

**Documento de Trabajo 2002-01**  
**Facultad de Ciencias Económicas y Empresariales**  
**Universidad de Zaragoza**

**Evolution of the Spanish Urban Structure**  
**during the Twentieth Century**

Luis Lanaspá, Fernando Pueyo and Fernando Sanz

*Department of Economic Analysis, University of Zaragoza*

**Abstract.** In this paper we analyse the evolution of the Spanish urban structure during the period 1900-1999. The methodology employed allows us to obtain two main results. First, that this evolution has not been homogenous, with a divergent pattern of growth being identified for the period running from 1900 to 1970, and a convergent one for the period from 1970 to 1999. Secondly, that the intradistribution movements in the Spanish urban hierarchy were very significant during the course of the 20<sup>th</sup> century. Apart from these two results, we have also detected the existence of spatial correlation for those cities that have either lost or acquired population during the period in question.

**Keywords:** Spanish urban structure, Zipf's law, Markov chains.

**JEL classification:** R11, R12

**Acknowledgements:** The authors would like to express their thanks to Luis Germán and Jesús Mur, as well as to two anonymous referees, for their helpful comments and suggestions on the contents of this paper.

Address: Fernando Sanz. Departamento de Análisis Económico. Facultad de CC. Económicas. Gran Vía 2. 50005 Zaragoza. Spain. Fax: 34 976 761996 E-mail: [fsanz@posta.unizar.es](mailto:fsanz@posta.unizar.es)

## **1. Introduction**

One of the most salient features of our present-day world is its essentially urban nature. Whereas a hundred years ago living in a city was the exception, today this has become the norm. Indeed, approximately half of the world's population is now urban, a percentage that rises to 75% in the case of the developed countries. These statistics would appear to justify the increasing interest that has been shown on the part of economists and geographers in studying a whole range of questions related to the phenomenon of cities. As a result, the discipline of Urban Economics has become an area of expansion whose subject of study involves an increasing volume of population. For an excellent panoramic of the current state of research in Urban Economics, see Huriot and Thisse (2000).

An essential characteristic of the city environment is that it is susceptible to change. Thus, the urban structure and hierarchy that existed only a few decades ago might bear little relationship to the current situation, due to the significant changes that have taken place to it. If this premise is indeed the case, then any study on the urban structure of a given country or geographical area should try to consider the longest time horizon possible, provided that the available data so permits. Only by adopting such an approach is it possible to offer a complete vision of what has really occurred.

Similarly, an urbanisation process has taken place throughout the length of the 20<sup>th</sup> century. This phenomenon, at least as far as developed countries is concerned, has been accompanied by significant processes of industrialisation and economic growth which, although showing an irregular pattern over the course of the years, nevertheless demonstrate a clearly increasing tendency. In other words, the present-day urban structure of countries cannot be considered solely as the consequence of its own evolution, but rather as one that has had important connections with other parallel phenomena of an economic character. Against this background, the aim of this paper is to study the evolution of the Spanish urban hierarchy throughout the length of the 20<sup>th</sup> century. We have earlier defended the need to consider sufficiently long periods of analysis, and thus in this paper we reflect the maximum period for which data is available, that is to say, covering the years 1900 to 1999.

A proposal of this type is not new to the literature, and other countries have been the subject of analysis. Thus, Eaton and Eckstein (1997) have considered the urban evolution of Japan and France, with Dobkins and Ioannides (2000) doing the same for the case of the USA<sup>1</sup>. Whilst both studies represent relevant antecedents for what we are setting out to do, to the best of our knowledge no work has yet considered the Spanish case using the methodology employed in this paper<sup>2</sup>.

The main traits of the change experienced by the Spanish population over the course of the 20<sup>th</sup> century are as follows. First, in 1900 the population of Spain totalled some 17,803,000 inhabitants, whilst by 1999 this figure had reached 40,202,000, that is to say, it had multiplied by a factor of 2.26. Secondly, the largest population nucleus at both dates was represented by the capital city of Madrid, which had 540,000 inhabitants in 1900 and 2,829,000 in 1999, i.e. its population had multiplied by a factor of 5.33. Thirdly, whereas in 1900 some 33.39% of the population resided in metropolitan centres with more than 10,000 inhabitants, this percentage has risen to 75.67% in 1999. However, these macro-statistics hide certain characteristics that are particular to the Spanish case, and thus the rest of the paper will be devoted to a more detailed examination of the changes experienced by the Spanish urban population over the course of the last century.

To that end, Section 2 shows that when analysing a sample of the one hundred largest cities in Spain in each of the years considered in the study, the distribution of their size is well described by a Pareto distribution, as is usual in the literature. The estimation for each year of a linear version of this distribution allows us to draw a first conclusion, namely that urban growth in Spain was divergent from 1900 to 1970, but convergent from 1970 to 1999. That is to say, the differences between the size of the cities increased in the first stage, whilst they decreased in the second. Additional empirical evidence would appear to corroborate this important result.

However, this first approach corresponds to a static point of view that does not permit any analysis of the movements of the cities in the sample from one period to another. This question is solved in Section 3, where we assume that the intradistribution

---

<sup>1</sup> Overman and Ioannides (2001) carry out an alternative analysis of the evolution of the US urban structure by estimating stochastic kernels.

<sup>2</sup> Lasuén et al. (1967), Rotllant and Soy (1993), Brañas and Alcalá (2001) and Brañas et al. (2001) constitute complementary approaches to the characteristics of the Spanish urban system.

movements in the size distribution of the cities from one period to another can be modelled by way of a Markov chain. This dynamic approach allows us to obtain the estimation of the corresponding transition matrix, as well as the vector of ergodic probabilities which defines the long-run distribution of cities. The main conclusion we draw from this analysis is that, perhaps a little surprisingly, the movements within the Spanish urban hierarchy have been significant throughout the period in question, with some cities even experiencing absolute falls in their population and others enjoying intense growth.

The contents of Section 4 follows as a natural continuation of the previous result. Once having detected the existence of an important number of cities whose relative position changes either up or down, and with these being described as winners or losers, the immediate question is whether or not these are grouped in the geographic space represented by Spanish territory. The corresponding spatial dependency tests offer a clearly affirmative answer to this question. Section 5 closes the paper with a summary of the main conclusions.

## 2. The size distribution of cities

Only a limited number of statistical distributions have been used in the literature to describe the urban structure of countries. Essentially, there have been just three: the lognormal distribution, the Pareto distribution and a particular case of the latter, the rank-size distribution, known popularly as Zipf's law (Zipf, 1949). The size distribution of cities follows a Pareto distribution (or, which is the same, comes described by a potential law) when it holds that

$$R = aS^{-b} \tag{1}$$

where  $R$  is the rank or number of cities with a population of size  $S$  or larger,  $S$  is the population of the cities and  $a$  and  $b$  are parameters, with the latter being the Pareto exponent, always positive by construction. The size distribution of cities is more or less equal, depending on the value of  $b$ . At the limit, if  $b$  tends to the infinite, then all the cities will be of an equal size. When  $b$  is equal to one, we obtain the well known rank-size rule or Zipf's law, which postulates that if the largest city of a geographic area has,

for example, one million inhabitants, then the second in size would have half a million, the third three hundred and thirty three thousand and so on<sup>3</sup>. Thus, the Pareto exponent can be interpreted as a metropolitanisation index (taken from Suarez-Villa, 1988), in such a way that values which are falling over time indicate relatively more important roles -increasing weight- for the largest cities and, therefore, a greater metropolitan concentration. By contrast, an increasing tendency represents a greater dispersion of the population outside the large metropolitan areas and a more balanced population distribution between urban centres of different sizes.

Let us now consider the results for the Spanish case. Departing from (1), we take logarithms in both sides and estimate the resulting linear expression:

$$\ln R = \ln a - b \ln S \quad (2)$$

The data has been obtained from the Instituto Nacional de Estadística ([www.ine.es](http://www.ine.es)) for the years 1900, 1910, 1920, 1930, 1940, 1950, 1960, 1970, 1981, 1991 and 1999, with the latter being a projection. For each year we have taken the largest 100 Spanish cities at that moment in time. An important question is that related to the appropriate sample size. Its importance comes from the work of Rosen and Resnick (1980), Guérin-Pace (1995) and Urzúa (2000), which show the possible sensitivity of the results to the sample size employed. In our case, and in addition to the analysis of the 100 largest cities, we have tested with the largest 300, 500 and 700, finding that the results on the time evolution of the estimated Pareto exponent are robust to these alternative samples.

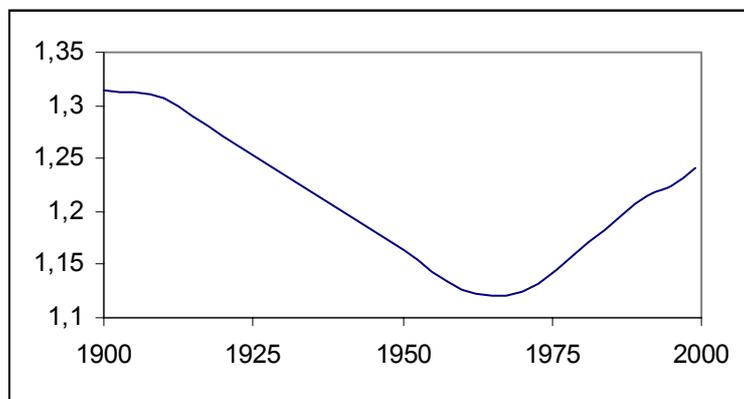
Figure 1 shows the time evolution of the estimations of the Pareto exponent. This parameter is always significant and the degree of adjustment is optimal, with  $R^2$  oscillating between 0.98 and 0.99. From Figure 1 it is clear that the estimation over time of the metropolitanisation parameter displays a U-shape, with the minimum value being reached in 1970. As a result, we can deduce two different patterns of behaviour over the course of the 20<sup>th</sup> century: from 1900 to 1970, the size distribution of the cities in increasingly divergent, whilst from the 1970s to the end of the period this distribution becomes more and more equal. In summary, from this first approach to the data we can draw one of the most fundamental conclusions of the paper, namely the existence of a

---

<sup>3</sup> The Zipf<sup>l</sup> law has been the origin of a wide literature including empirical approaches (Rosen and Resnick, 1980, Guerin and Pace, 1995 and Eaton and Eckstein, 1997) as well as theoretical ones (Simons, 1955, and Gabaix, 1999, among others).

break in the evolution of the Spanish urban structure around 1970s, with this coinciding approximately with the end of the Francoist dictatorship and the beginning of the democratic regime<sup>4</sup>. In the first phase, the growth of the cities is divergent, whereas in the second, it is convergent. Particular emphasis should be placed on this result, given that it is not capable of generalisation to the other developed countries that have been analysed. Thus, Eaton and Eckstein (1997) find that for the case of France from 1876 to 1990 and for that of Japan from 1925 to 1985, urban growth exhibited an essentially parallel character for the two countries. That is to say, all the cities increased their population, independent of their size, at rates that were not significantly different in statistical terms.

**Figure 1.** Evolution of the estimations of the Pareto exponent (N = 100)



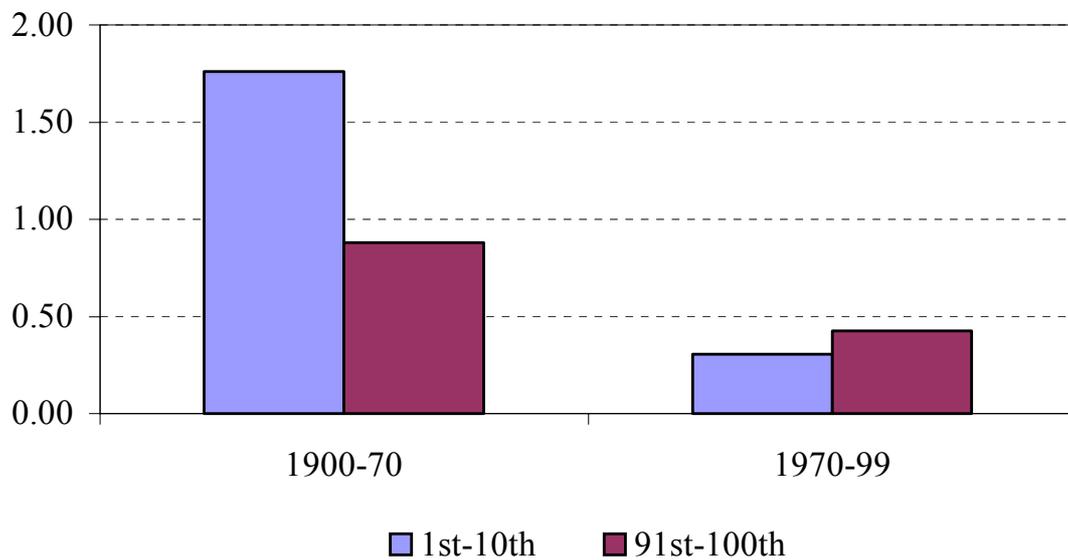
We can offer additional empirical evidence that would appear to corroborate this conclusion. Such evidence is presented following two different approaches. In the first, and for the purpose of illustrating the heterogeneous behaviour of cities with different sizes, we have computed the average inter-annual growth rate of the ten largest Spanish cities in 1900, as well as of the cities occupying the 91<sup>st</sup> to 100<sup>th</sup> position in the same year. We distinguish two sub-periods: 1900-1970 and 1970-1999, and the results are set out in Figure 2.

Note how until 1970 the largest cities grew at rates that were significantly greater than those of the smallest population nuclei, exhibiting what we have called a divergent pattern of growth. However, from 1970 onwards we can appreciate a change in this behaviour, since it is now the smallest cities that were growing at the fastest rate, thus

<sup>4</sup> Obviously, the absence of annual data prevents us from determining the precise date of such a change in the pattern of behaviour.

generating a convergent pattern of growth that erodes the differences in size between cities. A further conclusion that can be deduced from Figure 2 is that, in general, and whatever the size of the cities, the end-of-century growth rates are significantly lower than those found at the beginning of the century. This indicates that the urbanisation process in Spain has consolidated, with the ambit of potential growth of the main population nuclei being limited at the end of the century<sup>5</sup>.

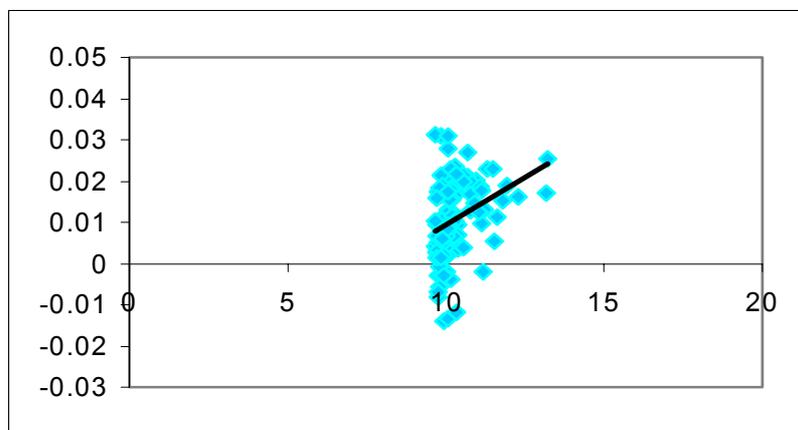
**Figure 2.** Average inter-annual growth rate (%) of the top ten and the bottom ten cities among the 100 largest in 1900



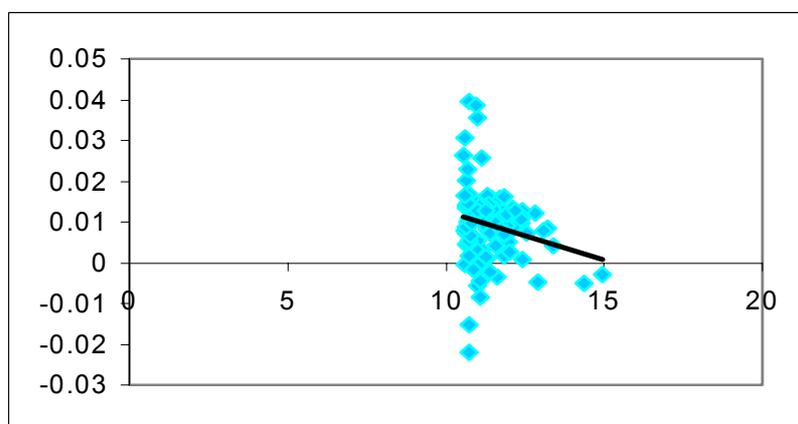
The second approach, which corroborates that growth was divergent until 1970 and convergent since that date, is related with what the empirical literature on economic growth describes as beta-convergence. Applying this concept to the case of cities, Figure 3 shows the dispersion diagram and the corresponding regression line of the inter-annual average growth rate from 1900 to 1970 of the one hundred largest Spanish cities in 1900 on the Neperian logarithm of the population at the beginning of the period under consideration. Figure 4 does the same for the period 1970 to 1999. In both cases, the corresponding estimated parameter is statistically significant.

<sup>5</sup> A detailed analysis of the data shows that for the largest cities, i.e. for the 10 or 15 largest, the rates of growth have even been negative starting in the 1990s.

**Figure 3.** Beta convergence of population growth. 100 largest cities in 1900. Period 1900-1970



**Figure 4.** Beta convergence of population growth. 100 largest cities in 1970. Period 1970-1999



This result is the expected one. Up to 1970 there is a positive correlation between the initial size of the cities and their rates of growth, whereas as from 1970 onwards the correlation between both variables is negative, as corresponds to convergent growth.

### 3. Movements within the distribution

In the previous Section we have analysed the size distribution of cities from an essentially static point of view, that is to say, without paying attention to the movements that take place within the distribution. This is a deficiency that must be corrected, given

that it is difficult to accept that the urban structure has not undergone profound changes over the course of a century. In other words, it could be the case that the distribution as such could remain essentially unchanged for long periods of time, but that significant movements could have taken place in the relative position of the cities within that distribution. In any event, this is a question that merits our attention and to which this Section will be dedicated.

To that end, we have used a dynamic methodology that is able to explicitly capture the movements of cities between periods. This allows us to predict the future size distribution of the cities and to describe the behaviour of the whole of the sample, rather than of just an average element, as in the case in the classic regression analysis. Specifically, we have used the approach proposed by Quah (1993) as an alternative to the so-called sigma-and-beta-convergence analysis used in empirical studies on the convergence in growth rates between countries or regions. This is an approach that has already been used, to the best of our knowledge, by Eaton and Eckstein (1997) for the case of French and Japanese cities and by Dobkins and Ioannides (2000) for the case of the USA.

### *Methodology*

The approach adopted by Quah rests on the assumption that the intradistribution dynamics of a statistical distribution from one period to the next can be modelled by means of a Markov chain. A Markov chain is a stochastic process, that is to say, a family of random variables  $\{Z_t\}$ . The values that  $\{Z_t\}$  can take are their states and the changes in these values are called transitions between these states. Subscript  $t$  denotes a point in the parametrical space  $T$  and  $Z_t, \forall t \in T$ , is a point in the space of states  $E$ . A Markov chain, which we will denote by  $\{X_t\}$ , is a stochastic process with a parametrical space  $T$  of discrete time and a finite space of states<sup>6</sup>  $E$  that verifies the following so-called Markov property:

$$P[X_t = j / X_{t-1} = i, X_{t-2} = c, \dots, X_1 = a] = P[X_t = j / X_{t-1} = i] = p_{ij}(t) \quad (3)$$

---

<sup>6</sup> In fact  $E$  can be finite or infinite numerable. As in the empirical application the number of possible states will be finite, we assume without loss of generality that there are  $M$  possible states:  $E = \{1, \dots, M\}$ .

where  $j, i, c, a \in E$ . This conditions implies that the future of the process depends only on its present state and not on its history.

The transition probabilities  $p_{ij}(t)$  capture the probability that the state  $i$  in  $t-1$  is followed by state  $j$  in  $t$ . We assume that the Markov chain is stationary, that is to say, that it verifies:

$$p_{ij}(t)=p_{ij}$$

We can group all the possible transition probabilities in a stochastic matrix  $P$ , called the transition matrix:

$$P = (p_{ij}) = \begin{pmatrix} p_{11} & p_{12} & \dots & p_{1M} \\ p_{21} & p_{22} & \dots & p_{2M} \\ \dots & \dots & \dots & \dots \\ p_{M1} & p_{M2} & \dots & p_{MM} \end{pmatrix} \quad (4)$$

with  $p_{ij} \geq 0$ ,  $i, j \in E$  and  $\sum_{j \in S} p_{ij} = 1$ ,  $\forall i \in E$ . This matrix  $P$  completely characterises the intradistribution dynamics, thereby defining the law of movement from one period to the next.

In our empirical application we do not obtain the probabilities  $p_{ij}$ , which have a parametric sense, but rather the estimations  $\hat{p}_{ij}$  of these probabilities. If the matrix  $P$  is invariant over time its elements can be estimated from the observed frequencies in the changes of state from one period to another. Thus, following Amemiya (1985) or Hamilton (1994), although using the notation of the former, the maximum likelihood estimator of  $p_{ij}$  is:

$$\hat{p}_{ij} = \frac{n_{ij}}{\sum_{j \in S} n_{ij}} \quad (5)$$

where  $n_{ij}$  is the observed number of cities that, having been initially in state  $i$ , appear in state  $j$  in the immediately following period.

A knowledge of the matrix of movements from one period to the next,  $P$ , allows us to try to determine the direction that the system will take in the long-run. That is to say, what is the probability of every state once the changes represented by matrix  $P$  are repeated an arbitrarily large number of times:

$$\lim_{Q \rightarrow \infty} P^Q \quad (6)$$

It can be demonstrated (see Hamilton, 1994) that, under certain conditions, the limit given in (6) exists and is equal to  $\pi \mathbf{1}$ , where  $\pi$  is a  $M \times 1$  vector and  $\mathbf{1}$  is a  $1 \times M$  row vector of ones. The vector  $\pi$  which, if it exists, is unique, is given the name of equilibrium distribution, long-run or steady state distribution or, more usually, the vector of ergodic probabilities, with the latter being the terminology we have adopted here. This vector  $\pi$  offers very interesting information, given that it describes the future distribution of the cities if the movements observed in the sample period are repeated to infinity.

#### *Results for the Spanish case*

Before presenting the transition matrix for the Spanish case, there are two issues that should be resolved. The first of these is quite simple. We have just described a first order transition matrix, that is to say, a matrix that describes movements from one period to the next. Nevertheless, our data corresponds to only one year for each decade and, as a result, the matrix we present reflects transition between consecutive periods, which are separated by ten years.

The second aspect is more important. The use of Markov chains requires making the space of states discrete, in order to manage a finite number of categories. Furthermore, the number of states has to be chosen carefully, given that the results might not be absolutely neutral to such a choice. In this sense, following Eaton and Eckstein (1997), we define six states which are related with the average size of the sample for each period<sup>7</sup>. These states are as follows: cities whose population is below 30% of the average, between 30% and 50% of the average, between 50% and 75% of the average, between 75% and 100%, between the average and twice the average and,

---

<sup>7</sup> We have also tried other alternatives to divide the range of sizes, which give rise to results that do not qualitatively differ from those presented in the text.

finally, more than twice the average. As in the previous Section, this average has been obtained from the 100 largest Spanish cities in each period.

Nevertheless, our approach differs from that of Eaton and Eckstein (1997) in an essential point, related to the different urban behaviour exhibited by France and Japan, on the one hand, and by Spain, on the other. Eaton and Eckstein (1997) argue that both the French and Japanese urban structure has remained stable over the course of the 20<sup>th</sup> century, with a parallel growth on the part of cities. As a consequence, the sample of cities they consider (39 for France and 40 for Japan), always contain the same cities over time. However, this result does not hold for the Spanish case. As we will show later and in more detail, of the 100 largest Spanish cities in 1900, only 58 remain in the sample of the 100 largest corresponding to 1999. In other words, the Spanish urban hierarchy has experienced significant changes over the century, even to the extent of some cities losing population over the years and others growing at very high rates.

In order to explain the changes undergone by the Spanish urban structure, we must pay attention to the fact that a significant number of cities either enter or leave the sample of the largest 100 at some moments in time. This is a possibility that need not be considered by Eaton and Eckstein. We have taken this feature into account by adding a seventh state to the transition matrix, with this state including, for each period, the 100 cities ranked from position 101 to 200. The transitions from the other states to this seventh correspond to the cities that leave the sample of the 100 largest, whilst the transitions from the seventh to the others reflect the cities that enter it.

With these aspects now being reflected, Table 1 shows the first-order transition matrix, with jumps of ten years, for the period 1900-1999. The seventh state we have just referred to is described as “Rest”, and the last row offers in bold type the ergodic probabilities corresponding to each state.

We can draw a number of interesting conclusions from the contents of this matrix. First, the persistence, which is given by the values in the diagonal, is greater in the extreme states. In other words, the probability of moving out of the initial state is lower for the largest and the smallest cities than it is for those of a medium size. Secondly, the possibility of parallel growth has to be rejected, since several values in the diagonal are

significantly different from one<sup>8</sup>, which indicates that the size distribution of Spanish cities have undergone significant intradistribution movements during the 20<sup>th</sup> century. In this sense, we should draw particular attention to the fact that cities whose size is less than 30% of the average have only a 40% possibility of remaining in the same state and the same probability of leaving the sample of the 100 largest (entering the “Rest” group). Thirdly, the ergodic probabilities show that, in the long-run distribution, the most probable state is “Rest”, indicating a tendency towards a greater proliferation of cities with small populations. This result is in accord with the fact that the largest Spanish cities were experiencing negative growth rates at the end of the 20<sup>th</sup> century. That is to say, the urbanisation process of the largest agglomerations would appear to have reached its limits, thereby favouring the future appearance of smaller population nuclei enjoying lower levels of congestion.

**Table 1.** Transition matrix for Spanish cities (1900-1999)

|          | Upper endpoints |              |              |              |              |              |              |
|----------|-----------------|--------------|--------------|--------------|--------------|--------------|--------------|
|          | $\infty$        | 2            | 1            | 0.75         | 0.5          | 0.3          | Rest         |
| $\infty$ | 0.960           | 0.040        | 0            | 0            | 0            | 0            | 0            |
| 2        | 0.007           | 0.901        | 0.085        | 0.007        | 0            | 0            | 0            |
| 1        | 0               | 0.146        | 0.744        | 0.110        | 0            | 0            | 0            |
| 0.75     | 0               | 0.010        | 0.077        | 0.734        | 0.179        | 0            | 0            |
| 0.5      | 0               | 0            | 0.003        | 0.085        | 0.722        | 0.094        | 0.096        |
| 0.3      | 0               | 0            | 0            | 0.019        | 0.173        | 0.404        | 0.404        |
| Rest     | 0               | 0            | 0.002        | 0.002        | 0.040        | 0.034        | 0.922        |
|          | <b>0.024</b>    | <b>0.134</b> | <b>0.078</b> | <b>0.094</b> | <b>0.158</b> | <b>0.051</b> | <b>0.461</b> |

#### 4. Cities that enter and leave the sample: a spatial analysis

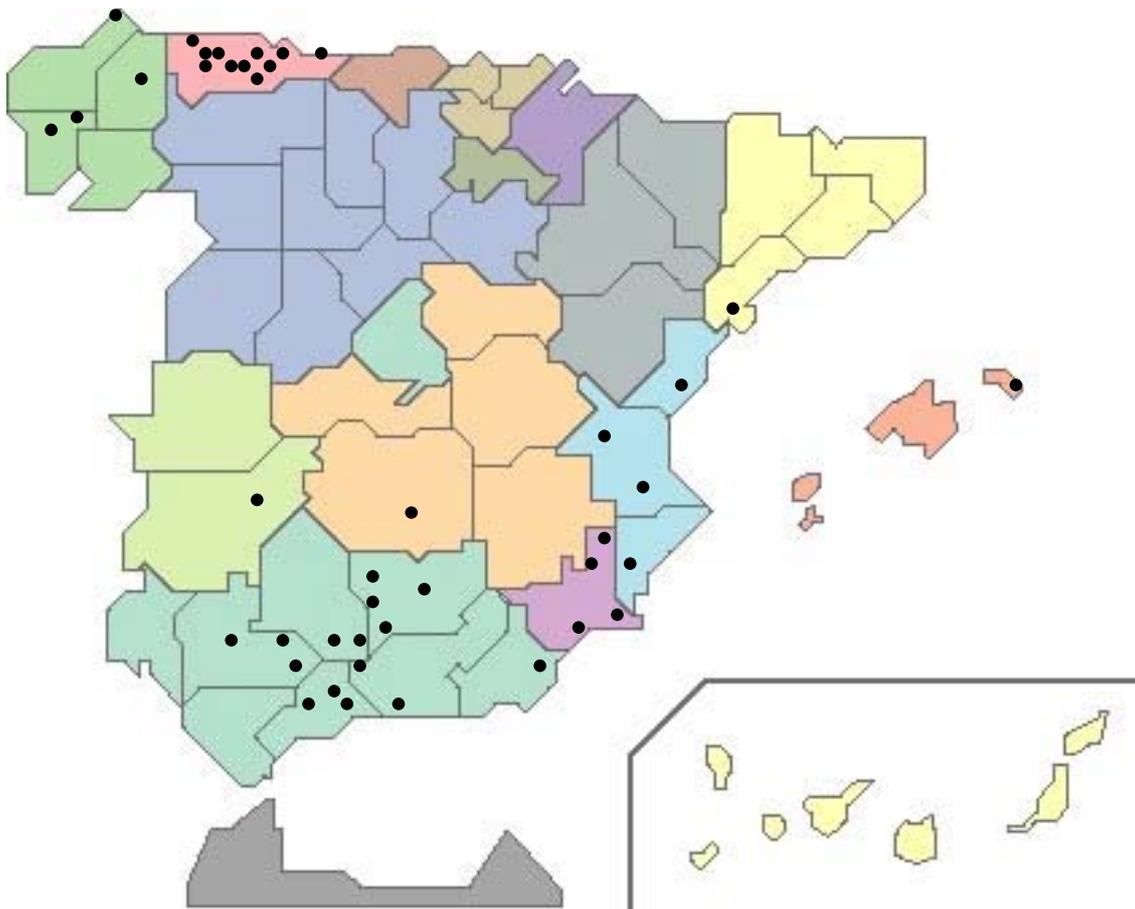
The previous Section has made clear, amongst other results, that the hierarchy of Spanish cities underwent significant changes during the course of the 20<sup>th</sup> century. As

<sup>8</sup> The average value of one element of the diagonal in the Spanish case is 0.770. For Japan (see Eaton and Eckstein, 1997) it is 0.879, that is to say, significantly higher.

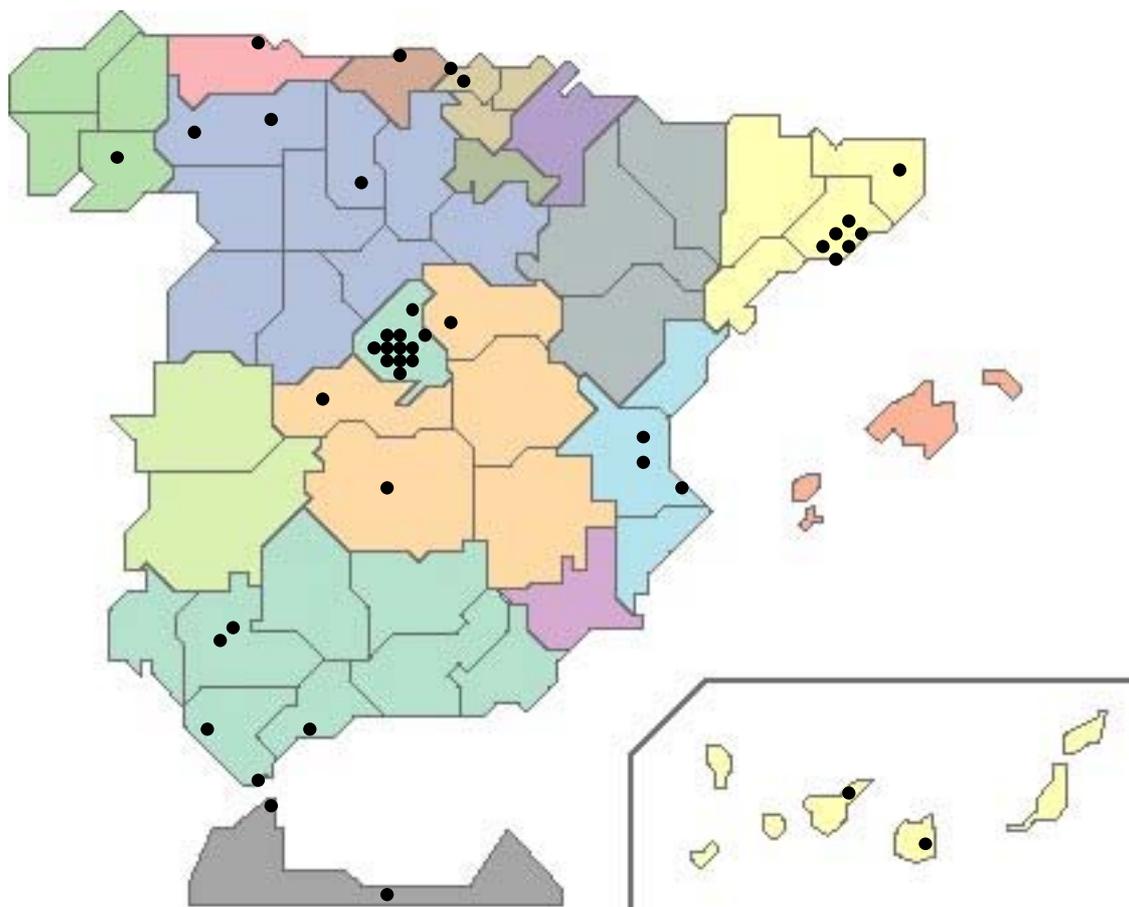
we have noted previously, of the 100 largest cities in 1900, only 58 maintain this condition in 1999. Therefore, 42 of them have left the sample, being replaced by the same number of new emergent cities.

It is worth verifying whether these cities, both those that enter and those that leave, present some type of spatial regularity in the Spanish territory. If this is indeed the case, i.e. if the cities that enter and/or leave the sample are in some way grouped in the space, then the spatial pattern could provide some guide in detecting the possible causes of the changes. Map 1 shows the geographical location of the 42 cities that have left the sample of the largest 100 during the course of the century, whilst Map 2 includes those that have replaced them.

**Map 1.** Cities that have left the sample of the largest 100



**Map 2.** Cities that have entered the sample of the largest 100



These maps confirm our *a priori* suspicion of some spatial regularity in the main urban changes that have taken place. Thus, of the 42 cities that have left the sample, 26 are concentrated in just two of the seventeen Autonomous Regions into which Spain is divided. Of these, 11 are to be found in the reduced geographical space of Asturias (which lies on the Cantabrian coast of Northern Spain) and 15 in Andalusia, a region located in Southern Spain. What explanation can be offered for this result? Following Germán (2001) y Germán et al. (2001), the decline suffered by previously important cities in these regions is directly related with a history of industrial failure. The existence of abundant coal deposits favoured the development of the iron and steel industry in Asturias. However, the sector was unable to successfully weather the serious crisis of the 1970s, thereby consolidating a clear process of industrial decline that has yet to be arrested by substitute strategic alternatives. For its part, the appearance of industry in Andalusia was related to two facts, namely its function as intermediary

between Castille and the Americas and its role as a minerals exporter. Once Spain had lost its last American colonies, following the Spanish-American War of 1898, the mineral extraction industry showed itself to be incapable of constructing and supporting a solid industrial sector. As a result, during the course of the 20<sup>th</sup> century Andalusia lost almost all its industrial activity and became converted into a region specialising in the primary sector.

When we turn our attention to the emergent cities, we find that Map 2 also offers clear guidelines. Here, we can note that the cities entering the sample are concentrated around Barcelona, which lies on the Mediterranean coast of North-east Spain, and, above all, around Madrid, located in the centre of the country, with 13 of the new cities included in the 100 largest being located in the outskirts of Spain's capital city. These are fundamentally commuter cities, lying close to these two large population centres, but enjoying lower levels of congestion. During the 20<sup>th</sup> century Barcelona consolidated itself as Spain's great Mediterranean port, at the same time as exhibiting a marked industrial specialisation. Both Madrid and Barcelona share the character of the industrial and services cities. The former also stands out (again, see Germán, 2001) by virtue of its geographical location in the centre of Spain, its role as the host of the main political and administrative institutions of State, its location as the radial centre of the transport network and its position as the main financial and business centre in Spain. The respective characteristics of Madrid and Barcelona have granted them economic primacy and made them the urban nuclei enjoying the most intense economic activity in Spain, which explains their capacity to generate the new emergent cities located in their outskirts.

So far the analysis we have made has been merely descriptive. Nevertheless, the results derived from Maps 1 and 2 suggests the existence of some degree of spatial dependency or autocorrelation, given the geographic concentration presented by the cities in both maps. The intuitive idea is simple. The null hypothesis to be tested is the presence of spatial independence, that is to say, that the location of the 42 cities that either enter or leave the sample is completely random along the space and does not present any type of regularity. The rejection of this hypothesis implies admitting that the cities which present a certain characteristic, in our case that of leaving or entering the sample, are spatially grouped in certain regions.

Let us now carry out the test in more detail, following Cliff and Ord (1981). Let  $X_i$  be a binary variable that takes the value 1 when the city leaves the sample of the 100 largest and 0 when it remains in it. Further, let  $W = \mathbf{W}_{ij}$  be the generalised matrix of weights, a  $100 \times 100$  matrix, whose order agrees with the number of cities and whose elements take the value 1 when the cities  $i$  and  $j$  belong to the same Autonomous Region, and 0 otherwise. This matrix  $W$  captures the possible spatial correlation and, as is usual in the literature, is normalised so that the elements of each row add up to one. The statistic- $H$  that measures the spatial dependency comes given by the number of times that two cities whose variable  $X_i$  takes the value 1 (that is to say, both either enter or leave the sample) are located in the same region:

$$H = \frac{1}{2} \sum_{i=1}^n \sum_{j=1, i \neq j}^n w_{ij} x_i x_j \quad (7)$$

where  $x_i$  and  $x_j$  are values of the variable  $X_i$  and  $n$  is the number of cities. This joint-count statistic is asymptotically distributed as an  $N(\mu, \sigma^2)$ . Under the strategy of nonfree sampling (the more appropriate in our case), the values of  $\mu$  and  $\sigma^2$  have been taken from expressions (1.31) and (1.32) in Cliff and Ord (1981). The test consists in verifying whether the value of  $S$  belongs to the region of rejection or acceptance of the null hypothesis of the absence of spatial correlation in a typified normal distribution.

The results of this test are as follows. The typified value of  $H$  corresponding to the cities that leave (map 1) and enter (map 2) the sample are 3.12 and 2.78, respectively. In both cases, the existence of spatial independence is clearly rejected at the 5% level of significance. We can therefore conclude that the location of the cities that either leave or enter the sample of the 100 largest is not random. By contrast, in the first case these cities are concentrated in the Autonomous Regions of Asturias and Andalusia, and in the second around the large cities of Barcelona and, above all, Madrid.

## 5. Conclusions

The long-run analysis of the size distribution of cities in a given country is an interesting exercise because, besides offering information about what has occurred, it

allow us to glimpse what might be the future behaviour. In this paper we have carried out such an analysis for the case of Spain during the period 1900-1999.

Three main conclusions emerge from our empirical analysis. First, a structural break in the evolution of the Spanish urban structure took place in the 1970s, coinciding with the end of the Francoist dictatorship and the beginning of the democratic regime. The period from 1900 to 1970 corresponds to a first stage which exhibits a divergent pattern of growth and in which the differences between the sizes of the cities become more marked. Conversely, from 1970 to 1999 we find that the urban growth is convergent, with the intermediate and small cities growing at the fastest rates. This result contrasts with that obtained by Eaton and Eckstein (1997) for France and Japan, where these authors find a parallel growth for both countries over the whole century. They argue that this is a feature that could be expected in countries whose surface has not changed over time and which has had stable population during long periods of time. Following this line of argument, and given that Spain shares these characteristics with France and Japan, a result confirming parallel growth should also be expected, with a limited number of changes in the hierarchy of cities. Nevertheless, we have shown that the data clearly rejects this possibility in the case of Spain.

Secondly, the rates of population growth present a significantly decreasing trend from the 1980s or 1990s. This has occurred mainly in the case of the largest cities, some of which have even experienced negative growth, and reflects the tendency on the part of individuals to move away from population centres with high levels of congestion in favour of intermediate sized cities.

Finally, intense intradistribution movements have taken place in the ranking of cities during the course of the 20<sup>th</sup> century. This is again a result that differs from that obtained for other developed countries. For the Spanish case, we have found that of the 100 largest cities in 1900, only 58 remain within that category in 1999. Further, we have detected that the changes are spatially correlated; specifically, the 42 cities that have left the sample are concentrated in a statistically significant form in Asturias and Andalusia, two Autonomous Regions characterised by a history of industrial decline throughout the 20<sup>th</sup> century. By contrast, the 42 new cities that have entered the sample of the 100 largest are spatially grouped, again in a statistically significant manner, around the great

urban centres of Barcelona and, above all, of Madrid, generating a constellation of commuters towns around both these dynamic cities.

## References

- Amemiya, T. (1985). *Advanced Econometrics*. Harvard University Press.
- Brañas, P. y F. Alcalá (2001). Aglomeración urbana española (1900-2000): Estimaciones “rank-size” vs. tests LM. Working Paper 0101/nº 19. Universidad de Jaén.
- Brañas, P., Rodero, J. y F. Alcalá (2001). A definition of urban nucleus based on a rank-size test. Working Paper 0104/nº 22. Universidad de Jaén.
- Cliff, A.D. y J.K. Ord (1981). *Spatial Processes. Models & Applications*. Pion.
- Dobkins, L.H. y Y.M. Ioannides (2000). Dynamic evolution of the size distribution of U.S. cities. En *Economics of Cities* de Huriot, J.M. y J.F. Thisse (Ed.). Cambridge University Press.
- Eaton, J. y Z. Eckstein (1997). Cities and growth: Theory and evidence from France and Japan. *Regional Science and Urban Economics*, 27, 443-474.
- Gabaix, X. (1999). Zipf’s law for cities: an explanation. *Quarterly Journal of Economics*, 114, 739-767.
- Germán, L. (2001). Especialización sectorial y trayectorias económicas de las regiones en España durante el siglo veinte. Mimeo. University of Zaragoza.
- Germán, L. y otros (2001). *Historia económica regional de España. Siglos XIX y XX*. (Ed.). Crítica.
- Guérin-Pace, F. (1995). Rank-Size distribution and the process of urban growth. *Urban Studies*, 32, 551-562.
- Hamilton, J.D. (1994). *Time Series Analysis*. Princeton University Press.
- Huriot, J.M. y J.F. Thisse (2000). *Economics of Cities*. Cambridge University Press.
- Lasuen, J.R., Lorca, A. y J. Oria (1967). Desarrollo económico y distribución de las ciudades por tamaño. *Arquitectura*, 101.

- Overman, H.G. y Y. M. Ioannides (2001). Cross-Sectional evolution of the US city size distribution. *Journal of Urban Economics*, 49, 543-566.
- Quah, D.T. (1993). Empirical cross-section dynamics in economic growth. *European Economic Review*, 37, 426-434.
- Rosen, K.T. y M. Resnick (1980). The size distribution of cities: An examination of the Pareto law and primacy. *Journal of Urban Economics*, 8, 165-186.
- Rotllan, E. y A. Soy (1993). El papel de las ciudades en el desarrollo regional. El caso español. *Papeles de Economía Española*, 55, 60-81.
- Simon, H. (1995). On a class of skew distribution functions. *Biometrika*, 42, 425-440.
- Suárez-Villa, L. (1988). Metropolitan evolution, sectoral economic change, and the city size distribution. *Urban Studies*, 25, 1-20.
- Urzúa, C.M. (2000). A simple and efficient test for Zipf's law. *Economics Letters*, 66, 257-260.
- Zipf, G.K. (1949). *Human behaviour and the principle of least effort*. Harvard University Press.

## **Documentos de Trabajo**

**Facultad de Ciencias Económicas y Empresariales. Universidad de Zaragoza**

**2002-01:** “Evolution of Spanish Urban Structure During the Twentieth Century”. Luis Lanaspá, Fernando Pueyo y Fernando Sanz. Department of Economic Analysis, University of Zaragoza.

**2002-02:** “Una Nueva Perspectiva en la Medición del Capital Humano”. Gregorio Giménez y Blanca Simón. Departamento de Estructura, Historia Económica y Economía Pública, Universidad de Zaragoza.

**2002-03:** “A Practical Evaluation of Employee Productivity Using a Professional Data Base”. Raquel Ortega. Department of Business, University of Zaragoza.

**2002-04:** “La Información Financiera de las Entidades No Lucrativas: Una Perspectiva Internacional”. Isabel Brusca y Caridad Martí. Departamento de Contabilidad y Finanzas, Universidad de Zaragoza.