing for the 1991-2001 decade some groups of high growth LLMs across the Mediterranean Coast and the islands, and low growth LLMs in particularly in the North- West quarter of the country; meanwhile, the 2001-2011 decade shows high growth clusters concentrated in Madrid Metropolitan Area and surrounding LLMs as well as in the Ebro Corridor (i.e., from Barcelona to Zaragoza Metropolitan Areas).

In order to explain employment growth at local level in Spain, a methodology that accounts for the spatial concentration and allows the inclusion of the characteristics of the neighbors is a must. A review of the various methods for addressing spatial processes can be found in the seminal works from Anselin (1988), Cliff and Ord (1981), Griffith (1988, 2003), Haining (1990), Anselin et al. (2004) and LeSage and Pace (2009). This papers contains several options to deal with spatial autocorrelation, being the most traditional ones:

(i) Spatial lag model (also called spatial autoregressive model (SAR)), considering the influence of the dependent variable of neighboring areas.

(ii) Spatial error model (SEM), introducing a spatial auto-regressive component which allows for the existence of spatial dependence in the error term.

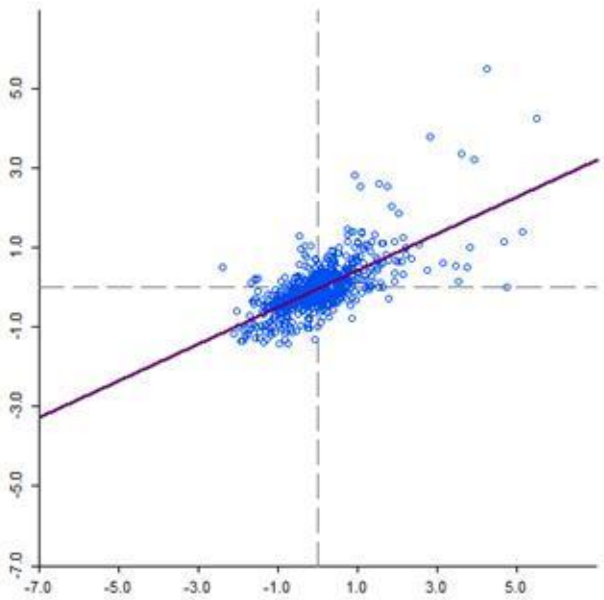
(iii) Spatial Durbin model (SDM), accounting for spatial dependence through the dependent and the independent variables simultaneously, trying to cope with the issue of omitted variables that might be behind spatially correlated errors.

The SDM, as advocated in the work of LeSage and Fischer (2008), is an appropriate specification for growth studies because it clarifies when the spatial autocorrelation is due to the clustering of high or low employment growth areas by means of a spatially lagged dependent variable. Moreover, this model tries to uncover the source of possible bias in the residuals related to the features of the neighbors through the inclusion of lagged explanatory factors. This kind of models are meant to address specifically spillover effects, which is the main concern of this type of analyses, while others focus on the disturbances.

Figure 18. Moran scatterplots for employment growth by LLMs between 1991 and 2001, and 2001 and 2011

1991-2001

Lag employment growth 1991-2001

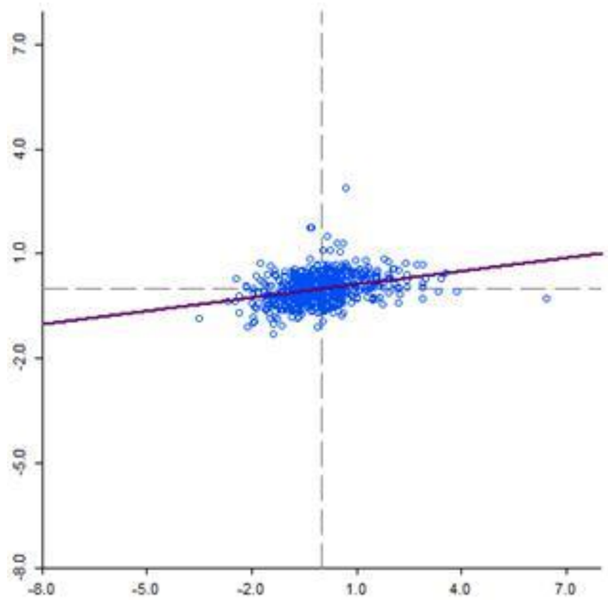


Employment growth 1991-2001

Global Moran's I: 0.4646

2001-2011

Lag employment growth 2001-2011

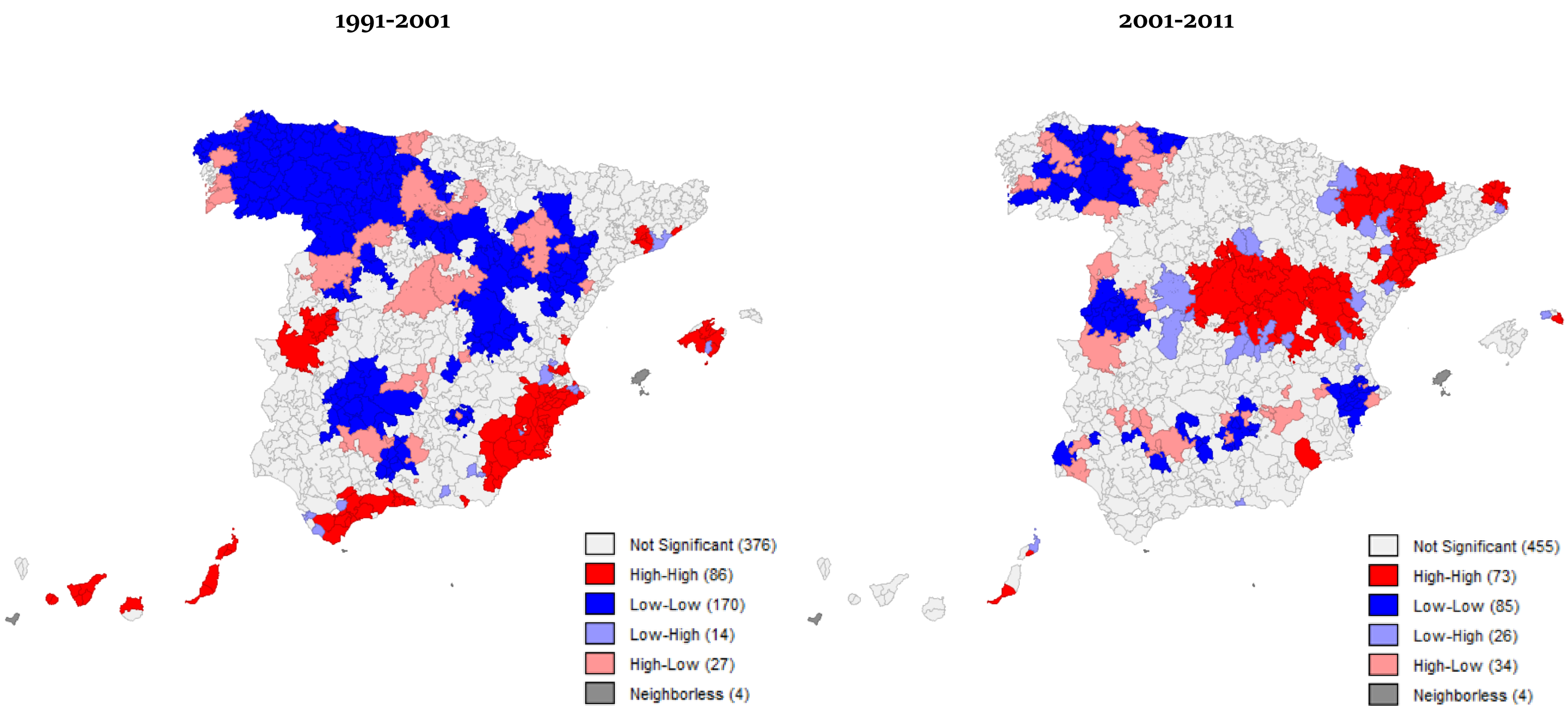


Employment growth 2001-2011

Global Moran's I: 0.1282

Source: Own elaboration with data from the National Geographical Institute (IGN) and the National Statistics Institute (INE).

Figure 19. LISA cluster maps of employment growth in Spanish LLMs between 1991 and 2001, and 2001 and 2011



Source: Own elaboration with data from the National Geographical Institute (IGN) and the National Statistics Institute (INE).

The general formulation of SDM is:

$$y = \alpha + \rho W y + \beta X + \gamma W X + \varepsilon$$
$$\varepsilon \sim N(0, \sigma^2) \qquad [11]$$

Where in our case y is the employment growth (G_{emp}) experienced in each Spanish LLMs and the neighbors' employment growth is considered through the term Wy. The matrix containing the remaining explanatory variables related to the socio-economic, geographical and political characteristics of the LLM is named *X* and the term *WX* embraces the neighbors' characteristics to the model.

The parameters will be estimated through maximum likelihood, but their interpretation is subject to the feedback effects of the explanatory variables between a region and its neighbors, induced by the spatially lagged factors. This means that the change in a factor will have a direct effect on the employment growth of the own region, but also an indirect effect on the neighbors, implying that the marginal effects are not represented by the β coefficients as in linear models, and need to be computed after the estimation of the parameters in order to draw correct conclusions about the relationship between employment growth and the rest of the variables.

Following the procedure developed by LeSage and Pace (2009), the figures that represent the actual effects of the factors in the SDM, and that will be interpreted in this study, can be summarized as:

(i) Average direct impact: mean of the impacts on the dependent variable in region i of a change in the *rth* factor of that same region. Technically, it equals the mean of the diagonal elements ($\partial y_i / \partial x_i^r$) of the partial derivatives matrix.

(ii) Average indirect impact: mean of the impacts on the dependent variable in region i of a change in the *rth* factor of the other regions (j≠i). It translates into de mean of the off-diagonal elements ($\partial y_i / \partial x_j^r$) of the matrix mentioned before.

(iii) Average total impact: sum of the average direct impact and the average indirect impact.

5. Results: particularities of the Spanish case

In this section the empirical analysis of local employment growth is undertaken by applying a spatial approach through the SDM, given the spatial dependence found in the preliminary section pointing out to the existence of a functional relationship between processes occurring in a LLM and those occurring in neighboring LLMs. First, the linear estimation of the model will be briefly discussed as an initial step to diagnose spatial autocorrelation. Then, the spatial bias will be dealt by means of the SDM, and finally the marginal effects of this specification will be presented, describing first the impacts in 1991-2001 and 2001-2011 separately, and then comparing both periods.

The ordinary least squares (OLS) estimation of the empirical model follows the specification shown in Equation 10. The Moran's I test indicates that the null hypothesis of random spatial distribution of the dependent variable should be rejected in both decades (0.08 and 0.06 respectively, being both values significant at 0.01% level), corroborating the existence of spatial autocorrelation in the residuals and giving further statistical support to the use of spatial econometric models in Equation 3. Under this circumstances, the lack of parameters that bring the spatial structure into the model deems the coefficients obtained in the OLS estimation as biased.

According to the suggestion made by LeSage and Fischer (2008), the specification of the employment growth model should include not only the neighbor's employment growth but also their characteristics, leading to the formulation of a SDM, which estimation is shown in Table 8. As noted in the previous section, the coefficients obtained at this stage should not be directly interpreted as marginal impacts of the factors on the dependent variable. Nonetheless, this first estimation provides useful information about the degree of spatial association, the bias correction achieved through the inclusion of the spatially lagged factors, and the relative performance of the model. For this reason, the preliminary results before adjusting the marginal effects are reported and discussed.

Table 8. Spatial Durbin Model estimation results, 1991-2001 and 2001-2011

	SDM 1991-2001		SDM 2001-2011	
	Factor	*Spatial Lag*	*Factor*	*Spatial Lag*
Constant	0.247	–	-0.523	–
Rho	0.319 ***	–	0.108	–
Initial Employment rate	-1.560 ***	-1.018	-0.594 ***	2.172 ***
Population (ln)	-0.051 ***	-0.032	-0.030 ***	0.079 ***
ID_{LLM1}	-0.011	0.024 **	-0.011	0.004
ID_{LLM2}	-0.026	0.040	-0.012	0.026
ID_{LLM3}	-0.164 ***	0.016	-0.060	0.368 ***
Socioec. condition index	0.745 ***	1.223 ***	-0.002	-0.997 ***
Sec. education (%)	-0.074	-0.253	0.202 ***	-0.412
Ter. education (%)	1.390 ***	2.400	0.826 ***	2.352 ***
Foreign pop. (%)	1.343 ***	-2.721 ***	0.095	-0.506
Pop. under 16 years of age (%)	3.922 ***	1.328	0.097	2.593
Pop. over 65 years of age (%)	-0.830 ***	1.572	-0.188	1.710 **
Specialisation index	0.089 ***	-0.050	0.015	-0.144
Specialisation index (square)	-0.014 ***	0.006	-0.005	0.015
LQ agriculture	-0.0002 ***	-0.0002	-0.0002 ***	0.000
LQ industry	-0.001 ***	0.001	-0.001 ***	0.000
LQ Public services	-0.002 ***	-0.001	-0.0004	-0.004 ***
Dist. Coast	-0.040	0.098 ***	-0.022	-0.017
Rain	-0.013 **	-0.007	-0.006	0.003
Temp. max. July	-0.003	-0.007	0.012 **	-0.012
Temp. min. January	0.003	0.018 ***	-0.0003	0.015 **
Capital city (1/0)	0.015	-0.054	-0.031	-0.025
Akaike Information Criterion (AIC)	*-660.05*		*-679.49*	
Log Likelihood	*375.02*		*384.74*	
Likelihood Ratio Test	*10.85 ****		*0.96*	
LM Test for residual sp. autocorrel.	*0.15*		*2.74 **	

The rho coefficient is low for both decades, but it is only significant for the first one. It is worth recalling that the analogous linear models for each decade present spatially correlated error terms in both periods, showing also a lower degree of spatial association between 2001 and 2011. Interpreting both results jointly, it seems that the model deals with the spatial autocorrelation better in the 1991-2001 period, accounting entirely for it, as indicated by the non-rejection of the null hypothesis (absence of spatial autocorrelation in the residuals) of the Lagrange Multiplier test. Interestingly, the Akaike information criterion and the log likelihood say that the specification explains better the employment growth during the second decade, despite there are still traces of spatial bias.

To find the marginal effects of the spatial spillover specifications, the methodology described by LeSage and Pace (2009) was followed, as mentioned in Section 4. Table 9 shows the direct, indirect and total effects derived from the SDM. Direct effects represent the impacts of the explanatory factors on employment growth, while indirect effects are the accumulative spillover impacts coming from the neighboring regions. Both impacts together result in the total effect of a given factor on the dependent variable.

Table 9. Direct, indirect and total effects, 1991-2001 and 2001-2011

	SDM 1991-2001			SDM 2001-2011		
	Direct	*Indirect*	*Total*	*Direct*	*Indirect*	*Total*
Initial employment rate	-1.593 ***	-2.195 **	-3.788 ***	-0.581 **	2.349 ***	1.768 **
Population (ln)	-0.053 ***	-0.071	-0.123 **	-0.030 ***	0.085 **	0.055
ID_{LLM1}	-0.010	0.030 **	0.020 **	-0.011	0.003	-0.008
ID_{LLM2}	-0.025	0.046	0.021	-0.011	0.028	0.016
ID_{LLM3}	-0.165 ***	-0.052	-0.218	-0.058 *	0.403 ***	0.345 ***
Socioec. condition index	0.777 ***	2.115 ***	2.892 ***	-0.009	-1.111 **	-1.120 **
Sec. education (%)	-0.080	-0.400	-0.480	0.199 *	-0.434	-0.235
Ter. education (%)	1.452 ***	4.116 *	5.569 **	0.841 **	2.720 **	3.561 ***
Foreign pop. (%)	1.293 ***	-3.318 **	-2.024	0.092	-0.552	-0.461
Pop. under 16 years of age (%)	3.979 ***	3.735	7.714 ***	0.113	2.901	3.014
Pop. over 65 years of age (%)	-0.801 **	1.891	1.090	-0.178	1.883 *	1.705 *
Specialisation index	0.089 **	-0.031	0.058	0.014	-0.159	-0.145
Specialisation index (Square)	-0.014 ***	0.003	-0.012	-0.005	0.016	0.011
LQ agriculture	-0.0002 **	-0.0004	-0.001 *	-0.0002 **	-0.0002	-0.0004 *
LQ industry	-0.001 ***	0.0004	-0.001	-0.001 ***	-0.0001	-0.001 ***
LQ Public services	-0.002 ***	-0.002	-0.004 *	-0.0005	-0.004 ***	-0.005 ***
Dist. Coast	-0.038	0.123 **	0.085 ***	-0.023	-0.022	-0.044 *
Rain	-0.013 *	-0.016	-0.029 ***	-0.006	0.002	-0.003
Temp. max. July	-0.003	-0.011	-0.015 *	0.012 *	-0.012	0.001
Temp. min. January	0.003	0.027 ***	0.030 ***	-0.0002	0.017	0.016 **
Capital city (1/0)	0.014	-0.072	-0.058	-0.031	-0.032	-0.063

Taking into account the socioeconomic factors, in the first decade (1991-2001) the initial employment rate has a negative total impact on the employment growth, driven mainly by the influence of neighboring regions (indirect effect). In magnitudes, a 1% increase in the initial employment rate translates into a 3.79% decrease in employment growth of which 2.19% comes from changes in the initial employment rate of the neighbors. The negative effect is also at work when the own initial population size is regarded, bearing a fall of

0.12%. This result, in combination to the previous one, might be a hint to the existence of agglomeration diseconomies that in turn could be signaling a convergence effect of smaller areas in terms of employment in the first decade.

However, when looking at the relative location of the LLM within the urban hierarchy (ID_{LLM1}, ID_{LLM2} and ID_{LLM3}), results show that the incremental distance to the big metropolitan areas (Madrid and Barcelona LLMs) exerts a positive indirect effect over a LLM's employment growth and overall a positive total impact. Remarkably, the direct effect of the distance to the third urban tier (defined as LLMs between 500,000 and 50,000 inhabitants) is significant and negative (-0.16%), which might suggest that LLMs included in this category and also those ones relatively close to medium size urban areas and relatively far from the big metropolitan areas were benefitting from higher employment growth.

On the other side, the socioeconomic index (2.89%), the initial capital endowment, measured by the proportion of population with tertiary education attainment, (5.57%) and the share of young population (7.71%) have a relatively large positive total impact, driven mainly by the change in the neighbors' features of the two first factors (2.11% and 4.11% respectively). The shares of foreign and older population, the specialization index variables and the location quotient for industry have a significant direct impact, negative in the case of the older population and the industry indicator, while the indirect and total effects are not significant. The location quotients of agriculture and public services have significant negative direct and total effects. Regarding the sectorial structure, the employment concentration on the sectors included lessens local employment growth.

With respect to the geographical features, the average rain rate is the only factor showing a significant direct effect, with a negative total impact of 0.03% on employment growth. Conversely, the distance to the coast and the minimum temperature in January of the neighbors have a positive influence. The maximum temperature on July doesn't display a significant direct or indirect effect, but its total effect bears a decreasing impact of 0.01%.

In regard to the second decade[10], while the estimates for the initial employment rate and population of the LLM display a negative direct impact (0.58% and 0.03% decrease respectively), the indirect effect and as a consequence the total effect, is positive. On the other side, the spillover effects are positive and larger in value, contributing to employment growth at 2.35 % and 0.08% respectively. Only the initial employment rate has a significant total effect, bearing a 1.77% increase in employment growth. Thus, in the second decade local growth seems to be concentrating in those areas that already were enjoying high employment rate figures, suggesting a distinct path towards economic divergence.

However, the relative location of the LLM in the urban hierarchy that LLMs near to small-medium sized urban areas, in combination with the effect of having neighbors spotted relatively far from them, result in a positive total effect of 0.34% over employment growth.

The average socioeconomic condition index has negative indirect and total impacts in this second decade, while the secondary and tertiary education have significant and positive effects. A 1% growth in secondary education attainment in the own region is responsible for a 0.19% increase on employment growth, a 0.84% in the case of tertiary education, which also has an indirect effect of 2.62% and a remarkable total effect of 3.56%. The increase in the share of older population in the neighbors has a positive effect of 1.88%, being the total impact of the factor 1.70%. The location quotients for agriculture, manufacturing and public services have a significant and negative total effect over local employment growth, but the magnitude of effect is almost irrelevant.

Geo-structural factors reflect that, overall, proximity of a LLM to the coast and warmer weather conditions enhance employment growth in this period.

[10] It is important to note that the current crisis hit Spain in 2008, affecting deeply the economic scenario. As an example, the change in the unemployment rate went from 10.94% in the first quarter of 2001, to 21.08% in the first quarter of 2011. Certainly, this fact affects the results of this study provided the lack of suitable local data for inter-census years, preventing us from including such an important shock in the analysis. In this situation, caution is due in the interpretation of the effects during the second period. Nevertheless, the framework depicted so far allows us to set a comparative scope between the two periods considered, and we believe that it adds to the understanding of the general picture of local employment growth in the country.

After applying the same model to both decades and comparing results, it can be seen that the Spanish local employment growth is sensitive to the period of analysis. Direct impacts are clearly dominant during the first decade, while the number of significant direct and indirect effects evens out in the second one. In other words, the spillover effects gain relative strength over local employment growth in the second period. Probably the inclusion of more features (own and spatially lagged) might reduce the remaining spatial autocorrelation.

The direct negative influence of initial employment rate and population is still operative in the second decade, but the indirect and total impacts turn positive in the second decade, implying that being large both in terms of initial employment and population deters employment growth. Unlike the first decade, being located close to other LLMs with a high score on those features is an advantage in the second decade. Thus, local employment growth in the second decade seems to be closely related to the proximity to size (or relative location of the LLM to the big metropolitan areas), while in the first decade both size and relative location exerted a negative effect over employment growth, suggesting that agglomeration diseconomies were in place during that decade.

A similar change is observed also in the socioeconomic condition index, which indirect and total effects become negative for the second decade, indicating that an increase in the socioeconomic condition of the neighbors hinders own employment growth, opposite result than the one obtained for the previous decade. Although significant, the direct and indirect effect of the LLM's initial level of population with at least tertiary education diminishes in magnitude for the second decade. This circumstance also affects the shares of foreign and young population, which effects become insignificant in this decade. On the contrary, the positive spillover and total effects of the older population share become significant. The own sectorial specialization no longer affect local employment growth while the employment concentration in public services of the neighbors have a negative impact on it. In general, the variables related to natural amenities lose their influence in the second decade. Regarding the location with respect to the urban hierarchy, the impact of the relative proximity to the metropolitan LLMs existing in the first period, disappears in the next decade, while the indirect effect of the distance to the small-medium areas gain significant weight in the positive total outcome.

6. Summary and main conclusions

Economic growth theories collide with the ambiguous empirical results obtained when working at local level. The conclusions about factors that explain local employment growth differences tend to be indefinite – and at times opposed –, as they rely on the period of analysis, the level of disaggregation, the data limitations and the empirical approach embraced. Our results for Spain confirm how sensitive the empirical analysis is, and the direct and/or indirect effect of some structural and socioeconomic factors over local employment growth may vary across time.

Given the regional and intertemporal heterogeneity observed in the Spanish case, this chapter tried to identify some patterns in the distribution of local employment growth for two decades with persistent growth: 1991-2001 and 2001-2011. The empirical approach considered takes into consideration socioeconomic and structural factors, as well as the regional context and their spatial interactions. Dealing with spatial autocorrelation and disaggregated local information, an empirical model was proposed with the intention of assessing the contribution of the LLM's own and neighboring characteristics over employment growth for each decade.

Some of the local determinants included in this study show the expected positive effect over local employment growth and in line with endogenous growth theories, but their magnitude varies along time. Both your own initial human capital endowment (measured as percentage of population with tertiary education) and your neighbors exerts a significant and positive effect over local employment growth in both decades,; however, the overall effect is greater for 1991-2001, before the Spanish economy's growth relied extensively on the construction sector.

Nevertheless, some other determinants show the opposite (and significant) effect depending on the period under consideration, which requires a combined analysis for both decades. Results for the variables related to initial size and relative location of the Spanish

LLMs show that local employment growth in the second decade is not apparently linked to size, but it is indeed affected by the proximity to size. Surprisingly, in the first decade both size and proximity exert an adverse effect over employment growth. At first sight these results could suggest that during the two decades under analysis the Spanish cities were not necessarily the ones with better employment opportunities, or exhibiting a higher economic dynamism. Working with local (instead of regional) data allows to uncover the existence of agglomeration diseconomies (significant and negative effect of own initial population size), but the use of incremental distances simultaneously unveils what urban size tier is yet relevant. In the 1990s, LLMs far away from a large metropolitan areas (those LLMs with more than 2.5 million inhabitants) see their employment growth affected in a positive way. In the 2000s, the longer the distance to a small-sized urban area (LLMs from 500,000 to 50,000 inhabitants), the higher local employment growth. Size by itself is not the driver of local employment growth, but the proximity to a certain urban tier.

So what were the local factor explaining local growth in these two decades?

While in the second decade local employment growth was positively and strongly affected by the initial levels of employment and labor conditions of the neighboring LLMs – but not by those of your own –, in the first decade the impact of the initial employment levels was significant and negative.

The aforementioned irrelevance of size in both decades, together with the impact of the initial levels of employment, suggest a differentiated local growth pattern for each decade. During the first decade local employment growth was positively affected by the proximity to small-medium sized urban areas that might benefit from the relocation of economic activities away from the influence of large neighbors. Moreover, according to the results, during the first decade the larger effect over local employment growth was exerted by the existence of low initial employment rates, an outcome in line with the suburbanization of the manufacturing industry to areas with lower labor and land cost. In the second decade, although the LLM's own initial employment level and (population) size continues to show a negative effect, the positive indirect effects imply that, in regards to growth, it is convenient to be located close to large urban areas and to LLMs with a high employment levels.

Results about the variable reflecting the LLM's standards of living are puzzling – The factor is significant in both decades, but shows a negative effect on employment growth in the second decade –, although it might be justified by the massive arrival of immigrant workers to Spain during the 2000s (more than 5 million in just a decade) and their location choices, a question which merits further research. The two decades under study reflect two very different patterns of local growth. While in the first decade the economy grew at a remarkably high annual rate, in the second half of the second decade the deep economic crisis fostered higher internal migration, higher international emigration (including return flows), and overall population loss.

There are not universal recipes for local growth and local policies must be designed paying attention to local characteristics. Besides considering the specific characteristics of the area, any effective local policies should also take into account the temporal dimension, as the influence of any factor over employment growth may vary (even being the opposite) depending on the time frame. This result reinforces the flexible framework set by the new European Union Cohesion Policy and the smart specialization philosophy, outlining the necessity of development policy projects tailored at more local level and revised in shorter periods of time.

CHAPTER IV. MOVING ALONG THE SPANISH URBAN HIERARCHY IN THE CRISIS CONTEXT: SOME FINDINGS ON THE URBAN WAGE PREMIUM

1. Introduction

The urban wage premium is defined as the difference between the wages of cities regarding their size, being the gap wider when comparing larger urban areas to other locations. This gap has been calculated at around 30% for the U.S. (Glaeser and Maré, 2001; Gould, 2007). For Spain between 2004 and 2009, the study from De la Roca and Puga (2016) revealed that there was a wage premium of 5% associated with doubling the size of the city where an individual works. Following Yankow (2006), the urban wage premium can be decomposed into static (price effect, improved firm productivity and ability sorting) and dynamic effect (enhanced human capital accumulation and job-worker matching), each of which accounts for half of the gap in the case of Spain, as De la Roca and Puga (2016) have estimated.

However, despite the urban wage premium has been widely investigated from a theoretical and empirical perspective, the question of how the urban wage premium reacts to a shock like the current crisis has not been addressed in the existing literature. More concretely for the Spanish case, the crisis changed the economic conditions and the migration patterns, affecting the labor supply and consequently the wage-setting scheme.

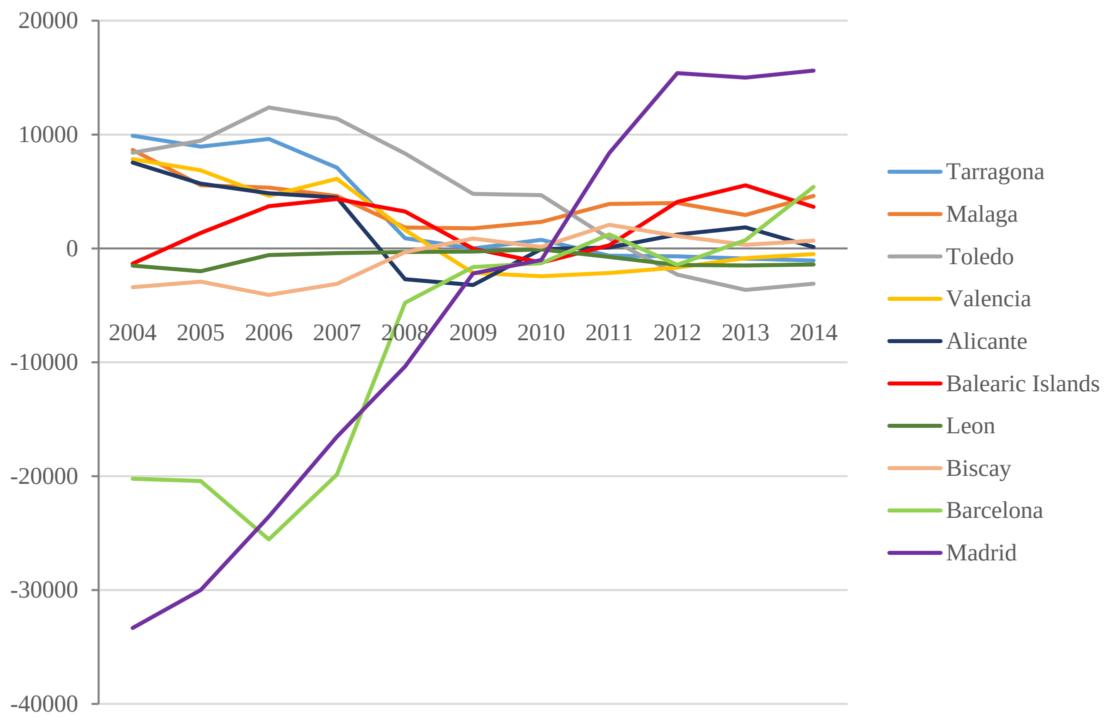
According to the Spanish Annual Wage Structure Survey (Table 10), wages in Spain grew at a 2.73% per year between 2004 and 2007. From the beginning of the crisis, wages remain fairly stable (0.61% growth annually between 2008 and 2013, 0.17% if 2008 is removed from the interval) in contrast to the previous pattern of continuous increase. Moreover, this change differs substantially across sectors. Wages in the financial intermediation sector were growing at a 1.60% annually, while during the years of the crisis they decrease at a rate of -0.82%. Another examples are the hospitality sector (1.64% and -0.13%) and educational activities (5.01% and 0.83%). The shift in the economic situation is reflected also in the fast growth of the unemployment rate (from a 7.93% in the third quarter of 2007 to a 17.24% a year and a half later, reaching its peak in the first quarter of 2013 at 26,94%). Taking into account that economic activities are not located homogenously across the territory and that the crisis hit some sectors harder than others, an initial hypothesis is that this change in wages might arise at the city level with different intensities, affecting the urban wage premium.

Table 10. Annual average wage growth by sectors

	2004-2013	2004-2007	2009-2013
TOTAL	*2.17*	*2.73*	*0.17*
Mining and other extractive industries	3.12	2.20	3.02
Manufacturing	2.33	2.40	1.73
Electricity, gas and water supply	1.35	-0.05	1.05
Construction	2.81	2.06	0.85
Trade and Repairing of motor vehicles	1.92	2.59	-0.44
Hospitality services	0.55	1.64	-0.21
Transportation and communication activities	1.10	2.08	1.17
Financial intermediation	0.82	1.60	-1.10
Real estate activities	0.91	2.34	-0.63
Education	2.20	5.01	0.22
Health and social work	1.96	2.99	-0.65
Other community and personal services	0.83	3.40	1.07

The notion of a change in the wage premium from a regional perspective induced by the crisis rises the point of a possible change also in the relative attractiveness of certain cities. If the wage gap between larger metropolitan areas and the rest of the urban structure changes, the incentives to move of the workers might change too. The current evidence with respect to provincial net migration reveals a shift in the in-outmigration balance. Figure 20 shows the five Spanish provinces (NUTS 3) with the highest net migratory balance and the five with the lowest one in 2004. By the end of the period 2004-2014, most of provinces have switched sides regarding the horizontal axis, a shift that happens between 2009 and 2011, right after the onset of the crisis, and by the end of the graph the dispersion of the series is lower than at the beginning. These dynamics show that the crisis affected the locational behavior of Spanish population, and suggests that there was actually a change in the existing balance of the urban wage premium for the different tiers in the Spanish urban hierarchy.

Figure 20. Provincial net migration by year



In view of the aforementioned events, this chapter aims to analyze the possible changes in the urban wage premium among the different tiers of the Spanish urban hierarchy as a response to the crisis. With the previous framework in mind, panel data techniques will be used in order to disentangle the drivers behind the urban wage premium and how the crisis might have affected them.

To distinguish the impacts on the urban wage premium before and during the crisis, data from the Continuous Sample of Employment Histories (MCVL in Spanish), released by the Ministry of Employment and Social Security of Spain, will be used. This database contains comprehensive information about the labor records of more than a million individuals, including their spatial location at a highly disaggregated level (municipalities with more than 40,000 inhabitants), which enables the identification of city-size effects. Moreover, it provides individual characteristics of the population as the Social Security records are complemented with data from the Municipal Registry.

In the next Section of this chapter the urban wage premium and the existing evidence on this phenomenon will be presented. Then, Section 3 will show the characteristics of the database and the variables that will appear in the empirical models that will be estimated through panel techniques in order to observe the changes in the differential wage. Section 4 includes the results obtained through the analysis. Finally, the main conclusions are summarized in Section 5.

2. Urban wage premium

Understanding and measuring the positive effects of agglomeration observed in large cities are fundamental issues in Regional Science, more specifically in the urban economics field. One of the clearest signs of the benefits of agglomeration is the enhanced productivity that firms located in large areas exhibit as a consequence of scale economies (Henderson, 2003).

Another effect that stands out is that the average wages in large cities are higher than those of smaller ones. This gap receives the name of urban wage premium in the literature, and it account for around a 30% percent difference in earnings in the U.S. (Glaeser and Maré, 2001; Gould, 2007), between a 15% and 60% in France depending on the urban tiers compared (Combes et al., 2008). After taking into account some basic characteristics, the estimations of this gap for Britain lies at 14% (D'Costa and Overman, 2014). For the Spanish case, there is little empirical evidence on local wage differentials due to the lack of data, but there are some authors that tried to explain regional wage differentials, providing with a valuable background for this analysis. Simón et al. (2006) found that there exist marked regional wage differences between workers with the same skills at the autonomous community (NUTS 2) level, induced by the collective bargaining system. These disparities increase or decrease following the upturns and drops of the business cycle, more intensively than in other European countries, as discussed by Bande et al. (2008). On the same line, Ramos et al. (2015) show that there are significant spatial spillover effects between the provincial (NUTS 3) wage curves, and that wage differentials are low but persistent in time as a consequence of the rigidities of the Spanish labor market. Finally, the main reference for this chapter is the paper by De la Roca and Puga (2016), as this represents the first analysis of the urban wage premium for Spain. Raw figures indicate a wage difference around 55% and 21% depending on the tiers of the urban hierarchy regarded, while for the period 2004-2009 the estimations show that relocating to a city that doubles the population size of the previous one translates into an increase in earnings of around 5%.

The urban wage premium is an interesting research subject since it raises the question about the rationale behind the higher wages of larger cities. The reasons regarded in the literature can be divided into static and dynamic effects, as reviewed in Yankow (2006) and in De la Roca and Puga (2016).

The static components of the urban wage premium are granted to the individual as soon as he/she starts working in a large city, and are lost when the worker moves back to a smaller area. Into this class falls the compensation that firms should offer because of the higher cost of living (price effect). If controlling for price differentials evens out the difference in wages, it can be said that the urban wage premium is a nominal effect. Another static component is the higher firm-level productivity stemming from the advantages of agglomeration, which translates into a higher marginal labor contribution. This effect has been widely studied, as shown by Ciccone and Hall (1996), Duranton and Puga (2004), Rosenthal and Strange (2004) or Holmes (2010), among others. The last component in this category is the sorting effect caused by the preference for the largest cities of more productive workers, which has an impact on the skill composition of the labor force of different areas. This sorting is based on unobserved abilities as observable characteristics are considered in previous steps of the estimation. The empirical evidence suggests that this effect accounts for a large share of the urban wage premium (Fuch, 1959; Glaeser and Maré, 2001; and Combes et al., 2008).

On the other hand, the dynamic effects are related to the interaction of workers and firms in the urban setting. Large cities offer a wider scenario for the creation and development of connections that become inherent to the worker, therefore the part of the wage premium associated to this dimension is only acquired through time, but it remains with the worker as she/he relocates. Included in this group is the learning effect, coming from the speeding up of human capital accumulation that takes place in cities, which becomes more valuable experience (Glaeser, 1999; Duranton and Puga, 2001); and the coordination effect, related to the improved job search and job-worker matching (Helsey and Strange, 1990; Sato, 2001).

For the Spanish case, De la Roca and Puga (2016) found that the static and dynamic components are balanced in terms of explanatory power regarding the urban wage premium, being more significant the effects of agglomeration on the static share, and the effects of learning in the dynamic one.

3. Data and methodology

To conduct the analysis we use the Continuous Sample of Working Histories (MCVL), a non-stratified random sample from the Social Security records, accounting for 4% of the population involved with the Social Security in a given year, including employees and people receiving unemployment or retirement benefits. In total, the sample contains information about more than one million individuals. For the regarded population, the dataset provides a full profile of each spell, and some characteristics of the employer. Moreover, the data includes a related file with some personal features such as age, sex, education, residence area or nationality.

The MCVL is available, upon request, in two versions: with and without tax data. Although both versions include comparable information, their identifiers are different, preventing their combined use. In this case, the version at hand is the MCVL with tax data, despite the fact that the tax information is not considered in this preliminary study.

Regarding the information related to wages that can be extracted from the sample, the contribution base for common contingencies works as an appropriate proxy in this analysis provided that it is closely associated to the gross earnings that the employee receives (excluding overtime). Its importance as an item of the final wage is officially acknowledged since it is the reference amount used to compute the retirement and unemployment benefits. There are other sources that provides information about wages in Spain, namely the Wage Structure Survey or the Labor Cost Survey, but in these cases the data is referred to the national or autonomous community (NUTS 2) level, leaving the MCVL as the only source that makes the analysis of wages at the city level feasible, as it allows to identify the municipality of work of the individual when the workplace is located in an area of more than 40,000 inhabitants.

As pointed out in the introduction, this chapter focuses on the wage differentials along the urban hierarchy in Spain, with a special interest in how the balance between the wage premia of different municipalities, classified according to their population sizes, was affected by the onset of the crisis since 2008. Following the procedure described in Arranz et al (2013), the waves of the MCVL from its release in 2004 up to 2013 are combined to build a monthly panel of workers under the general regime of the Social Security (the most common regulatory framework in the Spanish labor market). Due to computational limitations, the sample will be restricted to those employees that started their working history from 2004 onwards, and that worked in anytime during the period covered by the panel, which finally contains information about 23,882 individuals and 1,171,411 job spells.

There is an important limitation in this data set: the earnings are censored as the contribution base is subject to a minimum and a maximum value. This means that for earnings below or above these values, there is a fixed amount set every year by the administration that does not reflect the actual earnings. Nonetheless, the comparison made in the work of De la Roca and Puga (2016) for the period 2004-2009 using the censored contribution base and the uncensored tax records revealed that there are not significant differences in the results, which justifies the use of this measure in this preliminary analysis.

The dependent variable that will be under examination here is the individual monthly labor earnings, and the main explanatory factor will be the municipal size, represented by a set of six categories, being the reference tier the one containing the largest cities in the country, namely Madrid and Barcelona. The second tier contains those municipalities with less than one million but more than 500,000 inhabitants, the third one goes down to a population size between 500,000 and 250,000, the fourth category accounts for those municipalities between 250,000 and 100,000 inhabitants, the fifth one gathers cities between 100,000 and 40,000, and the last tier includes cities of less than 40,000 inhabitants. The estimation of the parameters associated to each tier in this size hierarchy will show the differences in wage levels, contrasting the existence of the urban wage premium. Furthermore, if the panel is divided into pre-crisis and during-crisis (current) periods, the changes that the wage differentials might suffered could be unveiled, contributing to the empirical evidence about the response of the Spanish urban system to generalized shocks.

Other control variables will be taken into account in order to observe the effect of the population size as clear as possible. This variables are factors widely used in wage equations, related to the worker and the company: the age and the age squared of the worker, his/her educational attainment (primary, secondary or tertiary), the job skill level (low, medium-low, medium-high or high), the experience (worked days in all the job spells) and tenure (worked days in the last job spells), the sector (agriculture, industry, transportation, construction, trade, communications, private services, public services, and other public activities) and the number of employees in the company.

Additionally, and merely for exploratory purposes, some variables representing certain features of the municipality of work are included. Two socio-economic factors extracted from the census of 2001 are regarded: the average educational level, measured as the share of population with a tertiary degree, and the unemployment rate for ages between 29 and

59. Some other variables, related to the natural amenities endowment of the area, are also included: the average rainfall, the minimum temperature in January, the maximum temperature in July, and the distance to the coast. These factors show the relationship of the wage level with the municipal traits, but provided their static character or the lack of observations along time, they cannot be included in a dynamic approach as the panel estimation.

In this preliminary version, all the attention will be drawn to measuring the urban wage premium and the effect that the crisis had on it from a broad perspective, which means that the estimations will cover only the city size and worker effects, without further discrimination of the static or dynamic components involved. The analysis will be performed taking as benchmark a stepwise approach, following the structure proposed in Yankow (2006).

As an initial stage, an estimation of the impact of the city-size classification on the log monthly earnings is presented, without taking into account other features apart from seasonality. This first appraisal exploits the self-containment quality of the waves of the MCVL, in order to evaluate the changes that this relationship shows along the years, hence for each year the following equation is regarded:

$$ln\, w_{ict} = \alpha + \sum_{c=1}^{5} \beta_c CITY_{ict} + \sum_{t=1}^{11} \gamma_t\, MONTH_{ict} + \varepsilon_{ict} \qquad [12]$$

Where w_{ict} is the individual monthly wage earned in a city in category c of the size classification and in the month t. $CITY_{ict}$ represents the fixed effect associated to the size tier that the city of work belongs to. The reference tier is the highest one, containing the largest cities – Madrid and Barcelona –. MONTH accounts for the seasonal effect, being January the reference month. In this estimation, the wages will be included both as a non-deflated and a deflated series, using the provincial consumer price index as a deflator.

The next step will use a pooled panel to estimate the urban wage premium, first for the entire period (2004-2013), and then for the pre- and during-crisis scenarios. This estimation includes all the variables mentioned above as a first approximation to the relationship between the wage premium and the considered factors. The specified model is the following:

$$ln\, w_{ict} = \alpha + \sum_{c=2}^{6} \beta_c CITY_{ict} + \sum_{t=2}^{T} \gamma_t\, YM_{ict} + \sum_{i=1}^{n} \delta_{it} X_{ict} + \varepsilon_{ict} \qquad [13]$$

In this case, t index represents the combination of year and month for the whole period of the panel (i.e. from January 2004 to December 2013), so the γ_t parameter measures seasonality, while X_{ict} is a vector containing individual, job and city features. This specification will begin with a basic scheme, which will be gradually completed with the remaining sets of control variables within *X*.

Finally, to control for unobserved characteristic of the worker, an estimation analogous to the one specified in Equation 2 is performed, but including individual fixed effects (μ_i):

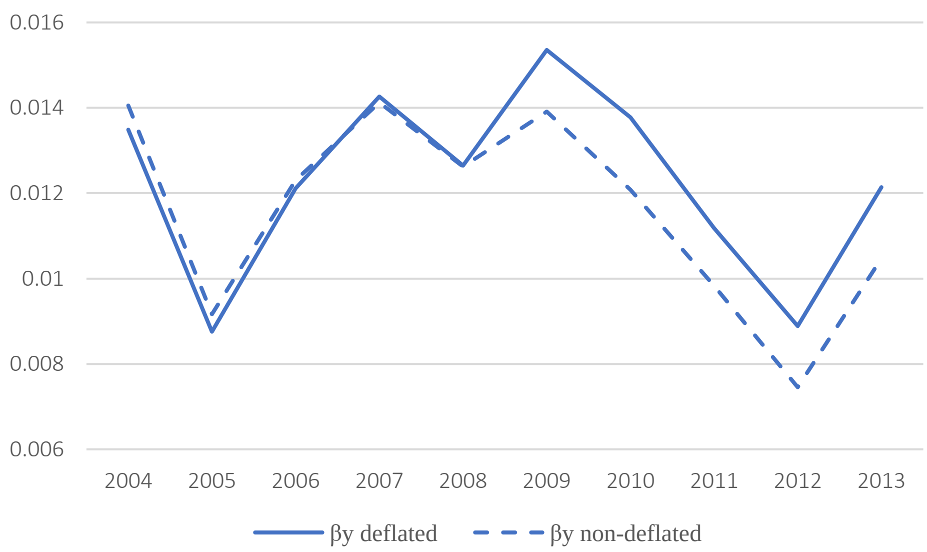
$$ln\, w_{ict} = \alpha + \mu_i + \sum_{c=2}^{6} \beta_c CITY_{ict} + \sum_{t=2}^{T} \gamma_t\, YM_{ict} + \sum_{i=1}^{n} \delta_{it} X_{ict} + \varepsilon_{ict} \qquad [14]$$

The idea behind this stepwise procedure is to explain the basic elements that might be behind the wage differential between Spanish cities progressively, as carried out in the related literature. Apart from the clearer distinction between the effects at stake, this approach allows to differentiate some of the static elements of the urban wage premium.

4. Results: local wage differentials and crisis

Using the aforementioned sample of workers who started their labor history in 2004, a first look to the relationship between monthly earnings and city size represented in Figure 21 shows that the coefficients are positive and significant for all the years considered. Only accounting for population size and seasonality (year-month dummies), the estimation returns that, overall, doubling the city size corresponds to an increase of 1.23% in the monthly earnings (1.16% regarding non-deflated earnings). It is also interesting to note that the difference between the deflated and non-deflated series gets wider from 2008, meaning that the share of the wage premium that can be attributed to a price effect is larger from that year compared to the previous ones, when both series basically overlap. Last but not least, basic to this study is the fact that 2008 pose a turning point for the effect of the population size on wages, provided that the aim of this chapter is to calculate if there are any significant differences in the urban wage premium before and during the economic crisis for Spain, and additionally, to check if the static and dynamic drivers behind the urban wage gap are also affected.

Figure 21. Annual evolution of OLS regression coefficients of deflated monthly earnings against Population size in 2001.



The positive relationship within wage and population size can be further explored if the population size is grouped into several classes, accounting also for the link not only with size as such, but also with the urban hierarchy. Taking the larger municipalities of Spain (Madrid and Barcelona, with more than 1,500,000 inhabitants) as the reference group, five additional size categories were built: municipalities between 1,500,000 and 500,000 inhabitants; between 500,000 and 250,000; between 250,000 and 100,000; between 100,000 and 40,000; and less than 40,000 inhabitants. Without considering any other factors apart from size and seasonality, the effect of each population category on the deflated earnings range from -0,032 to -0,002, but surprisingly this coefficients are not ordered according the size hierarchy.

The population size coefficient evolution depicted in Figure 21 can be also analyzed from an annual perspective for each of the classes, as shown in Figure 22, with respect to the first one, in order to see the dynamics of each tier in more detail. At the beginning of the period, the second and fourth categories present higher deflated monthly earnings than Madrid and Barcelona. The other three sit below, but the effect of size on earnings is not significant (black dot) in those cases. From 2005 to 2011 all the categories exhibit lower earnings than the largest cities, with a remarkable fall from 2007 to 2008 in most cases. This turning point in 2008 is presumably due to the effect of the crisis, while the changes in the order and the evolution of the different categories might be suggesting that the wage premium have a specific sensitivity in each tier, which in turn might be reflected by a different reaction of the wage premium components depending on the category considered.

Figure 22. Pooled OLS Coefficients for population size categories in each year

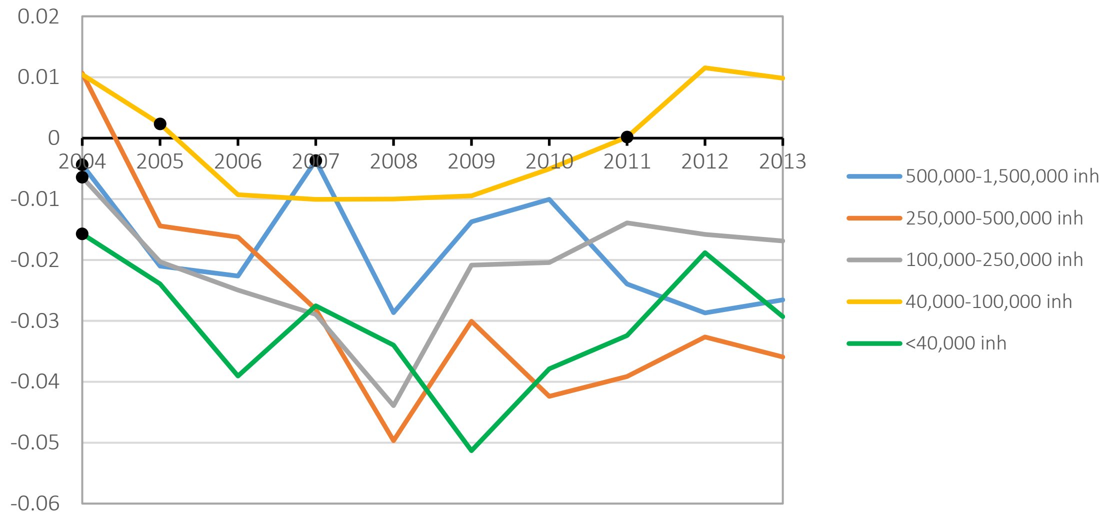


Table 11 contains pooled OLS gradual specifications that might shed some light on these findings, as they illustrate the response of the urban wage premium to changes in factors related to the worker, the job and the workplace. The estimation for the entire period (2004-2013) confirms the existence of such a premium for the largest metropolises in the country (Madrid and Barcelona), but as suggested in Figure 22, the wage gap does not necessarily follows the order of the urban hierarchy, especially when other factors are included in the analysis. According to these estimation, the gap is wider between the three highest tiers.

The worker and firm characteristics exhibit the expected behavior: earnings grow at a decreasing rate with age, the returns of education increase as the attainment level is higher, more demanding occupations in terms of skill are associated to a higher payments, and larger companies tend to offer higher wages too. The results regarding the features of the workplace offer mixed interpretations, but it is worth noting the negative effect of the distance to the coast, favouring coastal locations, and the positive effect of a high average educational level in the area.

Table 11. Pooled OLS estimations for the period 2004-2013

	(1)	(2)	(3)	(4)
Constant	6.717***	6.643***	6.476***	6.314 ***
City size				
1 million – 500,000 inh.	-0.019***	-0.039***	-0.026***	-0.030***
500,000 – 250,000 inh.	-0.031***	-0.044***	-0.028***	-0.035***
250,000 – 100,000 inh.	-0.023***	-0.033***	-0.014***	-0.024***
100,000 – 40,000 inh.	-0.002***	-0.021***	-0.006	-0.018***
< 40,000 inh.	-0.032***	-0.020***	-0.013***	-0.021***
Age		0.019***	0.019***	0.023***
Age²		-0.0003***	-0.0003***	-0.0003***
Educational level				
Secondary		0.039***	0.036***	0.034***
Tertiary		0.067***	0.059***	0.054***
Skill level				
Medium-low		0.060***	0.078***	0.082***
Medium-high		0.136***	0.161***	0.167***
High		0.349***	0.391***	0.402***
Number of workers			0.017***	0.017***
Average Rainfall				0.002***
Temp. max. July				0.004***
Temp. min. January				-0.009***
Dist. Coast				-0.037***
Educational level				0.048***
Unemployment rate (ages 20-51)				-0.004***
Adjusted R^2	*0.253*	*0.295*	*0.365*	*0.388*
F test	*3193.08 ****	*3628.08 ****	*3347.30 ****	*2508.19 ****

- Dependent variable: natural log of the individual monthly earnings (deflated)
- All specifications include month–year and industry indicators.
- */**/*** represent significance at the 10%/5%/1% level.

In the pre-crisis period (2004-2007), the disparities in the urban wage premium were slightly less marked if compared to the complete timeframe, as shown in Table 12. Regarding the individual and job factors, the effects just mentioned apply also to this period, but the returns are generally higher. However, there is a striking difference in the impact of the city educational level. During this period, a growth of 1% in the share of people with a university degree in the area is associated to a fall in the earnings of almost 5%.

Table 12. Pooled OLS estimations for the period 2004-2007

	(1)	(2)	(3)	(4)
Constant	6.715***	6.663***	6.436***	6.308***
City size				
1 million – 500,000 inh.	-0.013***	-0.039***	-0.024***	-0.034***
500,000 – 250,000 inh.	-0.017***	-0.037***	-0.013***	-0.022***
250,000 – 100,000 inh.	-0.023***	-0.035***	-0.016***	-0.025***
100,000 – 40,000 inh.	-0.004***	-0.026***	-0.017	-0.029***
< 40,000 inh.	-0.029***	-0.015***	0.007	-0.001
Age		0.018***	0.019***	0.020***
Age^2		-0.0002***	-0.0003***	-0.0003***
Educational level				
Secondary		0.041***	0.041***	0.041***
Tertiary		0.057***	0.053***	0.054***
Skill level				
Medium-low		0.051***	0.081***	0.103***
Medium-high		0.125***	0.169***	0.200***
High		0.390***	0.460***	0.497***
Number of workers			0.018***	0.018***
Average Rainfall				-0.001
Temp. max. July				0.006***
Temp. min. January				-0.008***
Dist. Coast				-0.040***
Educational level				-0.049***
Unemployment rate (ages 20-51)				-0.005***
Adjusted R^2	*0.253*	*0.295*	*0.365*	*0.388*
F test	*1034.48 ****	*1698.61 ****	*1276.37 ****	*968.30 ****

- Dependent variable: natural log of the individual monthly earnings (deflated)
- All specifications include month–year and industry indicators.
- */**/*** represent significance at the 10%/5%/1% level.

The crisis period (2009-2013), on the other hand, brings a reduction in the overall wage premium, in spite of the fact that the differences follow the same structure in the three sets of estimations, as can be seen in Table 13. Along with the reduction of the wage gaps, there is also a downturn in the effects of the skill level, but in this case, there is the counterpart of a higher earning in cities with higher educational level.

Table 13. Pooled OLS estimations for the period 2009-2013

	(1)	(2)	(3)	(4)
Constant	7.300 ***	6.878***	6.778 ***	6.631 ***
City size				
1 million – 500,000 inh.	-0.020***	-0.037***	-0.025***	-0.031***
500,000 – 250,000 inh.	-0.036***	-0.044***	-0.031***	-0.040***
250,000 – 100,000 inh.	-0.018***	-0.029***	-0.011***	-0.025***
100,000 – 40,000 inh.	0.001	-0.019***	-0.001	-0.014***
< 40,000 inh.	-0.034***	-0.020***	-0.018***	-0.026***
Age		0.019***	0.020***	0.025***
Age^2		-0.0003***	-0.0003***	-0.0003***
Educational level				
Secondary		0.036***	0.032***	0.029***
Tertiary		0.074***	0.065***	0.061***
Skill level				
Medium-low		0.060***	0.069***	0.060***
Medium-high		0.137***	0.151***	0.142***
High		0.327***	0.357***	0.352***
Number of workers			0.017***	0.016***
Average Rainfall				0.002***
Temp. max. July				0.004***
Temp. min. January				-0.009***
Dist. Coast				-0.037***
Educational level				0.076***
Unemployment rate (ages 20-51)				-0.003***
Adjusted R^2	*0.038*	*0.200*	*0.283*	*0.308*
F test	*371.22 ****	*2019.22 ****	*2142.95 ****	*1628.41 ****

- Dependent variable: natural log of the individual monthly earnings (deflated)
- All specifications include month–year and industry indicators.
- */**/*** represent significance at the 10%/5%/1% level.

The estimations presented support the suspicion of a change in the relative wage differences between cities, and that those differences might be the result of the disturbed economic framework brought about by the crisis. To look further into this matter, the following tables show comparable specifications, incorporating worker fixed-effects into the equations. Surprisingly, the inclusion of individual effects boost the effect of the urban tiers on wages, at least when the whole timespan is considered (Table 14), although the impact of the individual observable features is less intense. This effect fades as additional control variables are included in the analysis. The proximity to the coast also diminish its influence, while the average educational level stays at comparable levels regarding the related pooled OLS estimation.

Table 15 shows that in the pre-crisis period wages differentials in terms to city size were significantly wider, conversely to the results presented in Table 16 regarding the crisis situation, where the magnitude and the dispersion of the urban wage premium is smaller, in accordance with the results of Bande et al. (2008).

Table 14. Fixed-effects estimations for the period 2004-2013

	(1)	*(2)*	*(3)*	*(4)*
Constant	6.546***	6.728***	6.598***	6.314 ***
City size				
1 million – 500,000 inh.	-0.039***	-0.039***	-0.018***	-0.010***
500,000 – 250,000 inh.	-0.040***	-0.034***	-0.007***	0.003
250,000 – 100,000 inh.	-0.035***	-0.031***	-0.012***	-0.001
100,000 – 40,000 inh.	-0.008***	-0.013***	-0.005***	0.001
< 40,000 inh.	0.019***	0.015***	0.006***	0.012***
Age		0.015***	0.015***	0.017***
Age^2		-0.0003***	-0.0002***	-0.0003***
Educational level				
Secondary		0.025***	0.021***	0.023***
Tertiary		0.065***	0.053***	0.052***
Skill level				
Medium-low		0.045***	0.047***	0.051***
Medium-high		0.106***	0.105***	0.110***
High		0.231***	0.233***	0.236***
Number of workers			0.011***	0.012***
Average Rainfall				-0.003***
Temp. max. July				-0.002***
Temp. min. January				-0.0001
Dist. Coast				-0.010***
Educational level				0.051***
Unemployment rate (ages 20-51)				-0.001***
R^2				
Within	*0.469*	*0.369*	*0.380*	*0.363*
Between	*0.159*	*0.256*	*0.327*	*0.348*
Overall	*0.2514*	*0.273*	*0.317*	*0.328*
F test	*8162.59 ****	*4952.29 ****	*3468.62 ****	*2176.79 ****

- Dependent variable: natural log of the individual monthly earnings (deflated)
- All specifications include month–year, industry and individual indicators.
- */**/*** represent significance at the 10%/5%/1% level.

Table 15. Fixed-effects estimations for the period 2004-2007

	(1)	(2)	(3)	(4)
Constant	6.546***	6.588***	6.418***	6.295***
City size				
1 million – 500,000 inh.	-0.045***	-0.061***	-0.055***	-0.046***
500,000 – 250,000 inh.	-0.050***	-0.049***	-0.012***	0.007
250,000 – 100,000 inh.	-0.037***	-0.047***	-0.031***	-0.026***
100,000 – 40,000 inh.	-0.010***	-0.017***	-0.014***	-0.012**
< 40,000 inh.	0.028***	0.024***	0.008	0.010
Age		0.028***	0.030***	0.039***
Age²		-0.001***	-0.001***	-0.001***
Educational level				
Secondary		0.009**	-0.007	0.006
Tertiary		0.037***	0.024***	0.013
Skill level				
Medium-low		0.058***	0.065***	0.063***
Medium-high		0.106***	0.108***	0.108***
High		0.240***	0.262***	0.271***
Number of workers			0.010***	0.012***
Average Rainfall				-0.002
Temp. max. July				0.002
Temp. min. January				-0.001
Dist. Coast				-0.018***
Educational level				-0.032
Unemployment rate (ages 20-51)				-0.002***
R^2				
Within	*0.330*	*0.185*	*0.204*	*0.193*
Between	*0.044*	*0.109*	*0.198*	*0.227*
Overall	*0.110*	*0.133*	*0.210*	*0.244*
F test	*3818.71 ****	*1405.09 ****	*655.29 ****	*372.51 ****

- Dependent variable: natural log of the individual monthly earnings (deflated)
- All specifications include month–year, industry and individual indicators.
- */**/*** represent significance at the 10%/5%/1% level.

Table 16. Fixed-effects estimations for the period 2009-2013

	(1)	(2)	(3)	(4)
Constant	7.299***	7.010***	6.949***	6.875***
City size				
1 million – 500,000 inh.	-0.023***	-0.024***	-0.015***	-0.017***
500,000 – 250,000 inh.	-0.016***	-0.017***	0.001	-0.004
250,000 – 100,000 inh.	-0.016***	-0.015***	-0.0004	0.002
100,000 – 40,000 inh.	-0.011***	-0.015***	-0.006***	-0.005**
< 40,000 inh.	0.025***	0.023***	0.029***	0.027***
Age		0.015***	0.015***	0.017***
Age²		-0.0003***	-0.0002***	-0.0003***
Educational level				
Secondary		0.020***	0.020***	0.023***
Tertiary		0.043***	0.037***	0.039***
Skill level				
Medium-low		0.041***	0.037***	0.032***
Medium-high		0.101***	0.100***	0.097***
High		0.191***	0.195***	0.188***
Number of workers			0.012***	0.012***
Average Rainfall				0.001
Temp. max. July				0.0003
Temp. min. January				0.0002
Dist. Coast				-0.008***
Educational level				0.045***
Unemployment rate (ages 20-51)				-0.001*
R^2				
Within	*0.146*	*0.162*	*0.177*	*0.162*
Between	*0.009*	*0.174*	*0.245*	*0.250*
Overall	*0.037*	*0.174*	*0.230*	*0.243*
F test	*1549.16 ****	*1510.09 ****	*1129.80 ****	*679.47 ****

- Dependent variable: natural log of the individual monthly earnings (deflated)
- All specifications include month–year, industry and individual indicators.
- */**/*** represent significance at the 10%/5%/1% level.

5. Summary and main conclusions

One of the salient postulates of Urban Economics claims that there are certain dynamics related to agglomeration that foster the economy in the largest cities. According to this, the places at the top of the urban hierarchy benefit from improved efficiency and other related advantages, which in turn translates into a more productive system that attracts resources, especially labor, drawn by the expectation of higher wages. This difference in the earnings level between large cities and the rest of the urban system is called urban wage premium.

Among the few studies on local wages for the country, there are some evidence about the persistence in regional differences at an aggregated level (mainly induced by the regulatory framework of the labor market), and the existence of a wage gap between the tiers of the urban hierarchy. This chapter presents a preliminary analysis of the urban wage premium in Spain, trying to provide an empirical contribution with respect to the effect that the outbreak of the economic crisis had in the wage-population size relationship. Using the Continuous Sample of Employment Histories (MCVL), a panel was built for the period 2004-2013, containing information on the local earnings, and the individual, job and city features that might influence the wage level. The estimations focus on assessing the effect of the urban structure on earnings, separately for the pre- and during-crisis periods, through the gradual inclusion of control variables in a pooled OLS and a fixed-effects model.

Results show that the wage gap between the different urban tiers does not follow an ascending order according to size, being the premium wider when comparing areas at the top of the hierarchy. Another interesting finding is that the dispersion in the wage gap is larger in the pre-crisis term, supporting the notion of a positive association between the economic cycle and regional wage disparities. Variables as the age, the education attainment or the skill level have the expected influence on the monthly earnings, while the share of highly educated population in the area changes its effect depending on the period at issue: it has a negative impact in the pre-crisis estimations, and a positive one during the crisis. When other control variables are included in the analysis, the fixed-effects estimation accounting for each individual included in the panel shows a smaller impact of the urban tiers, in line with the sorting hypothesis, which states that part of the urban wage premium is explained by the preference of individuals to locate in larger cities according to certain personal unobservable characteristics.

Understanding the factors and dynamics behind the regional wage differentials is crucial from the policy perspective, as it provides with valuable information about the performance of the labor market, the possible resilience to economic shocks or the capacity of the area to attract population, relative to the position in the urban hierarchy.

GENERAL CONCLUSIONS AND FINAL REMARKS

Conclusions

This dissertation tried to capture part of the vast universe of processes that are at work in a system of cities like the Spanish one. The interrelationships between areas and the channels that structure those interactions offer a myriad of opportunities for research, although several challenges come along.

From the introduction to the last of the chapters, the evidence presents Spain as an interesting country to study the different stages of urban concentration, with a variety of situations regarding the intertwining of the economic and the demographic framework. The main objective of this dissertation is to add to the existing knowledge about the localization and agglomeration dynamics in Spain from a broad perspective, giving hints about the functioning of population and employment growth, always taking into account their spatial distribution and their connection with the specific features of the area regarded, but bearing in mind the importance of the neighborhood effects and the relative position within the urban hierarchy. Following this benchmarks, the first part of this thesis analyze two demographic effects, population growth and ageing convergence, while the second describe two labor related issues, employment growth and wage differentials. The four chapters present the same analytical profile, dealing first with the determinants that might explain the spatial arrangement of the mentioned phenomena, and then with the appropriate tools to capture the underlying spatial autocorrelation, the spatial heterogeneity or the relationship between the urban tiers, depending on the research interest of each chapter.

Chapter I analyzes the local determinants of population growth in Spain assuming that the different socio-economic, political or geographical factors considered affect global population growth, but their effects can be different across the space and through time. Both aspects are treated using Geographically Weighted Regressions for two decades (1991-2001 and 2001-2011), as a supplement to an OLS global approach.

Results show that global methodologies mask significant spatial variability in the case of several factors, as for the initial population size, educational level or employment. This factors exhibit a different effect in different parts of the country, either in both decades or in one of them. The estimations show that the population growth of the areas close to the Mediterranean coast depend less than other areas of the country on the socio-economic factors, displaying significant population increases despite any unfavorable initial condition. On the other hand, population growth in those local areas located south or west the peninsula has a positive relationship with the initial population size, and a negative one with the distance to large cities, suggesting that agglomeration economies are active in this area, conversely to the rest of the country. There are also factors with an even effect that is properly accounted for via global estimations, as is the case of the industrial specialization.

Overall, the results for the Spanish Local Labor Markets support the existence of temporal and spatial heterogeneity in the determinants of population growth. Fundamental factors from the regional science perspective, as distance or population size, change their effects depending on where and when they are estimated. Socio-economic factors, susceptible of policy intervention (employment for example), have an even more heterogeneous effect on population growth across space. This scenery calls for attention from the research side provided the compensation effect that might arise when working with large spatial conglomerates, and from the policy perspective at the time to adopt country-wide strategies.

Chapter II focuses on the study of the spatial dimension of ageing between 2001 and 2014, following the same methodological procedure of the previous considering the specific characteristics of the area, allowing for spatial heterogeneity and using the municipalities as the basic spatial unit. Here the focus is to analyze the local factors behind the high concentration and growth of local ageing population in Spain, applying a convergence specification in order to unveil if the municipalities are heading towards ageing at different speeds.

The global approach (OLS) shows a negative β, indicating the existence of a generalized ageing convergence process, in which population size and proximity to large cities act as a counterbalancing forces. However, as expected, the local estimation through GWR displays an uneven spatial framework where the north-eastern quarter of the country converge in ageing terms, while the areas in the south and the west have low convergence parameters, and some coastal municipalities shift to the divergence range. The same holds for the previously mentioned size and distance to size effects, that can foster ageing convergence depending on the area regarded. This different scope of situations suggests that the ageing process might not be under an intense influence of the birth-mortality ratio, as this measures have a stable behavior in the period considered. As a result, this chapter calls the attention to the spatially heterogeneous local ageing profile of Spain.

Based on the spatial and temporal heterogeneity observed in the local employment growth number, in Chapter III the intention is to identify the empirical regularities underlying local employment growth for the decades 1991-2001 and 2001-2011, focusing on the influence of structural factors (location, size and geography), and on the existing spatial interactions via a Spatial Durbin model. Again, results confirm the sensitivity of the empirical analysis to the timespan and to the neighboring features.

Some determinants behave in the expected way, as the positive effect of the region own and neighboring initial human capital endowment, but its magnitude changes from one decade to the. However, other factors show different signs or become insignificant depending on the period, as is the case for the initial population size or the relative location. The hint is that during the first decade, size and location were associated with agglomeration adverse effects, which hindered employment growth.

The estimations suggest a differentiated local employment growth pattern for each decade. During the first decade local employment growth was enhanced by the proximity to small-medium sized areas. In the second decade, although the own initial employment rate and population size still have a negative effect, the positive *indirect* effects indicates that, in regard to growth, locations close to large urban areas and to LLMs with a high employment levels are favored. These results should be interpreted having in mind the different economic growth profiles of the decades at stake, since the first one is an expansionary period, while the second is a crisis one.

The last chapter offer an early analysis of the urban wage premium in Spain, trying to assess the effect that the economic crisis had in the wage-population size relationship, and in the structure of the wage gap regarding the urban size tiers. This tasks will be carried out by means of a panel built for the period 2004-2013 with the information provided by the Continuous Sample of Employment Histories (MCVL).

The largest cities at the top of the urban hierarchy benefit from improved efficiency and other advantages, which translate into higher wages, but provided the difference in the industrial outline of each area, and the relative impact of the crisis over the different activities, it is expected that the differential wages changed from the pre- to the during-crisis period. The studies on local wages for Spain find a persistent difference in regional wages related to the bargaining system and confirm the existence of the urban wage premium.

An interesting finding is that that the wage gap between the different population size tiers does not comply with the rank of the urban hierarchy. Another finding is that the wage dispersion is larger in the pre-crisis term, supporting the positive link between the economic cycle and regional wage disparities. Other factors as the age, the education attainment or the skill level have the expected effect. The difference between the pooled OLS and the fixed-effects estimations second the hypothesis that cities attract more able workers regarding unobservable characteristics.

Policy implications

The analyses performed and their associated results all point in the same direction: spatial and temporal heterogeneity have a deep influence in each of the demo-economic aspect concerned in this dissertation. Undoubtedly, the ever-changing nature of spatial relationships is behind the lack of uniformity in these dimensions.

The complexity of the spatial patterns both from a cross-section or longitudinal perspective turns policy decision more complicated. In these sense, any measure aimed at enhancing population growth, reducing the demographic burden or promote employment requires local tailored actions, considering the range of determinants under policy control.

There are not universal formulas to foster local growth and policies must be designed paying attention to local characteristics. However, the influence of the characteristics of the neighbors is another factor to consider, as the extent to which the policies applied in other areas may affect the own is not that easy to account for. In this regard, not only timely knowledge of the local processes is required, but coordination between local administration, and between different political stances.

The results presented here lend empirical support, in the case of Spain, to the new approaches adopted in the European Union Cohesion Policy in terms of smart specialization strategies, stressing the urge to develop policy projects tailored at a more local level, and revised in shorter periods of time. Instead of advocating for global actions, this relatively new perspective requires a differentiated appraisal of the regional current and potential strength in order to adapt the political intervention to each place.

To summarize, this dissertation supports a local smart specialization approach as a counterpart of the traditional wider territorial perspective, guided by careful analyses taking into account the spatial structure of the policy subject.

References

Adsera, A. (2004). Changing Fertility Rates in Developed Countries: The Impact of Labor Market Institutions. *Journal of Population Economics*, 17(1), pp. 17-43.

Adsera, A. (2005). Vanishing Children: From High Unemployment to Low Fertility in Developed Countries. *The American Economic Review*, 95(2), pp. 189-193.

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

Ali, K., M.D. Partridge and M.R. Olfert. (2007). Can Geographically Weighted Regression 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.

Anselin, L. (1999). *Interactive techniques and exploratory spatial data analysis. Geographical Information Systems: Principles, Techniques, Management and Applications*, eds., P. Longley, M. Goodchild, D. Maguire, and D. Rhind. Cambridge: Geoinformation Int.

Anselin, L., Florax, R.J.G.M. and Rey, S.J. (Eds) (2004). *Advances in Spatial Econometrics: Methodology, Tools and Applications*. Springer, Berlin

Anselin, L. (2010). Thirty Years of Spatial Econometrics. *Papers in Regional Science*, 89 (1), pp. 3-25.

Arranz, J.M., García-Serrano, C. and Hernanz, V. (2013). How do We Pursue "Labormetrics"? An Application using the MCVL. *Estadística Española*, 55 (181), pp. 231-254.

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

Bande, R., Fernandez, M. and Montuenga, V. (2008). Regional Unemployment in Spain: Disparities, Business Cycle and Wage Setting. *Labour Economics*, 15, pp. 885-914.

Barro, R.J. and Sala-i-Martí, X. (1995). *Economic Growth*. McGraw Hill, Boston

Barro, R. J., X. Sala-i-Martí, O. J. Blanchard and R. E. Hall. (1991). Convergence across States and Regions. *Brookings Papers on Economic Activity*, pp. 107-182.

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

Baumol, W. J. (1986). Productivity Growth, Convergence, and Welfare: What the Long-run Data Show. *The American Economic Review*, 76 (5), pp. 1072-1085.

Becattini, G. (1962). Il Sistema Marshaliano. In: Becattini G (ed) *Il Concetto di Industria e la Teoria del Valore*. Boringhieri, Torino, p. 88-125

Becattini, G. (1975). Introduzione. Invito a una rilettura di Marshall. In: Marshall A, Marshall P (eds) *Economia della Produziones, a Cura di G. Becattini*. ISEDI, Milano, p. 9-114

Becker, G. S., T. J. Philipson and R. R. Soares. (2005). The Quantity and Quality of life and the Evolution of World Inequality. *The American Economic Review*, 95, pp. 277-291.

Beckstead, D. and Brown, W.M. (2003). From Labrador City to Toronto: The industrial Diversity of Canadian Cities, 1992 to 2002. In: Statistics Canada (Ed*) Insights on the Canadian Economy* 3, Ontario

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

Besag, J. (1974) Besag, J. (1974). Spatial Interaction and the Statistical Analysis of Lattice Systems. *Journal of the Royal Statistical Society*. Series B (Methodological), pp. 192-236.

Blasco, B. C. (2008). Desigualdades Territoriales en Relación con el Envejecimiento de la Población Española. *Documents d'Anàlisi Geographica*, 52, pp. 91-110.

Blien, U., Suedekum, J. and Wolf, K. (2006). Local Employment Growth in West Germany: A Dynamic Panel Approach. *Labour Economics*, 13 (4):445-458

Bloom, D.E., D. Canning and J. Sevilla. (2003). *the Demographic Dividend: A New Perspective on the Economic Consequences of Population Change*, Rand Corporation.

Boix, R. and Galletto, V. (2006) Sistemas Locales de Trabajo y Distritos Industriales Marshallianos en España. *Economía Industrial*, 359, pp. 165-184.

Boix, R and Trullén, J. (2010). Industrial Districts, Innovation and I-district Effect: Territory or Industrial Specialization*? European Planning Studies*, 18 (10), pp. 1705- 1727.

Bourguignon, F. and C. Morrison. (2002). Inequality among World Citizens: 1820-1992. *The American Economic Review*, 92, pp. 727-744.

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

Christaller, W. (1933). *Central Places in Southern Germany*. Jena: Fischer (English translation, London: Prentice Hall, 1966).

Ciccone, A. and hall, R.E. (1996). Productivity and the Density of Economic Activity, *American Economic Review*, 86, 54–70.

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.

Cliff, A.D. and Ord, J.K. (1973). *Spatial Autocorrelation*. London: Pion.

Cliff, A.D. and Ord, J.K. (1981). *Spatial Processes: Models and Applications*. Pion, London

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.

Coffey, W.J. and Shearmur, R. (1996). *Employment Growth and Change in the Canadian Urban System*. Canadian Policy Research Networks, Ottawa

Combes, P.P., Duranton, G. and Gobillon, L. (2008). Spatial Wage Disparities: Sorting Matters! *Journal of Urban Economics*, 63, pp. 723–742.

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.

Cooke, P., Heindenreich, M. and Braczyk, H.J. (2004). *Regional Innovation Systems: The Role of Governance in a Globalized Word*. Routledge, London

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.

Cressie, N. and Wikle, C. (2011). *Statistics for Spatio-temporal Data*. New York: Wiley.

D'Costa, S. and Overman, H.G. (2014). The Urban Wage Growth Premium: Sorting or Learning? *Regional Science and Urban economics*, 48, pp. 168-179.

Davis, D.R. and Weinstein, D.E. (2002). Technological Superiority and the Losses from Migration. National Bureau of Economic Research Working Paper Series No. 8971. http://www.nber.org/papers/w8971.

De la Roca, J. and Puga, D. (2016). Learning by Working in Big Cities. *Review of Economic Studies*. DOI: 10.1093/restud/rdw031.

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

Del Valle-Ramos, C. (2007). Envejecimiento Demográfico y Migraciones en La Ciudad de Sevilla (2000-2004). *Papeles de Geografía*, 45-46, pp. 187-201.

Doloreux, D., Shearmur, R. and Filion, P. (2001). Learning and Innovation: Implications for Regional Policy. An Introduction. *Canadian Journal of Regional Science*, 24(1):5-11

Dorius, S. F. (2008). Global Demographic Convergence? A Reconsideration of Changing Intercountry Inequality in Fertility. *Population and Development Review*, 34 (3), pp. 519-537.

Duranton, G. and Puga, D. (2001). Nursery Cities: Urban Diversity, Process Innovation, and the Life Cycle of Products. *American Economic Review*, 91 (5), pp. 1454–1477.

Duranton, G. and Puga, D. (2004). Micro-foundations of Urban Agglomeration Economies. In Henderson, J.V. and Thisse, J.F. (Eds.), *Handbook of Urban and Regional Economics,* Vol. 4. Amsterdam: North-Holland.

Easterlin, R. A. (2000). The Worldwide Standard of Living since 1800. *The Journal of Economics Perspectives*, 14, pp. 7-26.

Edwards, R. D. (2011). Changes in World Inequality in Length of Life: 1970-2000. *Population and Development Review*, 37 (3), pp. 499-528.

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

Eisinger, P.K. (1988). *The Rise of the Entrepreneurial State: State and Local Economic Development Policy in the United States*. The University of Wisconsin Press, Madison.

Europe. Eurostat. (2014). *Population Structure and Ageing. Statistics Explained* [Online]. Available From Http://Ec.Europa.Eu/Eurostat/Statistics-Explained/Index.Php/Special:Collection/Rendering/?Return_to=Population+Structure+and+AgeingandCollection_Id=andWriter=andStable=andQueue=0andPdfid=156.35.225.70_1432033219_74.Pdf

Feldstein, M. S. (2006). The Effects of the Ageing European Population on Economic Growth and Budgets: Implications for Immigration and Other Policies. National Bureau of Economic Research (NBER Working Paper No. 12736). Cambridge: National Bureau of Economic Research.

Florida, R. (2002). *The Rise of the Creative Class*. Basic Books, New York

Fotheringham, A.S., Brunsdon, C. and Charlton, M. (2000). *Quantitative Geography: Perspectives on Spatial Data Analysis.* London: Sage.

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

Fox, K.A. and Kumar, T.K. (1965). The functional economic area: delineation and implications for economic analysis and policy. *Papers and proceedings of the Regional Science Association*, 15: 57-85.

Fratessi, U. and M. Riggi. (2007). Does Migration Reduce Regional Disparities? The Role of Skill-Selective Flows. *Review of Urban and Regional Development Studies*, 19 (1), pp. 78-102.

Freeman, D.G. (2001). Sources of Fluctuation in Regional Growth. *The Annals of Regional Science*, 35(2):294-266

Fuchs, V. (1959). Differentials in Hourly Earnings by Region and City Size. NBER Occasional paper 101.

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

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.

Galaway B, Hudson J (Eds) (1994) *Community Economic Development: Perspectives on Research and Policy*. Thompson, New York

Gallego, R. and J. Subirats. (2012). Spanish and Regional Welfare Systems: Policy Innovation and Multi-Level Governance. *Regional and Federal Studies*, 22 (3), pp. 269-288.

García-Ballesteros, A. and B. Sanz-Berzal. (Coord.). (2002). *Atlas de la Comunidad de Madrid en el Umbral del Siglo XXI: Imagen Socioeconómica de una Región Receptora de Inmigrantes*. Madrid: Editorial Complutense.

Glaeser, E.L. (1999), Learning in Cities. *Journal of Urban Economics*, 46, pp. 254–277.

Glaeser, E.L. and Maré, D.C. (2001). Cities and Skills. *Journal of Labor Economics*, 19 (2), pp. 316-342.

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.

Goesling, B. and Firebaugh, G. (2004). The Trend in International Health Inequality. *Population and Development Review*, 30, pp. 131-146.

Goli, S. and Arokiasamy, P. (2014). Maternal and Child Mortality Indicators across 187 Countries of the World: Converging or Diverging. *Global Public Health*, 9 (3), pp. 342-360.

Gould, E.D. (2007). Cities, Workers, and Wages: A Structural Analysis of the Urban Wage Premium. *Review of Economic Studies*, 74, pp. 477-506.

Grafeneder-Weissteiner, G. and Prettner, K. (2013). Agglomeration and Demographic Change. *Journal of Urban Economics*, 74, pp. 1-11.

Graves, P. E. (1980). Migration and Climate. *Journal of Regional Science*, 20(2), pp. 227-237.

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

Gregory, T. and Patuelli, R. (2015). Demographic Ageing and the Polarization of Regions – An Exploratory Space-Time Analysis. *Environment and Planning A*, 47 (5), pp. 1.192–1.210.

Griffith, D.A. (1988). *Advanced Spatial Statistics: Special Topics in the Exploration of Quantitative Spatial Series*. Kluwer, Dordrecht

Griffith, D.A. (2003). *Spatial Autocorrelation and Spatial Filtering: Gaining Understanding through Theory and Scientific Visualization*. Springer, New York

Gunton, T. (2003). Natural resources and regional development: an assessment of dependency and comparative advantage paradigms. *Economic Geography*, 79: 67–94

Haining, R. (1990). The Use of Added Variable Plots in Regression Modelling with Spatial Data. *The Professional Geographer*, 42:336-44

Helsey, R.W. and Strange, W. (1990). Matching and Agglomeration Economies in a System of Cities. *Regional Science and Urban Economics*, 28, pp. 189–212.

Henderson, J.V. (2003). Marshall's Scale Economies. *Journal of Urban Economics*, 53, pp. 1–28.

Holmes, T.J. (2010). Structural, Experimentalist, and Descriptive Approaches to Empirical Work in Regional Economics. *Journal of Regional Science*, 50, pp. 5–22.

Hummels, D. (1999). Have International Transportation Cost Declined? Mimeo. University of Chicago.

Isard, W. (1956). *Location and Space-economy*. Cambridge: MIT Press.

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

Jacobs, J. (1984). *Cities and the Wealth of Nations*. Vintage, New York

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

Jones, J.P. and Casetti, E. (1992). *Applications of the Expansion Method*. London: Routledge.

Kanbur, R. and H. Rapoport. (2005). Migration Selectivity and the Evolution of Spatial Inequality. *Journal of Economic Geography*, 5 (1), pp. 43-57.

Keating, M. (1993). *The Politics of Modern Europe: The State and Political Authority in the Major Democracies*. Edward Elgar, Aldershot

Krugman, P.R. (1995). *Development, Geography, and Economic Theory*. The MIT Press, Cambridge

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.

Lardiés, R. (2009). Entre la Movilidad Residencial y la Migración. Factores y Razones para el Cambio de Residencia desde un Entorno Urbano entre la Población Mayor. *Estudios Geográficos*, 71 (268), pp. 177-201.

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

LeSage, J.P. and Fischer, M.M. (2008). Spatial growth regressions: Model specification, estimation and interpretation. *Spatial Economic Analysis*, 3 (3): 275–304

LeSage, J.P. and Pace, R.K. (2009). *Introduction to Spatial Econometrics*. Taylor & Francis, Florida

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

Lösch, A. (1938). The Nature of Economic Regions. *Southern Economic Journal*, 5(1):71-78

Lösch, A. (1940). *The Economics of Location*. Jena: Fischer (English translation, New Haven: Yale University Press, 1954).

Mackay, R.R. (2003). Twenty-five Years of Regional Development. *Regional Studies*, 37(3):303-317

Malecki, E. and Oinas, P. (1999). *Making Connections. Technological Learning and Regional Economic Change*. Ashgate, Aldershot

Markusen, A. (1996). Sticky Places in Slippery Space: A Typology of Industrial Districts. *Economic Geography*, 72(3):293-313

Markusen, A. (1999). Fuzzy Concepts, Scanty Evidence, Policy Distance: The Case for Rigour and Policy Relevance in Critical Regional Studies. *Regional Studies*, 33(9):869-884

Martin, R. and Sunley, P. (1998). Slow Convergence? The New Endogenous Growth Theory and Regional Development. *Economic Geography*, 74:201-227

Massey, D.B. (1995). *Spatial Divisions of Labor: Social Structures and the Geography of Production*. Routledge, New York

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.

Miller, H. and Han, J. (2009). *Geographic Data Mining and Knowledge Discovery*. Boca Raton: CRC Press.

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.

Neumayer, E. (2003). Beyond Income: Convergence in Living Standards, Big Time. *Structural Change and Economic Dynamics*, 14, pp. 275-296.

Neumayer, E. (2004). HIV/AIDS and Cross-National Convergence in Life Expectancy. *Population and Development Review*, 30, pp. 727-742.

Paelink, J. and Klaassen, L. (1979). *Spatial Econometrics*. Farnborough: Saxon House.

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.

Parr, J. (2001). Regional Economic Planning: Learning from Past Experience. In: Felsenstein D, McQuaid R, McCann P and Shefer D (eds) *Public Investment and Regional Economic Development*. Edward Elgar, Cheltenham

Parr, J. (2002). Agglomeration Economies: Ambiguities and Confusions. *Environment and Planning A*, 34:717-731

Parr, J. (2008). Cities and Regions: Problems and Potentials. *Environment and Planning A*, 40, pp. 3,009-3,026.

Parreño, J. M., Domínguez, J. and Díaz, R. (2014). El Impacto Territorial de las Variaciones Residenciales de los Jubilados Españoles. XIV Congreso Nacional de Población, Sevilla.

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. and Rickman, D.S. (2006). *The Geography of American Poverty: Is There a Role for Place Based Policy?* W.E. Upjohn Employment Institute

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.

Partridge, M.D., Rickman, D.S., Ali, K. and Olfert, M.R. (2009). Agglomeration Spillovers and Wage and Housing Cost Gradients across the Urban Hierarchy. *Journal of International Economics*, 78:126-140

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.

Philo, C., Mitchell, R. and More, A. (1998). Reconsidering Quantitative Geography: The Things that Count. *Environment and Planning A*, 30 (2):191-201

Polèse, M. (2009). *The Wealth and the Poverty of Regions: Why Cities Matters*. University of Chicago Press, Chicago.

Polèse, M., Rubiera, F. and Shearmur, R. (2007). Observing Regularities in Location Patterns: An Analysis of the Spatial Distribution of Economic Activity in Spain. *European Urban and Regional Studies*, 14 (2), pp. 157-180.

Polèse, 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.

Porter, M. (1990). *The Competitive Advantage of Nations*. The Free Press, New York

Porter, M. (1996). Competitive Advantage, Agglomeration Economies, and Regional Policy. *International Regional Science Review*, 19 (1-2):85-90

Porter, M. (1998). *Competitive Strategy: Techniques for Analyzing Industries and Competitors*. The Free Press, New York

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

Pradhan, M., D. E. Sahn and Younger, S. D. (2003). Decomposing World Health Inequality. *Journal of Health Economics*, 22, pp. 271-293.

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

Pujadas, I., J. Bayona-I-Carrasco and M. Rubiales. (2014). El Incremento de las Migraciones Internas en las Edades Avanzadas: Pautas Territoriales y Características Sociodemográficas. XIV Congreso Nacional De Población, Sevilla.

Quigley, J.M. (1998). Urban Diversity and Economic Growth. *Journal of Economic Perspectives,* 12 (2):127-138

Ram, R. (2006). State of the Life Span Revolution between 1980 and 2000. *Journal of Development Economics*, 80, pp. 518-526.

Ramos, R., Nicodemo, C. and Sanromá, E. (2015). A Spatial Panel Wage Curve for Spain. *Letters in Spatial and Resource Sciences*, 8 (2), pp. 125-139.

Rappaport, J. and Sachs, J.D. (2003). The United States as a Coastal Nation. *Journal of Economic Growth*, 8:5-46

Reig, E. (2007). *Competitividad, Crecimiento y Capitalización de las Regiones Españolas*. Madrid: Fundación BBVA.

Romer, P. (1990). Human Capital and Growth: Theory and Evidence. *Carnegie-Rochester Conference Series on Public Policy*, 32 (1): 251-286

Rosenthal, S.S. and Strange, W. (2004). Evidence on the Nature and Sources of Agglomeration Economies. In Henderson, J.V. and Thisse, J.F. (Eds.), *Handbook of Regional and Urban Economics*, Vol. 4. Amsterdam: North-Holland.

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.

Sala, G. (2014). Federalism without Adjectives in Spain. *The Journal of Federalism*, 44 (1), pp. 109-134.

Sala-i-Marti, X. (1996). The Classical Approach to Convergence Analysis. *The Economic Journal*, 106, pp. 1019-1036.

Sato, Y. (2001). Labor Heterogeneity in an Urban Labor Market. *Journal of Urban Economics*, 50, pp. 313–337.

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. (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.

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., Openshaw, S. and Wymer, C. (1997). Le Procedura di Identificazione dei Sistemi Locali del Lavoro. In: Sforzi F (Ed) *I Sistemi Locali del Lavoro*, p 235–242. ISTAT, Rome

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.

Shearmur, R. and Polèse, M. (2007). Do Local Factors Explain Local Employment Growth?: Evidence from Canada, 1971-2001. *Regional Studies*, 41 (4):453-471

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.

Simón, H.J., Ramos, R. and Sanromá, E. (2006). Collective Bargaining and Regional Wage Differences in Spain: An Empirical Analysis. *Applied Economics*, 38, pp. 1749-1760.

Smart, M.W. (1974). Labour market areas: uses and definition. In: Diamond DR, McLoughlin JB (eds) *Progress in planning*, vol 2. Pergamon Press, Oxford, p. 239-353

Solow, R. M. (1956). A Contribution to the Theory of Economic Growth. *The Quarterly Journal of Economics*, 70 (1), pp. 65-94.

Spain. Agencia Española De Meteorología. (2011). *Atlas Climático Ibérico* [Online]. Madrid: Ministerio de Medio Ambiente y Medio Rural y Marino. Available from: http://www.aemet.es/documentos/es/conocermas/publicaciones/Atlas-climatologico/Atlas.pdf

Spain. Instituto Nacional de Estadística. (1992). *Censo de Población, 1991* [Online]. Available from http://www.ine.es/jaxi/menu.do?type=pcaxis&path=%2Ft20%2Fe243&file=inebase&L=0

Spain. Instituto Nacional de Estadística. (2007). *Censo de Población, 2001* [Online]. Available From Http://Www.Ine.Es/Jaxi/Menu.Do?Type=PcaxisandPath=%2ft20%2fe242andFile=InebaseandL=0

Spain. Instituto Nacional de Estadística. (2012). *Censo de Población, 2011* [Online]. Available From http://www.ine.es/censos2011_datos/cen11_datos_res_col.htm

Spain. Instituto Nacional de Estadística. (2014). *Padrón Municipal de Habitantes. Estadística Del Padrón Continuo* [Online]. Available From Http://Www.Ine.Es/Jaxi/Menu.Do?Type=PcaxisandFile=PcaxisandPath=%2ft20%2fe245%2fp05%2f%2fa2014

Spain. Instituto Geográfico Nacional. (2001). *Mapas Digitales de los Municipios Españoles. Centro Nacional de Información Geográfica*. Available From http://centrodedescargas.cnig.es/CentroDescargas/equipamiento/BD_Municipios-Entidades.zip

Strambach, S. (2001). Innovation Processes and the Role of Knowledge-Intensive Business Services (KIBS). In: Koschatzky K, Kulicke M, Zenker A (Eds) Innovation Networks – Concepts and Challenges in the European Perspective. Physica Verlag, Heidelberg, p 53–68

Van Der Gaag, N. and De Beer, J. (2015). From Demographic Dividend to Demographic Burden: the Impact of Population Ageing on Economic Growth in Europe. *Tijdschrift voor Economische en Sociale Geografie*, 106 (1), pp. 94-109.

Vias, A.C. (2012) Micropolitan Areas and Urbanization Processes in the US. *Cities*, 29, pp. 24-28.

Vila, M. R. and Martín, T. C. (2007). Childbearing Patterns of Foreign Women in a New Immigration Country: the Case of Spain. *Population*, 62(3), pp. 419-450.

Von Thünen, J.H. (1826). *Von Thünen's Isolated State*. Hamburg (English translation, Oxford: Pergamon Press, 1966).

Weber, A. (1909). *Alfred Weber's Theory of the Location of Industries*. Tübingen: Verlag Mohr (English translation, Chicago: University of Chicago Press, 1929).

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.

Whittle, P. (1954). On Stationary Processes in the Plane. *Biometrika* 41, pp. 434–449.

Yankow, J., (2006). Why do Cities Pay More? An Empirical Examination of Some Competing Theories of the Urban Wage Premium. *Journal of Urban Economics*, 60 (2), Pp. 139-161.

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 Behaviour and the Principle of Least Effort: An Introduction to Human Ecology*. Adison Wesley Press, Cambridge.