ECONSTOR
Make Your Publications Visible.

A Service of

ZBW Leibniz-Informationszentrum Wirtschaft
Leibniz Information Centre for Economics

Liviano, Daniel; Arauzo-Carod, Josep-Maria

Conference Paper

Spatial Exploration of Age Distribution in Catalan Municipalities

52nd Congress of the European Regional Science Association: "Regions in Motion - Breaking the Path", 21-25 August 2012, Bratislava, Slovakia

Provided in Cooperation with:
European Regional Science Association (ERSA)

Suggested Citation: Liviano, Daniel; Arauzo-Carod, Josep-Maria (2012) : Spatial Exploration of Age Distribution in Catalan Municipalities, 52nd Congress of the European Regional Science Association: "Regions in Motion - Breaking the Path", 21-25 August 2012, Bratislava, Slovakia, European Regional Science Association (ERSA), Louvain-la-Neuve

This Version is available at:
https://hdl.handle.net/10419/120469

Standard-Nutzungsbedingungen:

Die Dokumente auf EconStor dürfen zu eigenen wissenschaftlichen Zwecken und zum Privatgebrauch gespeichert und kopiert werden.

Sie dürfen die Dokumente nicht für öffentliche oder kommerzielle Zwecke vervielfältigen, öffentlich ausstellen, öffentlich zugänglich machen, vertreiben oder anderweitig nutzen.

Sofern die Verfasser die Dokumente unter Open-Content-Lizenzen (insbesondere CC-Lizenzen) zur Verfügung gestellt haben sollten, gelten abweichend von diesen Nutzungsbedingungen die in der dort genannten Lizenz gewährten Nutzungsrechte.

Terms of use:

Documents in EconStor may be saved and copied for your personal and scholarly purposes.

You are not to copy documents for public or commercial purposes, to exhibit the documents publicly, to make them publicly available on the internet, or to distribute or otherwise use the documents in public.

If the documents have been made available under an Open Content Licence (especially Creative Commons Licences), you may exercise further usage rights as specified in the indicated licence.

WWW.ECONSTOR.EU

Mitglied der Leibniz-Gemeinschaft

Spatial Exploration of Age Distribution in Catalan Municipalities

Daniel Liviano ♣

Josep-Maria Arauzo-Carod ♠ ♦

Abstract

This paper takes the shelf and digs into the complex population's age structure of Catalan municipalities between 1999 and 2009. Catalonia is a very heterogeneous territory, and age pyramids vary considerably across different areas of the territory, existing geographical factors shaping municipalities' age distributions. By means of spatial statistics methodologies, this piece of research work tries to assess which spatial factors determine the location, scale and shape of local distributions. The results show that there exist different distributional patterns across the geography according to specific determinants.

Keywords: Spatial Models.

JEL Classification: C21.

This research was partially funded by CICYT's SEJ2010-19088/ECON and SEJ2010-16934/ECON), the "Xarxa de Referència d'R+D+I en Economia i Polítiques Públiques" and the SGR Program (2009-SGR-322) of the Catalan Government. The usual disclaimer applies.

(♣) Universitat Oberta de Catalunya (UOC); Av. Tibidabo 39-43; 08035 Barcelona (Catalonia); Phone: +34 932542169.

(♠) Universitat Rovira i Virgili, Departament d'Economia (QURE-CREIP); Av. Universitat 1; 43204 Reus (Catalonia); Phone: +34 977758902; Fax: +34 977759810.

(♦) Institut d'Economia de Barcelona (IEB); Av. Diagonal 690; 08034 Barcelona (Catalonia).

1 Motivation

Population Economics is an area that has attracted a growing number of both theoretical and empirical contributions in recent years and in a wide range of journals. Inside this area, many contributions rely on issues related with population growth, which is analysed from a broad type of perspectives like, among others, relationship between population and jobs' growth, internal and international migration, locational preferences of individuals, commuting, counterurbanization, etc. However, there are other areas that deserve further attention. One of these areas is age distribution, which refers to the distribution of population along the age dimension. Differences in terms of age distribution lead to different population pyramids, which is the most usual way to graphically represent age distribution. In this sense, growing populations show expanding pyramids (triangle-shaped with the base at the bottom), stable populations show stationary pyramids (square-shaped) and elder populations show contracting pyramids (triangle-shaped with the base at the top). According to population projections by United Nations Population Division (World Population Prospects, the 2010 Revision) the share of individuals aged 65 years and over is expected to grow from 19.40% in 2010 to 32.03% in 2050, as well as those between 15 and 64 are expected to shrink from 64.77% in 2010 to 52.75% in 2050[1]. These changes in age structure will obviously affect future societies in terms of, for instance, higher demands for health care and nursing (Uhlenberg, 2006), so age structure really matters and needs to be taken into account.

Yet another issue that deserves full attention is the role of space. Although space is a cornerstone for many demographic theories, and spatial effects in population dynamics have been largely theorized, many demographic studies lack a spatial perspective (Tiefelsdorf 2000). This is, when it comes to empirical research, the formal modeling of spatial effects is still a matter pending. This omission implies the assumption that population dynamics in a specific area has no relationship with what happens in neighbouring areas, which is clearly unrealistic and contradictory with, for instance, Waldo Tobler's first law of geography (1970), which states that *"everything is related to everything else, but near things are more related than distant things"*.

[1] In any case, it is important to notice that population forecasting can be biased for long term projections (Keilman, 1997).

However, recently scholars (see, among others, Chi and Marcouiller, 2012) have introduced into the analysis the spatial dimension of population by studying spatial dependence. Fortunately, there is still room for properly addressing spatial implications of Population Economics by using new technologies like Geographical Information Systems (GIS), which allow not only to collect and represent geographically referenced data (White and Lindstrom, 2006), but also to infer relationships among variables. Any revision of urban and regional economics literature will show that economic activity is strongly agglomerated[2]. Likewise, population is concentrated in dense urban areas (Fosset, 2006), and that extant agglomeration economies influence the way in which firms and individuals behave. Consequently, agglomeration influences both firm demography[3] and population demography.

In the case of Catalonia, Figure 1 shows the estimated density of the aggregated population's age for years 1999 and 2009. This evidence shows that the centre of the distribution has moved to the right side, which points out an underlying overall ageing process. However, the shape of both distributions reveal more information that that. For instance, there is a bump between age values 0 and 15 for the year 2009, showing a birth rate boost in the period 1999-2009. Likewise, the population percentage in the age interval 60-80 has decreased throughout this period.

However, the result depicted in Table 1 does not reveal to what extent population's age structure changes occur at a much more disaggregated level, and thus they are caused by local determinants, among which spatial effects are prominent. These effects can be categorized into two groups. On the one hand, spatial dependence[4] can be defined as a similarity (and also dissimilarity) measure between spatial variables located nearby. Spatial dependence is composed of large-scale spatial irregularities and local-scale spatial interaction effects. On the other hand, spatial heterogeneity is a much broader concept that refers to differences in a variable distributional parameters across space. Spatial heterogeneity often affects mean, variance and covariance structures of the data.

[2] See Parr (2002) for a review of agglomeration economies' classification.

[3] See, among others, Nyström (2007) for firm entry and exit at a regional level in Sweden, Fritsch et al. (2006) for firm survival at a district level in West Germany and Arauzo-Carod and Teruel-Carrizosa (2005) for firm entry at a urban level in Spain.

[4] Spatial dependence is also known as spatial autocorrelation and spatial interaction, although some geographers and demographers understand them differently. See (Chi and Zhu 2008) for a discussion.

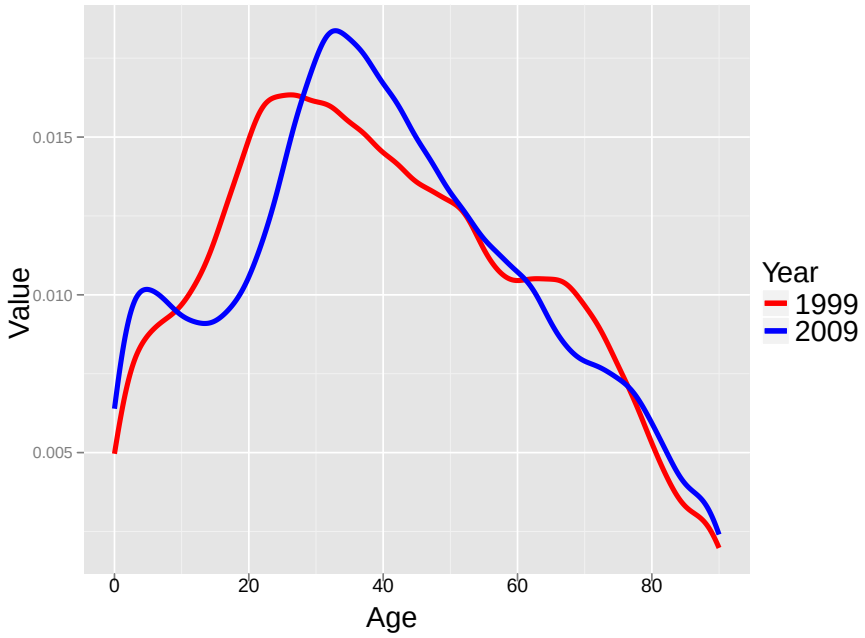

Fig. 1: Empirical density of Catalan population's age (1999 and 2009)

Therefore, although overall changes in age distribution (at the aggregate level) are expected to be important in incoming years, it is also necessary to analyse whether age structure differs from a spatial point of view in a cross section approach. In this sense, differences in age distribution across spatial units (municipalities, counties, provinces, states, etc.) have plenty of implications in terms of public policies affecting health, transport demand, education and social services. Concretely, typologies of public services demanded by individuals are strongly dependent on their age, so spatial units with different age distributions will differ in terms of the weight of public services demanded by them. Accordingly, it is of key importance to accurately explain how and why these age distributions vary across spatial units and, to the same extent, whether is it possible to predict future trends.

Apart from previous implications in terms of health care, as has been demonstrated by several scholars, age distribution has a wide range of other implications. Concretely, it influences population growth (Chi and Marcouiller, 2012; Rickman and Rickman, 2011; Partridge et al. 2009), population forecasting (Chi and Voss, 2011), suburbanization (Leichenko, 2001), migration patterns (Stillwell and García Coll, 2000; Nivalainen, 2004; Greenwood and Hunt, 1989; Greenwood, 1975), GDP levels

(Gómez and Hernández de Cos, 2008), income levels (Hashimoto and Tabata, 2010), fertility (Waldorf and Byun, 2005) and is an important determinant of residential preferences (Fuguitt and Brown, 1990). Therefore, policy makers should care about age distribution trends in order to properly account for all the consequences that derive from it.

This paper is structured as follows. In Section 1 contributions about population demographics are discussed and reviewed, as well as its relationship with spatial issues and whether the focus can be placed in other areas apart from population growth. In Section 2 the dataset and variables used in this paper are briefly presented. Section 3 explains and carries out the empirical application. Finally, Section 4 summarises the main conclusions.

2 Database and Variables

The data used in this article refer to local units (municipalities) in Catalonia for the period 1999-2009. On the one hand, data on the population classified into age intervals come from the Spanish Statistical Institute (INE), and on the other hand data referring to geographical characteristics of municipalities is taken from the Catalan Statistical Institute (IDESCAT). Moreover, the polygon map and related shape files have been obtained from the Catalan Cartographic Institute (ICC).

Formally, the data is divided into $N = 941$ municipalities[5]. Municipalities and years are denoted by the subscripts i and t, so that P_{it} is the total population in municipality i for the year t. $P_t = \sum_{i=1}^{N} P_{it}$ is the total population of Catalonia in year t. Population data for each municipality ans year is available in a frequency table with $n = 18$ age intervals. This structure is shown in Table 1.

Certainly, the available data is limited to age intervals, so that the exact age distribution has to be inferred. Here an estimation procedure is proposed. For each municipality i and year t[6], let p_k be the number of people whose age belongs to the k-th age category ($k \in 1, \ldots, n$). The first step is to create the vector $\vec{v}$, whose length is the total

[5] Data for five new municipalities (Gimenells i el Pla de la Font, Riu de Cerdanya, Sant Julià de Cerdanyola, Badia del Vallès and La Palma de Cervelló) have been left out due to lack of data.

[6] For the sake of simplicity, individual (i) and time (t) subscripts are omitted in this exposition.

Tab. 1: Original Structure of Demographic Data

Category (k)	Interval	Mid-point (x_k)	Category (k)	Interval	Mid-point (x_k)
1	[0,4]	2	10	[45,49]	47
2	[5,9]	7	11	[50,54]	52
3	[10,14]	12	12	[55,59]	57
4	[15,19]	17	13	[60,64]	62
5	[20,24]	22	14	[65,69]	67
6	[25,29]	27	15	[70,74]	72
7	[30,34]	32	16	[75,79]	77
8	[35,39]	37	17	[80,84]	82
9	[40,44]	42	18	[+85]	88

number of inhabitants, this is, $P = \sum_{k=1}^{n} p_k$. The elements of $\vec{v}$ are the mid-points x_k repeated p_k times, respectively. More formally, vector $\vec{v}$ can be expressed as the concatenation of n vectors $(v_1, \ldots, v_n)$, where each v_k is computed by means of the scalar product of mid-points by a ones vector $\vec{v}_k = x_k \cdot \mathbf{1}_{(1 \times p_k)}$. Once $\vec{v}$ has been computed, the next step is to estimate its density function, and this is done non-parametrically by means of a Kernel density estimation, which is a method for data smoothing based on a finite data sample. Analytically, the aim is to obtain an estimate of the density function $f(\cdot)$ that has generated the values in $\vec{v} = (v_1, ..., v_P)$. The shape of $f(\cdot)$ is approximated by:

$$\hat{f}(v) = \frac{1}{h \cdot P} \sum_{r=1}^{P} K\left(\frac{v - v_r}{h}\right). \tag{1}$$

In this function, $K(\cdot)$ is chosen to be a Gaussian density estimation, and h is the bandwidth or smoothing parameter, whose value has been selected so as to guarantee that $\hat{f}(\cdot)$ integrates to one, this is, $\int \hat{f}(v) = 1$. For each municipality, $\hat{f}(\cdot)$ is evaluated over a grid of 1,001 equally spaced nodes within the range $[0, 90]$[7]. The smoothness of the whole (estimated) density allows the computation of distributional parameters more efficiently than the analysis of a simple discrete histogram. These parameters, along with the remaining explanatory variables, are detailed below and summarised

[7] The upper limit of this range has been chosen rather arbitrarily, due to the fact that the data are right-censored because the last age interval is not closed (population above 85 years). Therefore, we are assuming that the end of the distribution is 90 years. Due to the low percentage of population above this age, this restriction is expected to have a negligible effect on the results.

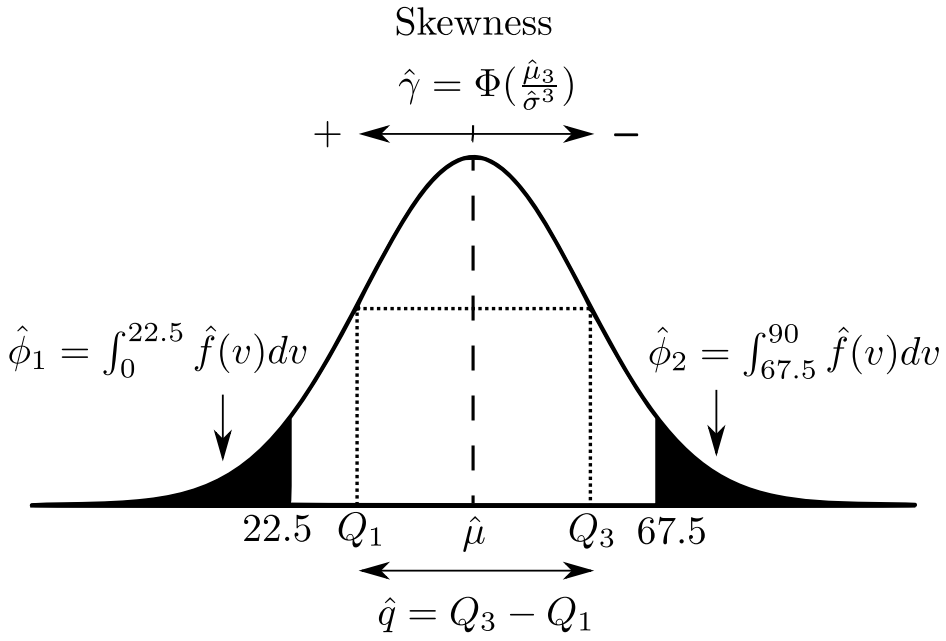

Fig. 2: Description of Distributional Parameters

in Table 2.

A) Distributional parameters

In order to capture and record the whole change in municipal age distributions, five different indicators have been selected. They are described next, and are graphically depicted and related to a density function in Figure 2.

1. **Mean** ($\hat{\mu}$): central tendency of the sample space, in this case simply the average of the age. This is a simple measure and indicates to what extent a municipality is, on average, young, middle-aged or old.

2. **Skewness** ($\hat{\gamma}$): measure of the asymmetry of the municipal age distribution. If the distribution is positively skewed, the mass of the distribution is concentrated on the left, this is, the largest share of the population correspond to age intervals below the mean, thus reflecting a young population. In the opposite case, a a negative skewness is the result of a mass distribution being concentrated on the right side of the mean, and it happens in municipalities with a high share of old people. Formally, the measure proposed here is based on the third standardized moment μ_3/σ^3 , where μ is the third moment about the mean and σ is

the standard deviation. This measure is centred around zero when there is no skewness/asymmetry, and take negative (positive) values if the mass of the distribution is on the right (left) side of the distribution. The fact that this measure takes both positive and negative values makes it difficult to compute variation rates between different years. To solve this shortcoming, a monotonic transformation is applied to the initial measure, so that the transformed values are bounded within the range $(0,1)$, the center of this range (this is0.5) being perfect symmetry. This transformation is the cumulative distribution function (CDF) of the Gaussian distribution, which applied to μ_3/σ^3 yields the desired skewness measure:

$$\hat{\gamma} = \Phi\left(\frac{\hat{\mu}_3}{\hat{\sigma}^3}\right) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\left(\frac{\hat{\mu}_3}{\hat{\sigma}^3}\right)} e^{-t^2/2} dt, \qquad \hat{\gamma} \in (0,1). \tag{2}$$

3. **Interquartile Range** ($\hat{q}$): also called midspread, it measures dispersion as the difference between the upper and lower quartiles, so that $\hat{q} = Q_3 - Q_1$. This way, it measures the difference between the ages including 50% of the probability mass. A high value of $\hat{q}$ indicates a highly dispersed population, whilst a low value points out a population where a large part of its distribution mass gathers around the mean.

4. **Kernel Densities ($\hat{\phi}$):** the aim of these empirical densities is to measure, on the one side, the population mass falling in the age interval $[0, 22.5]$. At the other side of the distribution, the population share above 67.5 years is measured. For each municipality, these two parameters are obtained by integrating the corresponding estimated density function $\hat{f}(v)$ over these age intervals:

$$\hat{\phi}_1 = \int_0^{22.5} \hat{f}(v) dv \tag{3}$$

$$\hat{\phi}_2 = \int_{67.5}^{90} \hat{f}(v) dv. \tag{4}$$

Considering that $\hat{f}(v)$ has been computed assuming the closed range $[0, 90]$, the first interval includes the first 25% of this range, whilst the second interval encompasses the last 25% of this range.

B) Human Variables

This group of covariates has three variables. POP is the logarithm of municipalities' total population, which proxies for size. DEN is the log of population density, computed as population divided by the urban area. POP and DEN are expressed in logarithms in order to reduce the effects of few outliers, one of them being the city of Barcelona, capital of Catalonia and by far the most densely populated municipality. Lastly, MIG stands for the share of immigrants. There are few sites where MIG takes the value zero, this is, there are no immigrants among the native population. The opposite situation is represented by some municipalities, where nearly 50% of inhabitants were born outside Spain.

C) Spatial Effects

One of the main assumptions made by contributions in spatial statistics and econometrics is that observations in space cannot be assumed to be mutually independent, so that the hypothesis made by the standard linear model stating that observations are not independent and identically distributed does not apply. In this sense, it makes perfect sense to think that characteristics of age distributions at the municipality level do not only depend upon variables from the same municipality, but also upon variables from neighbouring urban area. In order to take into account these effects, spatial lags are computed, analysed and included into the model's specification. These are computed by means of the product of a neighbourhood (or weights) matrix W with certain covariates, representing the average value of covariate values in neighbouring municipalities. In this analysis a distance weights matrix is considered such that $w_{ij} = (1/d_{ij})$, where d_{ij} is the distance between municipalities i and j, reflecting the idea that the closer these two municipalities are, the stronger the relationship between them.

D) Geographical Position

These are two variables controlling for the geographical position of each municipality. ALT is the average municipality's altitude with respect to sea level, which controls for accessibility. CL is a dummy variable with a value of one if the municipality is coastal, and zero otherwise.

Tab. 2: Description of Variables

Description	Variable
A) Distributional Parameters	
Mean	$\hat{\mu}$
Skewness	$\hat{\gamma} = \Phi\left(\frac{\hat{\mu}_3}{\hat{\sigma}^3}\right)$
Interquartile Range	$\hat{q} = Q_3 - Q_1$
Kernel Density in [0,22.5]	$\hat{\phi}_1 = \int_0^{22.5} \hat{f}(v)dv$
Kernel Density in [67.5,90]	$\hat{\phi}_2 = \int_{67.5}^{90} \hat{f}(v)dv$
B) Human Variables	
Log of Population	POP
Log of Population Density	DEN
Share of Immigrants	MIG
C) Spatial Effects	
Spatial Lag of POP	W-POP
Spatial Lag of DEN	W-DEN
Spatial Lag of MIG	W-MIG
D) Geographical Position	
Altitude	ALT
Coast Location	CL

Source: Own elaboration and Catalan Statistical Institute (IDESCAT).

3 Empirical Analysis

The empirical analysis conducted in this article is twofold. The first subsection consists in an exploratory analysis, whereby the main characteristics of the variables under study are summarised and studied, both at a univariate and a bivariate level. The second subsection is devoted to the specification and estimation of a spatial linear model, which is intended to unveil causality relationships between distribution characteristics (dependent variables) and human and geographic variables (covariates).

3.1 Exploratory Analysis

Table 3 shows the basic descriptive statistics of the variables under study. These variables are expressed in three ways: i) in levels for the year 1999, ii) in levels for the year 2009 and iii) in differences between these two years, thus showing the evolution over time. The first interpretation of these results is that there have not been dramatic changes over these 10 years. Regarding the distributional parameters, the most remarkable change corresponds to the interquartile range ($\hat{q}$), whose average value over all municipalities dropped by almost 3 years, which means that the bulk of local population's distributions is concentrating around the average age. This conclusion can also be drawn from Figure 1, where the aggregated distribution changes are depicted. The variables POB and DENS have also increased their mean values, showing a tendency of the population to be concentrated in the most densely populated areas, this is, a migration process from rural to urban areas. Besides, the variable MIG (share of immigrants) has also increased, which is a result of international migrations towards Spain. There are many theories explaining this phenomenon, one of the main ones being the availability of job positions in sectors like construction, tourism and services. Lastly, it is worth noting the high Moran's I statistic values, specifically in the variables in levels (1999 and 2009). This is due to the existing strong spatial patterns among municipal age distributions, and this result fully justifies the use of spatial statistics methods in the analysis.

Table 4 gathers information on the correlation among distributional parameters. The upper part of the table includes the Pearson correlation among distributional parameters in 1999, and it reveals that municipalities with a higher average age ($\hat{\mu}$) show negative skewness, this is, the distributional mass is located at the right side of the distribution. These high-aged municipalities also have higher dispersion around the mean ($\hat{q}$) and a higher percentage of the population above 67.5 years ($\hat{\phi}_2$). Interestingly, these relationships among distributional parameters also hold for the case of their differences in the period 1999-2009, shown in the central part of Table 4. Lastly, the bottom part of the table is devoted to the correlation among the initial value of parameters in 1999 and their differences between 1999 and 2009. All these correlation have negative values, and this result can be interpreted as a convergence process among different parts of the territory: populations with an initial higher mean age show a slower increase in this value, and the same applies for skewness

Tab. 3: Descriptive Statistics of Variables

	$\hat{\mu}$	$\hat{\gamma}$	$\hat{q}$	$\hat{\phi}_1$	$\hat{\phi}_2$	POB	DENS	MIG
A) Year 1999								
Mean	42,96	0,53	36,20	0,24	0,12	6,91	7,91	0,02
SD	4,25	0,07	3,49	0,05	0,05	1,63	0,94	0,03
Skewness	0,22	-0,42	-0,51	-0,42	0,39	0,70	0,38	3,33
Moran's I	0,59	0,54	0,51	0,51	0,58	0,50	0,22	0,36
B) Year 2009								
Mean	42,79	0,52	33,32	0,22	0,13	7,13	8,12	0,10
SD	4,32	0,06	3,19	0,05	0,05	1,69	0,90	0,07
Skewness	0,54	-0,55	0,06	-0,66	0,70	0,60	0,53	1,39
Moran's I	0,59	0,44	0,42	0,51	0,58	0,53	0,22	0,31
C) Difference 1999-2009								
Mean	-0,16	-0,01	-2,88	-0,02	0,00	0,22	0,21	0,08
SD	2,30	0,05	2,74	0,03	0,03	3,71	3,31	0,06
Skewness	-0,40	0,24	-0,01	0,43	-0,15	2,22	1,71	1,08
Moran's I	0,20	0,22	0,17	0,12	0,18	0,26	0,35	0,25

Note: Moran's I statistics are significant at the 1% level.

Tab. 4: Distributional Parameters' Correlations

Parameter	$\hat{\mu}$	$\hat{\gamma}$	$\hat{q}$	$\hat{\phi}_1$	$\hat{\phi}_2$
A) Parameters in levels 1999					
$\hat{\mu}$	1,00				
$\hat{\gamma}$	−0,94	1,00			
$\hat{q}$	0,67	−0,61	1,00		
$\hat{\phi}_1$	−0,94	0,88	−0,47	1,00	
$\hat{\phi}_2$	0,94	−0,85	0,73	−0,83	1,00
B) Parameters' difference 1999-2009					
$\hat{\mu}$	1,00				
$\hat{\gamma}$	−0,67	1,00			
$\hat{q}$	0,32	−0,34	1,00		
$\hat{\phi}_1$	−0,78	0,53	0,07	1,00	
$\hat{\phi}_2$	0,72	−0,40	0,44	−0,42	1,00
C) Correlation between levels and differences					
	−0,24	−0,53	−0,50	−0,39	−0,21

and dispersion.

A cartographic representation of the variables' spatial distribution is shown in Figures 3 and 4. In terms of population spatial distribution, Catalonia is quite heterogeneous, with most of people agglomerated around urban areas close to sea side in a strip of around 200 km long. Figures 3 and 4 show this pattern in a broader way, and there it is easy to identify three specific areas: i) sea-side locations with elder urban educated individuals, mainly working at service industries, ii) inner locations with younger rural less educated individuals, mainly engaged in both manufacturing (mainly) and service activities and, finally, iii) mountain locations (Pyrenees) in a transition pattern to elder neourban and educated individuals, increasingly working at leisure-oriented activities (i.e., tourism).

3.2 Spatial Regression

The aim of this section is to perform a confirmatory analysis in the form of spatial regressions. This is, the estimation of an equation for each distributional parameter

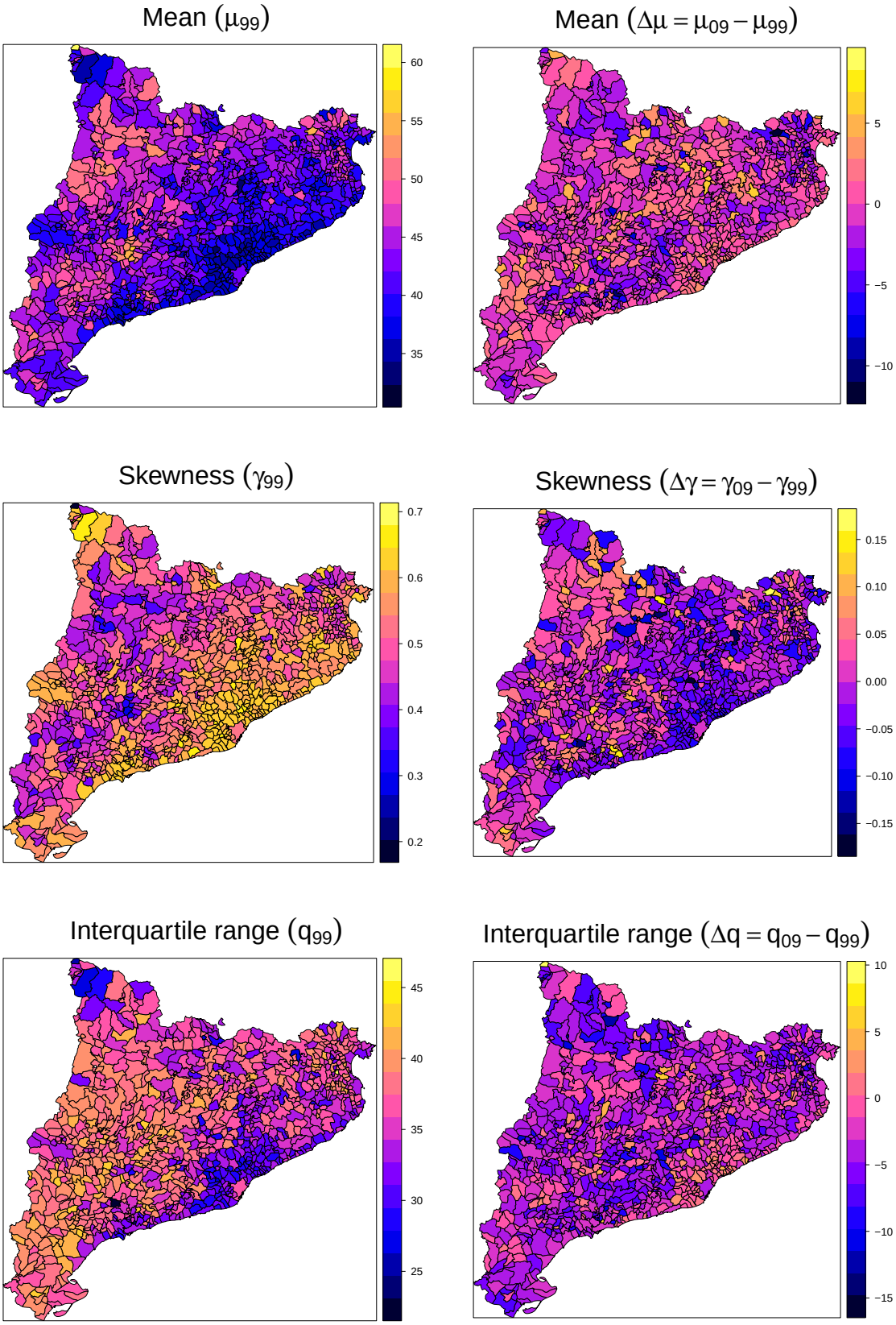

Fig. 3: Spatial Distribution of Parameters and Population

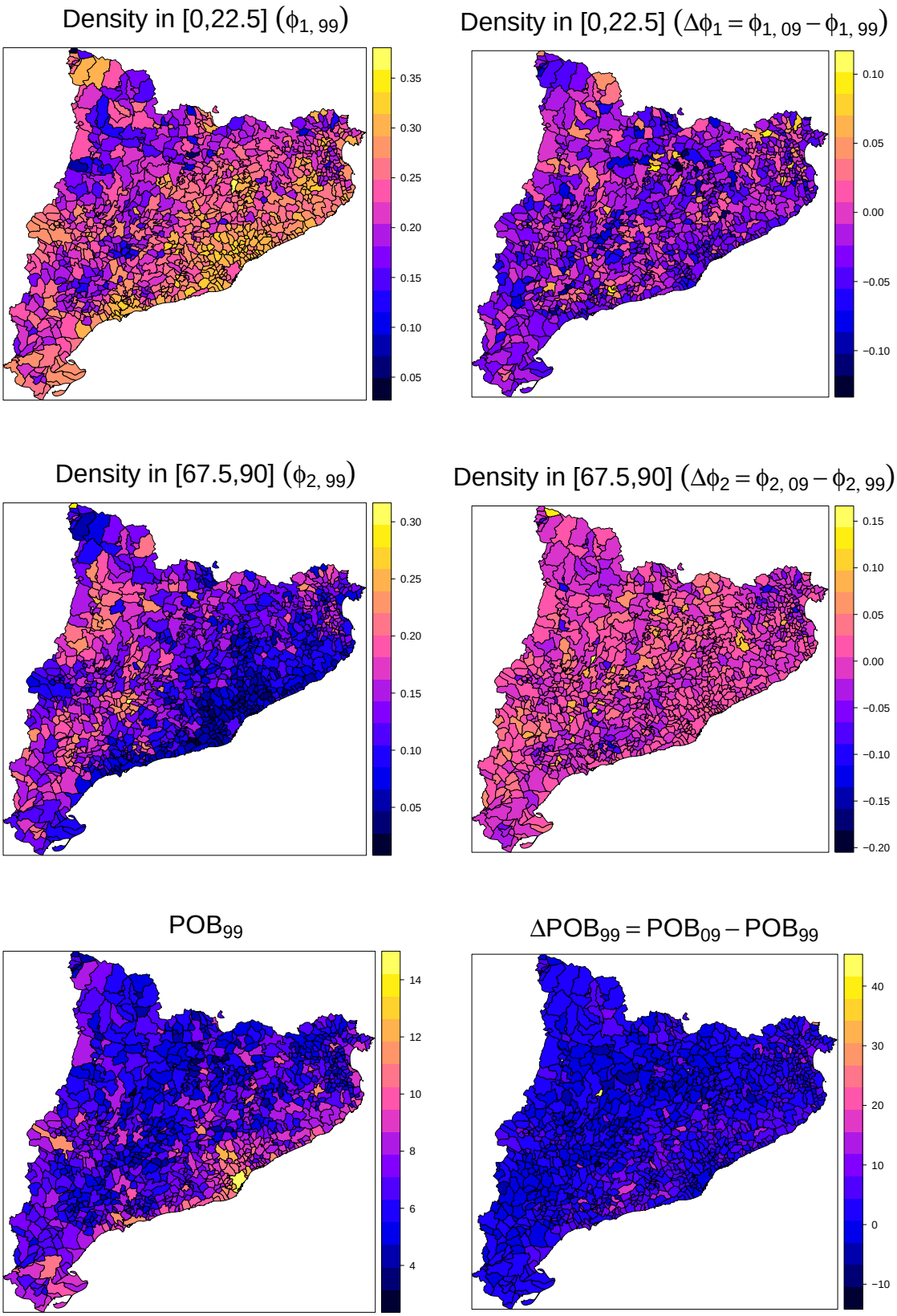

Fig. 4: Spatial Distribution of Parameters and Population (cont.)

is intended to shed light upon the effects of local covariates upon these parameter's changes over the period 1999-2009. Formally, the models to be estimated take the form

$$\triangle y_i = \alpha + x_i^{'}\beta + \epsilon_i, \qquad i = 1, \ldots, N. \tag{5}$$

In specification (5), the dependent variable is defined as $\triangle y_i = y_{2009,i} - y_{1999,i}$, and it is to be estimated for each distributional parameter, i.e. $y = \{\hat{\mu}, \hat{\gamma}, \hat{q}, \hat{\phi}_1, \hat{\phi}_2\}$. An obvious concern regarding the estimation of (5) is the presence of spatial effects. These effects are of different nature, but they have in common that they render classical inference unreliable[8]. A possible spatial effect is that the variance of the model is not distributed randomly across the space, and thus follows some pattern. Besides, the hypothesis of independence among sample observations is often violated, and thus causality relationships jump over each observation's borders and affect neighbouring observations.

In a regression framework, a common procedure to detect spatial effects is the analysis of estimated residuals' distribution. Table 5 shows the results of several Lagrange Multiplier (LM) diagnostics for spatial dependence. Specifically, these tests try to reveal whether the data generating process is a spatial lag model, a spatial error model, or both[9]. As it is shown in these three diagnostics, neither of these two ways of modelling spatial dependence can be discarded.

The estimation of the five empirical models relies on the implementation of Kelejian and Prucha (1999) estimator for the autoregressive parameter in a spatial model. It uses a Generalised Moments (GM) approach to optimise the autoregressive parameter (λ) and the variance of the model (σ^2) jointly. Regarding the covariates included in the model, only a subset of the variables described in Table 2 have been included. This choice is due to the existing multicollinearity existing if all variables were to be included simultaneously, and is has been made following a variance

[8] For seminal works on this matter, see Anselin (1988) and Cressie (1993).

[9] In spatial econometrics, a spatial lag model includes a spatial lag of the dependent variable, whilst in a spatial error model it is the disturbances that exhibit spatial dependence. For a detailed explanation of these models, see Anselin (1988), and for a description of diagnostics tests for spatial dependence, see Anselin et al. (1996).

Tab. 5: Lagrange Multiplier diagnostics for spatial dependence

	LM test for spatial error		LM test for spatial lag		LM test for spatial ARMA	
Dependent variable	Statistic	P-value	Statistic	P-value	Statistic	P-value
Mean ($\hat{\mu}$)	22,51	0,00	23,73	0,00	23,86	0,00
Skewness ($\hat{\gamma}$)	43,81	0,00	48,61	0,00	48,66	0,00
InterQ. Range ($\hat{q}$)	5,90	0,02	12,21	0,00	21,50	0,00
Density in [0,22.5]	12,60	0,00	12,30	0,00	12,60	0,00
Density in [67.5,∞]	10,19	0,00	11,91	0,00	12,15	0,00

inflation factor (VIF) analysis.

The regression results are shown in Table 6, and from these results one can better understand how spatial specific characteristics shape age distribution at a local level. One of the main results is that the variable POB_{99} has a positive effect on the ageing of municipalities, this is, those municipalities with a higher population in 1999 have shown an increase in the age mean as well as a decrease in the skewness (this is, the mass of the distribution has moved to the right). Interestingly, the increase of population ($\triangle$POB) has the opposite effect, i.e. those municipalities showing population growth have experience a rejuvenation in their age distribution. This result also applies for the variables $\triangle$MIG, this is, municipalities receiving immigration experience either a lower increase or a decrease of $\hat{\mu}$, an increase in the skewness, a decrease in the distribution dispersion ($\hat{q}$) and an increase of the population below 22.5 years ($\hat{\phi}_1$). The explanation for this result could be that increases in population are usually caused by increases in natality from young migrants. The spatial lags of POB and MIG have the same sign as these variables and their coefficients are significant, revealing the existence of both a spatial structure and spatial interactions between municipalities. Lastly, the coefficient of the dummy variable CL (coast location) is associated with an increase of $\hat{\mu}$ and $\hat{q}$, as well as a decrease in skewness ($\hat{\gamma}$). This result can be easily explained according to spatial distribution of highly demanded natural amenities (e.g., proximity to sea side, which implies accessibility to beaches, recreational water activities and so on) and profile of such individuals able to pay higher housing prices at these amenity-intensive sites, many of which are elder people.

Tab. 6: GM Spatial Simultaneous Autoregressive Error Model: Estimation Results

Dep. Var.	Mean ($\triangle\hat{\mu}$)		Skewness ($\triangle\hat{\gamma}$)		InterQ. Range ($\triangle\hat{q}$)		Density in $[0,22.5]$ ($\triangle\hat{\phi}_1$)		Density in $[67.5,90]$ ($\triangle\hat{\phi}_2$)	
	Estimate	Std. Error	Estimate	Std. Error	Estimate	Std. Error	Estimate	Std. Error	Estimate	Std. Error
CONSTANT	0,810	(0,8812)	0,253	(1,9214)	−3,889	(1,1346)***	−0,019	(0,0126)	0,012	(0,0112)
POB_{99}	0,231	(0,0573)***	−0,727	(0,1231)***	0,233	(0,0757)***	−0,004	(0,0008)***	0,003	(0,0007)***
△POB	−0,224	(0,0185)***	0,148	(0,0395)***	−0,065	(0,0248)***	0,002	(0,0003)***	−0,003	(0,0002)***
△MIG	−10,634	(1,1537)***	27,786	(2,4926)***	−13,216	(1,5032)***	0,064	(0,0166)***	−0,125	(0,0148)***
W-POB_{99}	0,173	(0,0827)**	−0,490	(0,1814)***	0,354	(0,1065)***	−0,002	(0,0012)**	0,002	(0,0010)
W-$DENS_{99}$	−0,269	(0,1108)**	0,672	(0,2412)***	−0,158	(0,1430)	0,004	(0,0016)**	−0,002	(0,0014)
W-△POB	−0,131	(0,0343)***	−0,010	(0,0760)	−0,136	(0,0436)***	0,001	(0,0005)*	−0,002	(0,0004)***
ALT	0,507	(0,2805)*	−0,599	(0,6355)	−0,530	(0,3424)	−0,005	(0,0039)	0,002	(0,0034)
CL	1,751	(0,2851)***	−2,576	(0,6265)***	1,249	(0,3624)***	−0,016	(0,0041)***	0,012	(0,0036)***
λ		0,221		0,288		0,118		0,170		0,148
LR Test		13,573***		23,235***		33,705*		74,492***		66,875**
ML Residual σ^2		3,491		15,907		61,788		0,001		0,001
AIC		3886		5325		4410		−4080		−4299

Note: ***, ** and * denote significance at the 1%, 5% and 10%, respectively.

4 Summary and Conclusions

In this paper spatial trends and implications of age structure have been addressed, which is a topic from population economics that has not received yet sufficient attention according to its importance. Besides, the existence of spatial heterogeneity in terms of age distribution and, consequently, the importance of properly taking into account these asymmetries have been highlighted and discussed. As for technicalities of empirical application, these age distribution heterogeneities have been approached by analysing and modelling several distribution parameters like mean, skewness, dispersion and kernel densities of both tails. Overall, this analysis roughly supports the prior expectations of such inequalities made in this paper, which fits perfectly into previous empirical evidence (both from the area analysed in this paper as well as from other similar areas) regarding spatial heterogeneity in terms of population distribution.

Further extensions of this research should point into specific determinants of age distribution, as well as better identify causality relationship among different economic activities and whether there is still room for policy measures affecting the shape of the distribution in the long term.

References

Anselin, L. (1988). *Spatial econometrics: methods and models*. London: Kluwer.

Anselin, L., Bera, A., Florax, R., and Yoon, M. (1996). Simple diagnostic tests for spatial dependence. *Regional science and urban economics*, 26(1):77–104.

Arauzo-Carod, J. and Teruel-Carrizosa, M. (2005). An urban approach to firm entry: The effect of urban size. *Growth and Change*, 36(4):508–528.

Chi, G. and Marcouiller, D. (2011). Isolating the effect of natural amenities on population change at the local level. *Regional Studies*, 45(4):491–505.

Chi, G. and Ventura, S. (2011). An integrated framework of population change: Influential factors, spatial dynamics, and temporal variation. *Growth and Change*, 42(4):549–570.

Chi, G. and Voss, P. (2011). Small-area population forecasting: borrowing strength across space and time. *Population, Space and Place*, 17(5):505–520.

Chi, G. and Zhu, J. (2008). Spatial regression models for demographic analysis. *Population Research and Policy Review*, 27(1):17–42.

Cressie, N. (1993). *Statistics for spatial data*. Wiley (New York).

Fossett, M. (2005). Urban and spatial demography. *Handbook of population*, pages 479–524.

Fritsch, M., Brixy, U., and Falck, O. (2006). The effect of industry, region, and time on new business survival–a multi-dimensional analysis. *Review of Industrial Organization*, 28(3):285–306.

Fuguitt, G. and Brown, D. (1990). Residential preferences and population redistribution: 1972–1988. *Demography*, 27(4):589–600.

Gómez, R. and Hérnandez de Cos, P. (2008). The importance of being mature: the effect of demographic maturation on global per capita gdp. *Journal of population economics*, 21(3):589–608.

Greenwood, M. (1975). Research on internal migration in the united states: a survey. *Journal of Economic Literature*, 13(2):397–433.

Greenwood, M. and Hunt, G. (1989). Jobs versus amenities in the analysis of metropolitan migration. *Journal of Urban Economics*, 25(1):1–16.

Hashimoto, K. and Tabata, K. (2010). Population aging, health care, and growth. *Journal of Population Economics*, 23(2):571–593.

Keilman, N. (1997). Ex-post errors in official population forecasts in industrialized countries. *Journal of official statistics*, 13:245–278.

Kelejian, H. and Prucha, I. (1999). A generalized moments estimator for the autoregressive parameter in a spatial model. *International economic review*, 40(2):509–533.

Leichenko, R. (2001). Growth and change in us cities and suburbs. *Growth and Change*, 32(3):326–354.

Nivalainen, S. (2004). Determinants of family migration: short moves vs. long moves. *Journal of Population Economics*, 17(1):157–175.

Nyström, K. (2007). An industry disaggregated analysis of the determinants of regional entry and exit. *The Annals of Regional Science*, 41(4):877–896.

Parr, J. (2002). Missing elements in the analysis of agglomeration economies. *International Regional Science Review*, 25(2):151–168.

Partridge, M., Rickman, D., Ali, K., and Olfert, M. (2009). Do new economic geography agglomeration shadows underlie current population dynamics across the urban hierarchy?*. *Papers in Regional Science*, 88(2):445–466.

Rickman, D. and Rickman, S. (2011). Population growth in high amenity nonmetropolitan areas: what's the prognosis? *Journal of Regional Science*, 51(5):863–879.

Stillwell, J. and Coll, A. (2000). Inter-provincial migration of the spanish workforce in 1988 and 1994. *Regional Studies*, 34(7):693–711.

Tiefelsdorf, M. (2000). *Modelling spatial processes*. Springer Berlin/Heidelberg.

Tobler, W. (1970). A computer movie simulating urban growth in the detroit region. *Economic geography*, 46:234–240.

Uhlenberg, P. (2005). Demography of aging. *Handbook of Population*, pages 143–167.

Voss, P. (2007). Demography as a spatial social science. *Population Research and Policy Review*, 26(5):457–476.

Waldorf, B. and Byun, P. (2005). Meta-analysis of the impact of age structure on fertility. *Journal of Population Economics*, 18(1):15–40.

White, M. and Lindstrom, D. (2006). Internal migration. *Handbook of population*, page 311.