Empirical Evidence of Population Concentration in Spain, 1900-2001

Goerlich, Francisco José and Mas, Matilde

Instituto Valenciano de Investigaciones Económicas, Universitat de València

2008

Empirical Evidence of Population Concentration in Spain, 1900-2001.

Francisco J. Goerlich

(Corresponding author) Universidad de Valencia, Departamento de Análisis Económico, Campus de Tarongers, Avda. de Tarongers s/n, 46022-Valencia (Spain) Phone: 963 82 82 46 / Fax: 963 82 82 49 E-mail: Francisco.J.Goerlich@uv.es and Instituto Valenciano de Investigaciones Económicas, C/ Guardia Civil, 22, Esc. 2, 1º, 46020 Valencia-Spain

Matilde Mas

Universidad de Valencia, Departamento de Análisis Económico, Campus de Tarongers, Avda. de Tarongers s/n, 46022-Valencia (Spain) and Instituto Valenciano de Investigaciones Económicas, C/ Guardia Civil, 22, Esc. 2, 1º, 46020 Valencia-Spain. Phone: 96-319 00 50 / Fax: 96 319 00 55. E-mail: matilde.mas@ivie.es

ABSTRACT

This short paper examines the evolution of the population density in Spain during the 20th century. Using a homogeneous database of the population at a municipal level – elaborated from the eleven censuses carried out between 1900 and 2001– the paper looks at the general characteristics of population concentration from various perspectives. Focusing on population density, we present empirical evidence that supports the hypothesis that, over time, Spain's population has undergone progressive concentration, a process that has not finished in the present days. Its main contribution is to offer quantitative support for phenomena which have already been well documented by specialists in more general terms.

Key Words: Population, Municipalities, Census, Agglomeration, Population Density.

JEL Classification: J10, J11

Correspondence: Francisco J. Goerlich, Universidad de Valencia, Departamento de Análisis Económico, Campus de Tarongers, Av de Tarongers s/n, 46022-Valencia. Spain. E-mail: Francisco.J.Goerlich@uv.es.

Acknowledgements: The authors are grateful to Pilar Chorén for help in processing the data, to the editors and to two anonymous referees for comments. Also to projects of the Ministerio de Ciencia y Tecnología/FEDER, SEC2005-02776, ECO2008-03813 and Fundación BBVA-Ivie research programme for financial assistance. Results mentioned in the text but not reported can be requested at Francisco.J.Goerlich@uv.es.

I. INTRODUCTION

This short paper examines localization patterns of the population in Spain at the municipal level during the 20th century. Over the course of this period the country underwent far-reaching social, economic and demographic changes. Some hesitant steps towards the industrialization process had been taken during the 19th century and the first three decades of the 20th, but the second half of the latter century would see the industrialization "miracle" spread to all parts of the country. The accompanying change in the "quality of life" was to affect each member of the population in varying and multidimensional ways.

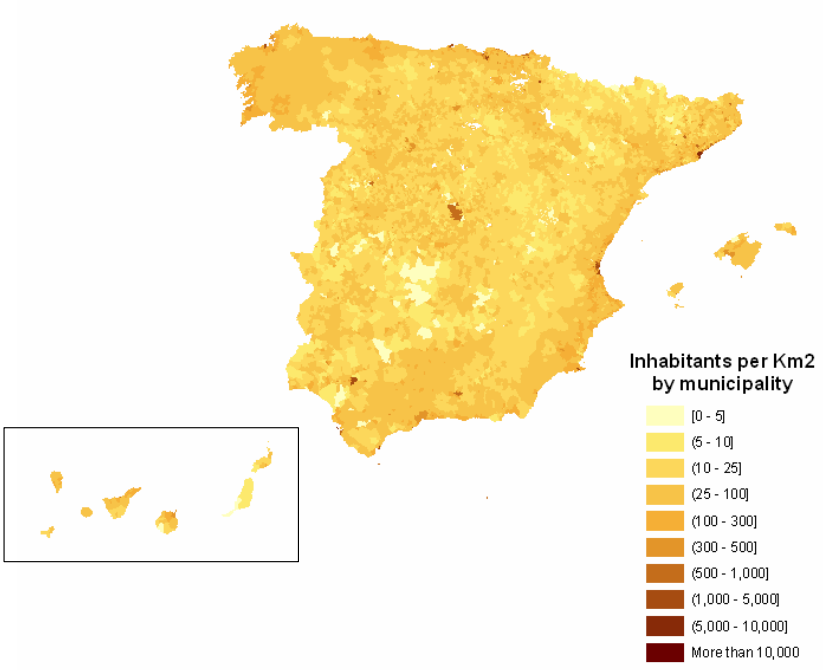
Two simple examples can serve to show the extent of this change. Life expectancy more than doubled over the course of the century, going from 34.76 years in 1900 (*National Statistical Institute* (INE) 1952) to 79.44 years at the end of the century (in 2000, http://www.ine.es). The income *per capita* increased by a factor of 10 in real terms, going from just under 250,000 pesetas in 1900 (valued in pesetas of 1995) to more than 2,000,000 pesetas in 2000 (Prados de la Escosura 2003), despite the fact that the population more than doubled during this period.

These changes did not come about in a uniform way as far as time and location are concerned. Over the course of the 20th century the population, basis of all economic activity, became increasingly concentrated (Zoido and Arroyo 2004: De Cos and Reques 2005). This uneven distribution of the population over the territory was already evident in 1900 (Map 1). The process of development and industrialization in the Spanish society has intensified this phenomenon to the extreme (Map 2). However, it must be pointed out that the economic development of the 20th century did not create an urban system out of nothing, rather it operated on a network of already existing cities that had been formed in the 18th and 19th centuries (and in some cases much earlier). What has changed in a radical way is the intensity of the process; due not only to the increase in population figures but fundamentally also in relative terms. The maps presented illustrate well these changes.

The Spanish experience has therefore been similar to that of the large European cities (de Vries 1984; Hohenberg and Lees 1995), the difference being that the process

was, to a certain extent, delayed. Our calculations corroborate those made on a regional scale by Ayuda, Collantes and Pinilla (2005, 2007). While the time perspective in these studies is more extended, the use of a larger unit of geographical analysis tempers the process of spatial concentration of the population. Martí-Hennenberg (2005) obtained similar results at a regional European level.

MAP 1: **Population Density. 1900**

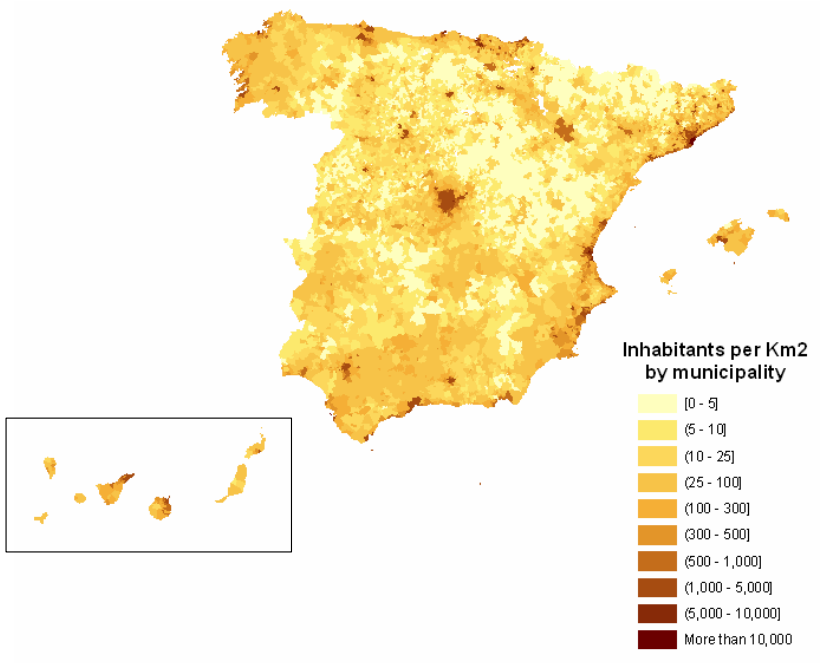


Source: INE, IGN and own elaboration.

An exhaustive description of the process of population localization in Spain at a municipal level during the 20th century is given in Goerlich, Mas, Azagra and Chorén (2006). That study charts the different rhythms in the progressive shrinking of the small villages (the rural environment) in contrast to the growth of medium-sized cities and the boom of the larger ones (the metropolitan areas). This shrinking and growth followed a marked spatial pattern: the interior underwent a process of depopulation and the periphery one of densification. Madrid, the state capital, is the most notable exception in this process of the movement of the population to the coast (Maps 1 and 2). This should come as no surprise, given that capital cities have always had their own demographic

dynamics (Ades and Glaeser 1995). However, we can observe that other regional capitals, clearly visible on the map of 2001, also add their little note of colour as poles of attraction in an "almost empty" interior. (Map 2)

MAP 2: **Population Density. 2001**



Source: INE, IGN and own elaboration.

This paper tries to probe somewhat deeper into these general patterns of population concentration. As is the case with other authors (Ayuda, Collantes and Pinilla 2005, 2007), our interest is not focussed only on large urban agglomerations or cities, as is the case with a large amount of the literature (van der Woude, de Vries and Hayami 1990; Reher 1990; Eaton y Eckstein 1997; Black and Henderson 2003). On the contrary, our analysis takes into account the small villages, of little importance in terms of the volume of population but abundant in number and surface area. These make up part of our wide-ranging geography. Their dynamics and their relation to the regional capitals and large urban areas are a reality that is impossible to ignore.

The structure of the paper is as follows: the next section gives details of the database used; Section 3 outlines the basic features of the process of population concentration at municipal level in Spain during the 20th century; Section 4 presents four extensions of the preceding analysis; and Section 5 offers some brief conclusions.

II. DATA SOURCES

As a primary source of information this paper uses the municipal population figures of the eleven Spanish censuses that cover the period from 1900 to 2001 (the date of the most recent available census). Among the administrative divisions of national territory, the municipalities are the smallest units that have clearly designated boundaries, so population density can be calculated.

The municipal structure of Spain has undergone substantial changes over the course of the 20th century. The number of municipalities has reduced considerably, from 9,267 in 1900 to the 8,108 listed in the 2001 census. Moreover, numerous modifications have been made to the municipal structure due to fusions, segregations and other kinds of alterations within the municipalities between the censuses.

As a result, Goerlich, Mas, Azagra and Chorén (2006) produced homogenised municipal populations statistics starting from *two basic principles*: (*i*) Populations are defined on the basis of a territorial criterion, the municipal boundaries, and (*ii*) the criterion that determines these territories is the existing municipalities recorded in the 2001 census.

Hence, this paper uses information on the homogenised resident municipal populations from the censuses conducted between 1900 and 2001, where this homogeneity is based on the municipal boundaries in existence in the 2001 census. Data on municipal surface is taken from the *National Geographical Institute* (IGN) municipal database. It is worth noting that Martí-Henneberg (2005) faces similar problems of homogeneity in comparing historical population density data in Europe at a regional level and he adopts a similar approach.

III. POPULATION CONCENTRATION: SOME BASIC FACTS

The population concentration shown in maps 1 and 2 can be quantitatively summarised in various ways. An index of concentration, with an interesting decomposability property, is the (second) Theil (1967) index, or mean logarithmic deviation, T^*, which can be applied to variables other than the one with which it is most frequently associated, i.e. the personal distribution of income. In our case we use the Theil index to study the population concentration as measured by population density at a municipal scale.

If d_i is the population density of municipality i, $d_i = \frac{Inhab_i}{Km_i^2}$, and s_i is its surface share, $s_i = \frac{Km_i^2}{\Sigma_{j=1}^{n} Km_j^2}$, we can define the Theil index, T^*, as

$$T^* = \sum_{i=1}^{n} s_i \log \frac{\mu}{d_i} = \log \frac{\mu}{\tilde{\mu}} \qquad (1)$$

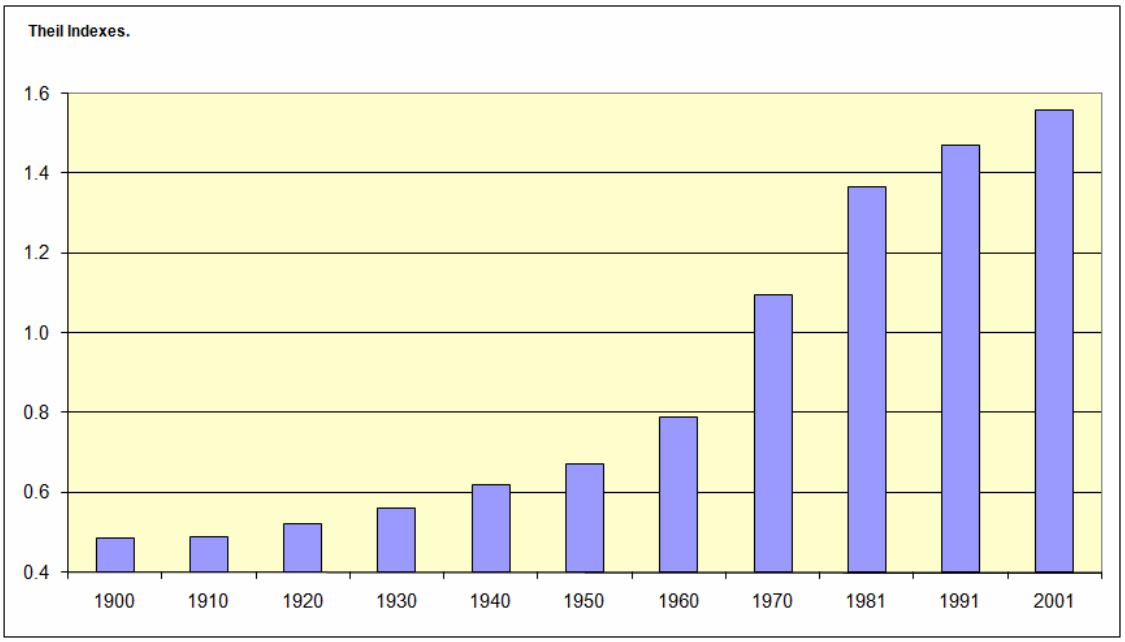
where μ is the arithmetic mean, $\mu = \Sigma_{i=1}^{n} s_i d_i$, $\tilde{\mu}$ is the geometric mean, $\log \tilde{\mu} = \Sigma_{i=1}^{n} s_i \log d_i$ and n is the number of municipalities. The mean logarithmic deviation has a minimum value of zero (if there is an even spatial distribution of population), but has no upper limit, so that a greater degree of concentration appears as a higher value on the index, but this is unbounded from above.

Figure 1 shows the evolution of T^* for Spain, and offers two pieces of highly relevant information. The first is the sustained increase in population concentration over the course of the 20th century. The second is the way in which the period from 1950-1981 (and in particular the decade of the 1960´s) stands out as that of the highest polarization. This is in fact the period of highest growth in Spain.

In this sense, it is interesting to relate the Theil index with the income *per capita*, also shown in Figure 1, to see that both indicators follow the same general pattern. In fact a simple correlation coefficient between both variables is as high as 0.96 and the highest growth is achieved in the 1960´s, which is when we find the highest increase in population concentration. The conclusion seems obvious; the concentration of people follows the economic development as the New Economic Geography theories, based on

increasing returns, predict (Fujita, Krugman and Venables 1999; Pons, Paluzie, Silvestre and Tirado 2007).

Figure 1. Theil index. Spain. Population density (Inhabitants/Km^2). 1900-2001.



Theil index. Spain.

	1900	1910	1920	1930	1940	1950	1960	1970	1981	1991	2001
Population density.	0.484	0.488	0.520	0.560	0.619	0.670	0.787	1.093	1.366	1.469	1.559

Soucs: INE, IGN and own calculations.

Real GDP *per capita* (1995 = 100). Spain.

	1900	1910	1920	1930	1940	1950	1960	1970	1980	1990	2000
Level	13.15	13.95	15.77	19.22	15.11	16.29	23.60	46.19	68.75	91.82	119.45
Growth by decade		5.91%	12.26%	19.78%	-24.06%	7.52%	37.07%	67.15%	39.77%	28.94%	26.31%

Sources: Prados de la Escosura (2003), Table A.13.5, p.- 682. Growth rates are calculated as logaritmic differences.

This tendency towards increasing concentration is not generally found at regional level. Table 1 shows the corresponding Theil indexes for provinces. As can be observed, the majority of provinces present minimum values in either 1900 or 1910. In addition, the majority also reach the maximum value in 2001.

While the level of concentration continued to rise throughout the 20th century in the majority of regions, those of Illes Balears, Barcelona, Madrid, Las Palmas and Vizcaya reached the maximum index value in 1981. From that year onwards the Theil index falls continuously. All these cases represent rich provinces of important economic development, so this is an indication of the level of saturation of the regions and the relative decline of their respective metropolitan areas. From other point of view, it can

be argued that this is an indication that municipal boundaries are not the right scale to study the population concentration in certain cases.

Table 1. Theil index (Mean logarithmic deviation). Provinces. Population density. 1900 - 2001.

	Province	1900	1910	1920	1930	1940	1950	1960	1970	1981	1991	2001
01	Álava	*0.271*	0.279	0.294	0.331	0.402	0.425	0.644	1.169	1.581	**1.703**	1.644
02	Albacete	*0.074*	0.076	0.081	0.093	0.110	0.121	0.141	0.253	0.409	0.503	**0.611**
03	Alicante/Alacant	*0.224*	0.234	0.267	0.294	0.338	0.376	0.443	0.581	0.705	0.754	**0.769**
04	Almería	0.228	0.225	*0.220*	0.225	0.275	0.291	0.350	0.533	0.742	0.903	**1.031**
05	Ávila	0.136	*0.134*	0.136	0.144	0.157	0.167	0.190	0.266	0.445	0.541	**0.653**
06	Badajoz	0.247	0.242	0.222	0.210	0.209	0.205	*0.204*	0.275	0.352	0.409	**0.474**
07	Balears (Illes)	*0.337*	0.344	0.359	0.365	0.440	0.475	0.533	0.660	**0.734**	0.702	0.693
08	Barcelona	*1.372*	1.388	1.461	1.632	1.728	1.873	2.070	2.366	**2.542**	2.493	2.329
09	Burgos	0.188	*0.188*	0.198	0.233	0.276	0.323	0.431	0.740	1.091	1.234	**1.326**
10	Cáceres	0.278	0.282	0.257	0.255	0.250	0.240	*0.231*	0.274	0.377	0.459	**0.531**
11	Cádiz	0.558	0.516	0.571	*0.506*	0.538	0.574	0.604	0.704	0.821	**0.848**	0.837
12	Castellón/Castelló	*0.287*	0.290	0.315	0.358	0.428	0.509	0.653	0.955	1.249	1.396	**1.525**
13	Ciudad Real	0.420	0.414	0.421	0.399	0.397	*0.391*	0.401	0.519	0.640	0.699	**0.778**
14	Córdoba	0.390	*0.378*	0.386	0.397	0.434	0.450	0.428	0.552	0.708	0.758	**0.807**
15	Coruña (A)	*0.248*	0.252	0.264	0.273	0.298	0.339	0.380	0.455	0.573	0.640	**0.701**
16	Cuenca	0.109	*0.104*	0.110	0.122	0.130	0.143	0.169	0.282	0.446	0.537	**0.615**
17	Girona	*0.379*	0.413	0.423	0.477	0.523	0.564	0.690	0.923	1.163	1.244	**1.272**
18	Granada	0.371	*0.358*	0.374	0.383	0.416	0.414	0.447	0.589	0.818	0.937	**1.038**
19	Guadalajara	*0.136*	0.138	0.146	0.160	0.184	0.192	0.226	0.483	0.918	1.079	**1.324**
20	Guipúzcoa	*0.281*	0.329	0.425	0.480	0.519	0.572	0.656	0.777	0.881	0.904	**0.919**
21	Huelva	*0.423*	0.439	0.433	0.435	0.449	0.452	0.474	0.610	0.780	0.857	**0.909**
22	Huesca	*0.174*	0.179	0.187	0.190	0.205	0.252	0.324	0.536	0.721	0.787	**0.804**
23	Jaén	0.249	0.235	0.239	*0.217*	0.218	0.226	0.225	0.291	0.382	0.433	**0.479**
24	León	*0.188*	0.192	0.204	0.234	0.294	0.348	0.430	0.606	0.819	0.980	**1.103**
25	Lleida	*0.346*	0.386	0.433	0.449	0.468	0.552	0.655	0.925	1.136	1.210	**1.231**
26	Rioja (La)	*0.345*	0.360	0.410	0.463	0.539	0.591	0.716	1.078	1.440	1.543	**1.645**
27	Lugo	0.071	0.077	*0.062*	0.067	0.083	0.106	0.133	0.185	0.252	0.346	**0.429**
28	Madrid	1.255	*1.230*	1.406	1.523	1.717	1.795	1.998	2.270	**2.322**	2.159	1.896
29	Málaga	0.311	*0.301*	0.314	0.348	0.409	0.437	0.473	0.629	0.877	0.965	**1.038**
30	Murcia	0.336	*0.325*	0.327	0.327	0.368	0.390	0.448	0.512	0.586	0.634	**0.654**
31	Navarra	0.280	*0.279*	0.298	0.329	0.394	0.463	0.597	0.940	1.178	1.249	**1.335**
32	Ourense	0.213	*0.210*	0.219	0.226	0.241	0.279	0.293	0.363	0.456	0.600	**0.723**
33	Asturias	*0.256*	0.282	0.328	0.371	0.436	0.510	0.627	0.845	1.035	1.135	**1.226**
34	Palencia	0.157	*0.157*	0.173	0.202	0.249	0.283	0.351	0.529	0.750	0.883	**1.011**
35	Palmas (Las)	*0.940*	1.038	1.095	1.138	1.238	1.324	1.302	1.548	**1.607**	1.435	1.174
36	Pontevedra	*0.177*	0.207	0.241	0.271	0.312	0.336	0.393	0.507	0.645	0.712	**0.796**
37	Salamanca	*0.288*	0.294	0.302	0.343	0.397	0.436	0.488	0.683	0.926	1.087	**1.206**
38	Sta. Cruz de Tenerife	*0.293*	0.358	0.315	0.347	0.371	0.456	0.494	0.620	0.740	**0.754**	0.725
39	Cantabria	*0.442*	0.488	0.527	0.577	0.643	0.685	0.821	1.084	1.351	1.472	**1.549**
40	Segovia	0.122	*0.115*	0.118	0.130	0.148	0.166	0.204	0.365	0.556	0.639	**0.711**
41	Sevilla	0.530	*0.522*	0.562	0.545	0.604	0.651	0.695	0.891	1.055	1.097	**1.131**
42	Soria	*0.068*	0.070	0.071	0.079	0.095	0.115	0.153	0.338	0.565	0.692	**0.817**
43	Tarragona	0.311	*0.294*	0.323	0.362	0.406	0.466	0.559	0.794	1.002	1.050	**1.106**
44	Teruel	*0.117*	0.131	0.134	0.145	0.165	0.180	0.222	0.334	0.519	0.605	**0.676**
45	Toledo	0.229	0.219	0.224	*0.213*	0.218	0.231	0.256	0.360	0.505	0.564	**0.644**
46	Valencia/València	*0.774*	0.782	0.802	0.915	1.049	1.099	1.202	1.487	1.687	1.750	**1.758**
47	Valladolid	0.286	*0.283*	0.297	0.350	0.408	0.438	0.568	0.946	1.307	1.425	**1.505**
48	Vizcaya	*0.658*	0.690	0.795	0.871	0.939	1.023	1.270	1.583	**1.781**	1.772	1.729
49	Zamora	0.167	0.166	*0.156*	0.159	0.187	0.213	0.243	0.330	0.463	0.579	**0.683**
50	Zaragoza	0.336	*0.329*	0.354	0.394	0.472	0.540	0.680	1.055	1.342	1.443	**1.525**
	Spain	*0.484*	0.488	0.520	0.560	0.619	0.670	0.787	1.093	1.366	1.469	**1.559**

Notes: The *minimum* value for each province is shown in *italics*
The **maximum** value for each province is shown in **bold**.
Sources: INE, IGN and own calculations.

In 1991 Álava, Cadiz and Santa Cruz de Tenerife reached their point of maximum concentration. It is interesting to observe that all of these provinces (except Álava and Madrid) are on the coast, which has seen a marked increase in its tourism industry. Note

that, among them, we find all of the insular regions of Spain. In the remaining provinces, the concentration has done no more than increase at a sustained rate over the course of the century, with maximum values on the Theil index in 2001.

What this table demonstrate is that in spite of this general trend towards concentration (and growth), the experience of the regions in Spain has been of a diverse nature (Collantes and Pinilla 2003). This is where the *additive decomposition property* of T^* can play an interesting role.

Let us assume that the municipalities of the country taken as a whole are made up of H different, exhaustive and mutually excluding groups represented by the index $h = 1, 2, 3, \ldots, H$. In our case, the h groups are the 52 Spanish provinces. Let n_h the number of municipalities of the h group and $\mathbf{d}^h = (d_1^h, d_2^h, \ldots, d_{n_h}^h)$ its population density vector, so that d_i^h is the population density of the municipality *i* of group *h*. Let us suppose that $\boldsymbol{\mu} = (\mu_1, \mu_2, \ldots, \mu_H)$ is the vector of group means, with μ_h being the average density of group *h*, $\mu_h = \Sigma_{i=1}^{n_h} s_i^h d_i^h$ where s_i^h is the surface share of municipality *i* with respect to group *h*. Using this notation we can write $\mu = \Sigma_{h=1}^{H} s_h \mu_h$, where s_h is the surface share of group *h* in the aggregate.

We can now express the global dispersion, as measured by T^*, as the sum of two components: (*i*) the dispersion that exists *within* each of the groups, *intra*-group, or *intra*-regional dispersion, and (*ii*) the dispersion that exists *between* the different groups, *inter*-group, or *inter*-regional dispersion.

Moreover, the dispersion within the groups is obtained as a weighted sum of the dispersion indexes applied to each of the groups, where the weights sum up to one and reflect the relative importance (in terms of surface) of each group. As far as the dispersion *between* groups is concerned, this is simply the application of T^* to the average density of each group. To be precise,

$$T^* = \underbrace{\sum_{h=1}^{H} s_h T^{*h}}_{\text{Within Component}} + \underbrace{\sum_{h=1}^{H} s_h \log \frac{\mu}{\mu_h}}_{\text{Between Component}} \qquad (2)$$

We can think of decomposition (2) in the following manner. Let us suppose that within each group (province) we redistribute the population in such a way that all the municipalities have the same density. This means that the dispersion observed would be given by the *inter*-group component in (2), given that *within* each group the municipalities would be identical, $T^{*h} = 0, \forall h$. Let us now suppose that the transfer of population occurs between groups until the average density of each group equalises, and *within* each group the population is reassigned in a proportional way that does not alter the dispersion *within* the group. This means that the dispersion observed would come from the *intra*-group component in (2), given that the average density of the different groups would be identical by construction, $\mu_h = \mu, \forall h$.[1]

This decomposition, applied to the municipalities grouped according to province, is presented in Table 2. On the basis of this table we reach the following conclusions. First, we see the confirmation that the most important component in dispersion is the one that has as its origin the difference in the municipal densities *within* provinces. Almost two-thirds of the dispersion observed according to T^* derives from this component. Consequently, in an analysis of population localization it is necessary to move to a more reduced geographical scale than the regional one. The second conclusion to be drawn is that a considerable portion of the concentration observed (the remaining one-third) derives from the differences in the average density at a provincial level, i.e. differences between regions. Finally, it is confirmed that the dispersion between the average regional population densities has grown substantially over the course of the 20th century, in line with the growth on the global Theil index (T^*). Therefore, the growth of both components (*inter*-group and *intra*-group) is parallel, and such that the relative participation of both is maintained approximately stable over time. In other words, the observed concentration of the population has not only accentuated the differences between the regions (Ayuda, Collantes and Pinilla 2005, 2007), but also the differences within the regions themselves (the *intra*-group or *intra*-regional component).

[1] The argument put forward is parallel to that of the transfer of income in the distribution of income analysis. The "transfers of population" are effectively possible due to the politics of development. During the 1950´s and 1960´s in Spain a large number of "colony towns" sprang up, which meant a real "transfer of population" between regions. An identical phenomenon occurred due to the construction of reservoirs.

Table 2. Theil index decomposition (Mean logarithmic deviation). Regional grouping. 1900 - 2001.

Component	1900	1910	1920	1930	1940	1950	1960	1970	1981	1991	2001
Inter-group/Between	*0.174*	0.176	0.193	0.214	0.236	0.254	0.313	0.446	0.535	0.558	**0.583**
%	*35.9%*	36.0%	37.1%	38.2%	38.1%	38.0%	39.8%	**40.8%**	39.2%	38.0%	37.4%
Intra-group/Within	*0.310*	0.312	0.327	0.346	0.384	0.416	0.474	0.647	0.831	0.912	**0.975**
%	**64.1%**	64.0%	62.9%	61.8%	61.9%	62.0%	60.2%	*59.2%*	60.8%	62.0%	62.6%
Total	*0.484*	**0.488**	**0.520**	**0.560**	**0.619**	**0.670**	**0.787**	**1.093**	**1.366**	**1.469**	**1.559**
%	100.0%	100.0%	100.0%	100.0%	100.0%	100.0%	100.0%	100.0%	100.0%	100.0%	100.0%

Notes: The *minimum* value for the period is shown in *italics*
The **maximum** value for the èriod is shown in **bold**.
Source: INE, IGN and own calculations.

The growth in the *intra*-regional component simply reflects the population concentration in the largest city within each region, generally speaking the capital,[2] and the accompanying evacuation of the smallest settlements (i.e. *the urbanisation effect*, to which it is necessary to add, in the case of the provinces that make up the interior of Spain, the *valley effect*, or a movement away from the mountain areas, and in the case of the provinces at the littoral, the *coastal effect*, or a movement to the coast line). In fact for Spain as a whole, the population density of the capitals was 5.5 times the density of the nation in 1900 and 10.8 times at the beginning of the 21st century.

With very few exceptions (Badajoz, Cáceres and Cuenca),[3] at the beginning of the 20th century all capitals show higher population density than the average of its region, and the tendency is again generally increasing, with only three exceptions. These are Barcelona, Cádiz and Madrid, which show lower population density in 2001 than in 1900. What these results demonstrate, once again, is that in some cases the municipal administrative division is not the appropriate one for the study of population concentration, given that saturation of the physical space is such in some cases that it is necessary to move to larger geographical units, such as metropolitan areas.

[2] There are only seven provinces in which the capital has not been the largest municipality at the time of any of the censuses of the 20th century. Outstanding among them is the case of Pontevedra, whose capital, Pontevedra, has always been second in size to the municipality of Vigo. The other cases are Cádiz, whose biggest municipality since 1950 has been Jerez de la Frontera; Ciudad Real, where between 1900 and 1950 the biggest municipality was Valdepeñas, and between 1950 and 1981 Puertollano; Jaén, whose biggest municipality between 1900 and 1930 was Linares; Asturias, where Gijón was the biggest municipality in several censuses (1910, 1930, 1940, 1950, 1970, 1981, 1991 and 2001); Tarragona, where Reus was the biggest municipality in 1910 and 1920; and finally Toledo, where the capital lost its importance in terms of size to Talavera de la Reina between 1970 and 2001.

[3] It is worth mentioning that the capital of Cáceres has the highest surface of all municipalities in Spain, and the capital of Badajoz is the third in the *ranking*.

IV. POPULATION CONCENTRATION: SOME EXTENSIONS

This section focuses on four additional characteristics of the spatial population distribution, which we think are of interest.

4.1. Average density, median density and asymmetry.

We have already seen how the differences between the regional population densities have grown continuously over time (Table 2), although generally, these densities have increased steadily. Nevertheless, nine provinces registered their lowest density in 2001, namely, Ávila, Cuenca, Huesca, Lugo, Ourense, Palencia, Soria, Teruel and Zamora. This occurred in Segovia, in 1991, and in Guadalajara, in 1981. The rest of the regions registered their minimum values at the beginning of the 20th century.

The situation is very different, however, if we consider the median density as the central measure of the spatial population distribution. That is to say, the population density which divides the distribution into two equal parts: half of the municipalities with a lower density, and half with a higher one. Table 3 shows the average and median density at a national level. The average density grows over time, indicating the increase in the population, but the median density, while showing a slight growth at the beginning of the century, begins a sharp falling off tendency in the second half of the 20th century. The figures could not speak more eloquently: in the present day, half of the municipalities have fewer than 15 Inhabitants/Km^2. This tells us quite clearly about the increasing asymmetry in the spatial population distribution. Moreover, the ratio between the median, $\xi_{0.5}$, and the average, $\frac{\xi_{0.5}}{\mu}$, can be taken as an indicator of concentration, given that it represents the slope of the Lorenz (1905) curve at the percentile 0.5. If the slope falls, as is continuously the case throughout the period under consideration, a smaller percentage of the population lives in the half of the municipalities with lower density, while a larger percentage lives in the half with higher density. In other words, some municipalities are heavily populated, whereas others are becoming empty.

The contrast between the average and median density tells us about the progressive shrinking of the settlements of a more rural nature, typically low density areas (Collantes 2005, 2007). Given that the difference between these two figures can

be taken as an indicator of asymmetry in a distribution, we can see a continuous increase in this asymmetric behaviour.

Table 3. Average density, median density and asymmetry. Spain. 1900 - 2001.

España	1900	1910	1920	1930	1940	1950	1960	1970	1981	1991	2001
Average density	*37*	40	44	48	52	56	61	67	75	77	**81**
Median density	21	23	24	25	26	**26**	25	21	17	16	*15*
Difference (Asymmetry)	*16*	17	20	22	27	30	36	47	58	61	**66**
***Ratio* Median/Average**	**0.575**	0.569	0.551	0.528	0.491	0.466	0.417	0.305	0.228	0.204	*0.180*

Notes: The *minimum* value for the period is shown in *italics*
The **maximum** value for the ériod is shown in **bold**.
Sources: INE, IGN and own calculations. Asymmetry index is the standarised difference between the mean and the median (Hotelling and Salomons 1932).

4.2. Homogeneity *versus* polarization.

An alternative and intuitive way of observing the same result is to look at the percentage of the population on either side of the median of the distribution and observe how this percentage decreases over time. This information is given in the first line of Table 4. The percentage of population that lives in the half of the municipalities with lower density –the Lorenz ordinate at the percentile 0.5, *L* (0.5)– drops from 17.7% in 1900 to a mere 4.0% in 2001. As a consequence, the 20th century has witnessed a continuous transfer of population from one half of the distribution to the other, a process that is especially intense in the period of highest growth, from 1950 to 1981. The last two lines of the table show the percentage of the population within two and four deciles, in a symmetrical fashion, on either side of the median.[4] In both cases, a continuous decrease in percentages is observed, with the maximum values being reached in 1900 and the minimum in 2001. Once again, the figures could not speak more eloquently of the progressive polarization of the spatial distribution of population and the disappearance of the central part of the distribution.

Table 4. Polarization indicators based on the Lorenz ordinates. Spain. 1900 - 2001.

Component	1900	1910	1920	1930	1940	1950	1960	1970	1981	1991	2001
***L* (0.5)**	**17.7%**	17.7%	17.0%	16.1%	14.9%	14.0%	12.0%	8.0%	5.5%	4.7%	*4.0%*
***L* (0.6) - *L* (0.4)**	**11.7%**	11.6%	11.1%	10.6%	10.0%	9.5%	8.4%	6.1%	4.7%	4.2%	*3.7%*
***L* (0.7) - *L* (0.3)**	**24.0%**	23.9%	22.9%	21.8%	20.5%	19.6%	17.3%	12.8%	9.9%	8.8%	*8.0%*

Notes: The *minimum* value for the period is shown in *italics*
The **maximum** value for the ériod is shown in **bold**.
Sources: INE, IGN and own calculations.

[4] This is just the difference of the ordinates of the Lorenz curve between the percentiles 0.6 and 0.4 in the first case, and 0.7 and 0.3 in the second.

4.3. Persistence and the importance of history.

Attention has more than once been drawn to the intense process of population concentration that has occurred in the last century. It is, however, interesting to find out how persistent this has been in terms of the municipalities involved, a question that relates to the *intra*-distributional mobility. A simple coefficient of correlation between the situation in 1900 and that in 2001 shows a moderate degree of persistence in levels, and a higher degree of persistence in *rankings*. In the aggregate this correlation is 0.53 for levels and 0.78 for *rankings* at a municipal scale, even though the period under consideration is over the course of 100 years.[5] Hence, from the aggregate point of view the persistence in the distribution is remarkable.

Looking at the details an important exception in persistence exists at the municipal level. If we look solely at the initial (1900) and final (2001) situation, six municipalities register population losses of more than 10,000 inhabitants: La Unión (Murcia, with 13,983 habitants lost), Valdés (Asturias, with 11,896), Tineo, (Asturias, with 10, 756) Fonsagrada (Lugo, with 10,643), Salas (Asturias, with 10,591) and Cuevas de Almanzora (Almería, with 10,086). The predominance of the mining industry and its geography (Asturias, in the north of Spain) is evident. The mining settlements have been the big losers (in absolute terms) as far as the population is concerned.

4.4. A divergent distribution.

Finally, an alternative way of analysing these results is by means of an equation that relates the initial population density with the rate of subsequent density growth. This is the equation of unconditional β-convergence of the economic growth literature (Barro and Sala-i-Martín 1995). A negative relation between the initial position and the subsequent growth indicates convergence, in the sense that the municipalities with lower density tend to attract more people than the municipalities with more population density, while a positive relation indicates divergence, the municipalities with higher density tend to attract more people, on average, than the ones with lower density. Consequently, we can see a tendency to population concentration in a limited number of settlements –the places, broadly speaking, which already had a larger density initially–.

[5] This correlation is much higher at a provincial scale, as high as 0.987 for 1900 and 2001 for levels.

For the period as a whole, and using logarithms, we obtain,

$$\log(d_{2001}) - \log(d_{1900}) = \hat{\alpha} + 0.4277 \log(d_{1900}) + \hat{u} \qquad n = 8{,}108$$
$$(0.0215) \qquad R^2 = 0.127 \tag{3}$$

where $\log(d_{2001}) - \log(d_{1900})$ represents the density growth over the course of the whole century, the equation is estimated by ordinary least squares, and the robust standard errors are shown in brackets (White 1980). We can see that the coefficient on the initial condition is positive and highly significant (*t*-ratio 19.47), which indicates a tendency towards the population concentration previously referred to.[6]

V. CONCLUSIONS

This short paper offers an insight into changes in the spatial distribution of the Spanish population over the course of the 20th century. To this end, we focus on the variable population density at a municipal scale and the data is provided by the population censuses between 1900 and 2001, conveniently homogenized according to the structure of the municipalities of the latest available census.

Contrary to many studies about the settlement of the population, we do not only focus on cities, however these may be defined, but on the whole of the existing settlements. Extremely small size settlements are numerous in Spain. Half of municipalities that currently exist today have less than 1,000 inhabitants. Consequently, they make up an important part of the geography of the rural environment.

Several basic characteristics of the population concentration in Spain (which in 1900 was already fairly concentrated in certain places) have been ilustrated. This concentration has only increased over the 20th century, particularly during the period from 1950 to 1981. Hence, the analysis presented support the thesis that, in terms of population density, inequalities on a municipal scale have been exacerbated over time. This population concentration has produced: (*i*) wide discrepancies between regions

[6] From the point of view of time series, equation (3) represents an unstable AR(1) process. In this case, the usual estimators do not have good properties to carry out statistical inference. Nevertheless, the equation (3) rests only on the cross-section dimension of our data and is perfectly valid for the statistical inference presented in the text. Ayuda, Collantes and Pinilla (2007) present similar results at the regional level, but over a somewhat longer time period.

which have increased over time, also, (*ii*) marked differences within regions, where the situations are very heterogeneous, (*iii*) a higher degree of polarization between the municipalities of the different regions, (*iv*) a quite strong persistence of the original positions (territorial inertia), (*v*) symptoms of divergence, given that municipalities with higher population density tend to attract more population than municipalities with lower population density, and (*vi*) an increasing role of regional capitals in the concentration process, extending the influences on their respective hinterlands. The population tends to locate today in the same places as in the past. What has changed in a fundamental way is the intensity of the concentration process.

To sum up, the economic development of the 20th century, especially in its second half, has manifested itself in a high concentration of the population in a reduced number of places, a process that has not finished in the present days.

REFERENCES

ADES, A. AND GLAESER, E. (1995) "Trade and circuses: Explaining Urban Giants". *Quarterly Journal of Economics,* 110, 195-228. Madrid.

AYUDA, M. A.; COLLANTES, F. AND PINILLA, V. (2005) "From locational fundamentals to increasing returns: The spatial concentration of population in Spain, 1787 - 2000". Documento de Trabajo 2005-05. Universidad de Zaragoza.

AYUDA, M. A.; COLLANTES, F. AND PINILLA, V. (2007) "Long-run regional population divergente and modern economic growth in Europe: A case study of Spain". Documento de Trabajo 310-2007, Fundación de las Cajas de Ahorro (FUNCAS).

BARRO, R. J. AND SALA-I-MARTÍN, X. (1995) *Economic Growth*, McGraw Hill, New York.

BLACK, D. Y HENDERSON, V. (2003) "Urban evolution in the USA", *Journal of Economic Geography*, 3, 4, (October), 343-372.

COLLANTES, F. (2005) "Declive demográfico y cambio económico en las áreas de montaña españolas, 1860-2000". *Revista de Historia Económica – Journal of Iberian and Latin American Economic History*, 23, 3, 515-540.

COLLANTES, F. (2007) "La desagrarización de la sociedad rural española, 1950-1991". *Historia Agraria. Revista de Agricultura e Historial Rural*, 42, (Agosto), 251-276.

COLLANTES, F. AND PINILLA, V. (2003) "La evolución a largo plazo de la población española, 1860-2000: Tipología provincial y análisis del caso aragonés". *Políticas Demográficas y de Población*, Gobierno de Aragón, 43-70.

DE COS, O. AND REQUES, P. (2005) "Los cambios en los patrones territoriales de la población española (1900-2001)". *Papeles de Economía española,* 104, 167-192. Madrid.

DE VRIES, J. (1984) *European Urbanization 1500 – 1800*, Methuen and Co. Ltd, London.

EATON, J. AND ECKSTEIN, Z. (1997) "City and growth: Theory and evidence from France and Japan". *Regional Science and Urban Economics*, 27, 443-474.

FUJITA, M.; KRUGMAN, P. R. & VENABLES, A. J. (1999) *The spatial economy: Cities, regions, and International trade*. Cambridge, MA: MIT Press.

GARCÍA FERNÁNDEZ, P. (1985) *La población de los actuales términos municipales 1900-1981. Poblaciones de hecho según los censos.* Madrid. INE.

GOERLICH, F. J.; MAS, M. AZAGRA, J. AND CHORÉN, P. (2006) *La Localización de la Población sobre el Territorio. Un Siglo de Cambios. Un Estudio Basado en Series Homogéneas 1900-2001.* Fundación BBVA. Bilbao.

HOHENBERG, P. M. AND LEES, L. H. (1995) *The Making of Urban Europe, 1000-1994*. Harvard University Press: Cambridge, MA.

INSTITUTO NACIONAL DE ESTADÍSTICA (1952) *Tablas de mortalidad de la población española. Años 1900 a 1940*. Madrid. INE.

LORENZ, M. O. (1905) "Methods of measuring the concentration of wealth". *Publications of the American Statistical Association* 9, 209-219.

MARTÍ-HENNEBERG, J. (2005) "Empirical evidence of regional population concentration in Europe, 1870-2000", *Population, Space and Place*, 11, 269-281.

PRADOS DE LA ESCOSURA, L. (2003) *El Progreso Económico de España (1850-2000).* Fundación BBVA. Bilbao.

PONS, J.; PALUZIE, E.; SILVESTRE, J. & TIRADO, D. A. (2007) "Testing the new economic geography: Migrations and industrial agglomerations in Spain". *Journal of Regional Science*, 47, 2, 289-313.

REHER, D. S. (1990) "Urbanization and demographic behaviour in Spain, 1860-1930". En Ad van der Woude, Jan de Vries y Akira Hayami (Eds.), 282-299.

THEIL, H. (1967) *Economics and Information Theory*. Amsterdam: North-Holland.

VAN DER WOUDE, A.; DE VRIES, J. AND HAYAMI, A. (1990, Eds.) *Urbanization in History. A Process of Dynamic Interactions*. Claredon Press. Oxford.

WHITE, H. A. (1980) "A heteroskedasticity-consistent covariance matrix and a direct test for heteroskedasticity", *Econometrica*, 48, 4, (May), 721-746.

ZOIDO, F. AND ARROYO, A. (2004) "La población de España", en A. Arroyo (Coord.) *Tendencias demográficas durante el siglo XX en España*, disponible en la web del INE, http://www.ine.es/prodyser/pubweb/tend_demo_s20/tend_demo_s20.htm, Madrid, Instituto Nacional de Estadística, pág. 22.