FEP WORKING PAPERS
FEP WORKING PAPERS

Research Work in Progress

n. 230, Sept. 2006

Long term evolution of the size distribution of Portuguese cities

Ana Paula Delgado

Isabel Maria Godinho

CEDRES - Centro de Economia, Desenvolvimento Regional, Estudos e Serviços

U.PORTO

FEP FACULDADE DE ECONOMIA UNIVERSIDADE DO PORTO

Long term evolution of the size distribution of Portuguese cities[♣]

Ana Paula Delgado

CEDRES / Faculdade de Economia, Universidade do Porto,

Rua Dr. Roberto Frias, 4200-464 Porto, Portugal.

E-mail: apaula@fep.up.pt

Isabel Maria Godinho

CEDRES. Rua Dr. Roberto Frias, 4200-464 Porto, Portugal

E-mail: isa_godinho@hotmail.com

Abstract

In this paper we study the evolution of the Portuguese urban system from 1864 to 2001. We apply the rank-size model and use rank-size estimates to describe the evolution of city-size hierarchy.

NonParetian behavior of the distribution is examined by adding a quadratic term to the basic equation of the model. Our results enhance two different processes in the evolution of urban system: until the middle of the twentieth century urban growth was accompanied by population concentration in the largest cities; afterwards growth benefits middle size cities, reinforced in the last decades by heavy population losses in the two largest cities.

From the association between the characteristics and evolving pattern of city size distribution and the spatial pattern of urban growth, it appears that the nonParetian behavior of city size distribution in the last decades can be linked to the particular growth process of cities located in the proximity of the central cities of the two metropolitan areas of mainland Portugal.

In order to obtain a better understanding of the dynamics of the Portuguese urban system we examine the movements in the ranking of cities, through a Markov chain process. We also analyse the existence of spatial correlation in the process of urban hierarchy restructuring.

Keywords: Urban hierarchy, rank-size distribution, urban growth, Markov processes

JEL classification: O18, R11, R12

[♣] We thank our colleague Rui Alves for helping us compute the ergodic probabilities.

1. Introduction

The aim of this paper is to analyse the long term evolution of the Portuguese urban system, from 1864 to 2001. Studies in this vein have been conducted, for example, by Guérin-Pace (1995) for France, Eaton and Eckstein (1997) for France and Japan, Dobkins and Ioannides (2000), Black and Henderson (2003) for the USA and by Lanaspa et al. (2003) for Spain. All of them revisit the rank-size model, which has been recognised as one of those stylised facts in spatial economics, and there is a general acceptance of that model as a good synthetic description of the hierarchical organisation of urban systems.

In a previous paper we have provided empirical evidence of the evolution of the rank size exponent and examined the effect of varying city size cut-offs on its estimated value. We studied further the deviations of the rank size distribution from linearity, which is seen as a violation of Gibrat's Law, since in order to generate a log-normal distribution, city growth rates must be independent of initial city size and also independence from one period to another. We concluded that, in Portugal, rather than the relationship between size and growth rates, deviation from linearity seemed to arise from autocorrelation in successive growth rates. From our results, we detected a pattern of urban growth characterised by concentration of population in the early phases of the period considered, followed by a decrease in concentration that appeared to result, in the last decades, from a process of selective growth beneficial to the same cities, in particular those that are closer to the central cities of the metropolitan areas of Lisboa and Porto. In this paper we develop that study and take a Markov chain process to describe mobility of cities within city size distribution, examining also the existence of spatial autocorrelation in the movements of cities.

In section 2 we present and compare the datasets that we employ and discuss some of the drawbacks arising from the concepts of urban unit that are used. In the following section we apply the rank size model to the analysis of evolution of the Portuguese urban system. We start with a brief characterisation of that system, enhancing its specific traits. We then use a Pareto distribution to estimate, in each census date, the size distribution of cities. We discuss slope sensitivity to sample threshold and to urban definition. We extend the rank size model by adding a quadratic term to the basic equation and analyse the long term evolution of the estimated parameter and the sensitivity of this estimates to sample threshold. In section 4, we apply a Markov chain process to describe the inter-census movements of cities within the distribution. We compute the year to year transition matrix, from which we calculate the average transition matrix and the associated ergodic probability vector. In section 5, we study the spatial pattern of movements within the distribution between 1864 and 1991. Since, during the more than a century-long period of our analysis, an important number of upward or downward movements occur and as there seems to occur a spatial pattern in these movements, we test for spatial autocorrelation. Finally, in the last section we present the main conclusions.

2. Description of the data set

Studies on urban hierarchy and rank size distribution are contingent on the definition of the unit of analysis. Thus, the characteristics of the urban system resulting from the analysis of the rank-size parameter estimates depend on the definition of urban units. From a theoretical perspective, the adequate definition would be one that considers the urban place as an integrated economic and functional unit. But as a rule, researchers are constrained by the lack of appropriate data.

Another problem concerns the definition of urban units and its consistency over time. In this paper we use two city-proper databases for mainland Portugal, where cities are defined according to administrative criteria. A drawback with a sample based on administrative definitions is that city boundaries may not coincide with the functional and economic boundaries of urban places. However, applying city definitions to prior decades in a single country study, instead of contemporaneous administrative definitions, minimises the problem of city definition and that of building consistent definitions over time[1].

Portugal is a country with long established national borders whose mainland urban system dates back to a number of centuries: many of the cities are several hundred years old and a number of them are even older than the nation. Over time, some of the older cities may have lost population and various urban functions. Still, they retain their administrative status. On the other hand, on early dates, some cities had zero population or were too small to be considered urban units. So, in order to define whether a place qualifies as a city, we use an absolute cut-off of 2000 inhabitants, on each census date.

The data set for the 1864-1991 period was developed by Albergaria (1999) and uses a consistent definition of cities, calculating the population for each city and each census using the 1998's administrative cities. In order to analyse the recent evolution of the urban system,

[1] We must note that in Portugal, as in many other countries, data constraints do not allow alternative approaches to city definition over time.

we use another city proper database[2], for the 1991-2001 period. This latter database uses the 2001 administrative classification of cities. As a consequence the number of cities for 1991, grows from 111 to 123 and, as we observed inconsistency between the two data sets, we considered them separately.

In short, our sample obeys two criteria: 1) urban places which in the 1998 or 2001 database have the administrative status of "city"; and 2) have at least 2000 inhabitants, on each census date.

For the 1991-2001 period we also analyse the sensitivity of the results to the definition of urban units, using data supplied by INE (National Statistics Institute) and referring to urban places[3], with at least 2 000 inhabitants. Differences between city proper and urban place databases arise mainly from the criteria that a place must observe in order to qualify as a city[4] and are reflected in the size of both data sets. In fact, for the urban places database, the number of urban units rises to 450, in 1991, and 531, in 2001.

3. Rank-size evolution of the Portuguese urban system

3.1. Some basic facts about the Portuguese urban system

The Portuguese urban system is characterised by a large number of very small cities – 50% of the cities had, in 1991, fewer than 14000 inhabitants - and two dominant cities, which are the central cities of Portugal's two metropolitan areas (Table 1). The long term evolution shows a slow increase in the number of cities, between 1864 and 1991, while city population more than quadruplicates in the same period. As a consequence, average city size increased from 8829 inhabitants, in 1864, to 29087, in 1991. In general, urban population grew faster than total population and the urbanisation rate[5], although moderate, increased from around 19%, in 1864, to 34% in 1991.

The growth of urban population is faster than that of the number of cities, suggesting an urbanisation process characterised above all by population concentration in existing cities. This process of concentration favours the two main cities, Lisboa and Porto. After 1940, the

---

[2] This database was built by Ferreira, Cardoso and Silva (2003) based on INE (2002) - Instituto Nacional de Estatística (INE) / National Statistics Institute

[3] Places are defined as continuous built up areas with at least 10 or more dwellings, which have an own assignment, regardless of whether they belong to the same basic administrative unit of the country ("freguesia").

[4] Nowadays in order to qualify as a city, places must have at least 8000 voters and possess a certain minimum set of functions and social infrastructure; the acquisition of that administrative status also depends on political criteria.

[5] Defined as the ratio of total city population (urban population) to total population, in a given year, expressed in percentage.

decline in the primacy index[6] portrays a process of decentralisation of urban growth, reinforced in the last decades by heavy population losses in the central cities of the Lisboa and Porto metropolitan areas. However, in 2001, 57% of the Portuguese urban population lived in the 28 cities that belong to the metropolitan areas of Lisboa and Porto.

Table 1 - Basic data on the Portuguese Urban System (Full sample), 1864-2001

Data Source	Census date	Number of Cities	Average size	Median Size	Minimum Size	Maximum Size	Urban Population	Urbanisation rate (%)	Top Two Primacy Index (%)
Albergaria database	1864	85	8829	4563	2013	190311	750496	18.83	37.26
	1890	91	11791	5469	2172	300964	1072970	23.02	41.70
	1900	97	12397	5815	2044	351210	1202476	24.05	43.07
	1920	101	14688	6851	2054	484664	1483455	26.17	46.31
	1940	105	19502	9277	2075	694389	2047756	28.37	46.54
	1950	108	21571	9755	2009	783226	2329644	29.41	45.70
	1960	109	23278	10206	2092	802230	2537248	30.60	43.58
	1970	108	25057	10520	2141	769044	2706118	33.31	39.73
	1981	110	29637	12457	2189	807937	3260069	34.92	34.82
	1991	110	29087	13248	2789	663394	3199601	34.14	30.19
Atlas database	1991	122	29546	13638	2487	661966	3604563	38.46	26.65
	2001	122	30895	15382	2578	564657	3769214	38.19	21.96

Source: Delgado and Godinho (2004: 9)

The image of the Portuguese urban system portrayed by the urban places data set is quite different (Table 2). The urbanisation rate in 2001 rises from 38% to 55% when we consider urban places instead of legal cities. As expected, the top two primacy index decreases.[7] The number of urban units is substantially higher, with an average size of around 10000 inhabitants, which is about one third of the average size for the Atlas database. These differences can be imputed to the legal requirements that a place must fulfill in order to qualify as a city. However, as Carter (1981: 20) points out: *"In older countries many towns which have long decayed retain their former status and chartered rights and fight energetically to maintain them; likewise newly grown towns find it a lengthy and cumbersome process to obtain the articles of recognition."* As a consequence, although the Atlas database contains urban units that do not conform to the size implicit criteria of 8000 voters, the relative importance of very small towns is lower.

The effect of urban definition in sample size is drastic: when we consider a 5000 inhabitant threshold we lose more than sixty percent of the number of urban units in the 2000 inhabitant

6 Defined as the ratio of resident population in the top two cities to total urban population, expressed in percentage.

7 We must note that in both data sets the top two urban units (Lisboa and Porto) are roughly the same size. In fact, in both cases, urban place and administrative city are synonymous.

sample for the urban places data set, whereas for the Atlas data set, the reduction in the sample size is less than 10% of the initial size. On the other hand, the effectiveness of *bureaucratic* and political barriers to access city status is reflected in the fact that, for the 10000 inhabitant threshold, the number of urban units is significantly higher when we consider the urban place data set.

Table 2 - Basic data on the Portuguese urban places, 1991-2001

Data Source	Census date	Number of Cities	Average size	Median Size	Minimum Size	Maximum Size	Urbanisation rate (%)	Top two primacy index (%)
INE Urban Places	1991	450	10103	3934	2004	662782	48.51	21.23
	2001	531	10270	4323	2001	563818	55.26	15.16

3.2. The rank-size model

According to the rank-size model, the size distribution of cities follows a Pareto distribution:

(1) $R_{it} = AP_{it}^{-\alpha}$ or, in logarithmic form, (1') $\log R_{it} = \log A - \alpha \log P_{it}$

where R_{it} is the rank of the i^{th} city in time period t, P_{it} is the size (population) of the i^{th} city in time period t, A is a constant and α is the Pareto/Zipf's exponent. This formula is known as the Pareto equation[8].

City size distribution is then characterised by the number of cities and two parameters: the exponent (α) and the constant term (A). The exponent is a measure of city size inequality in a given urban system and time period. Using Pareto's formula, when $\alpha > 1$ the rank-size curve is flatter and city sizes are more evenly distributed than that predicted by Zipf's law ($\alpha = 1$). In particular, considering the limiting value of $\alpha \to \infty$, all cities would have the same size. On the other hand, when $0 < \alpha < 1$, the rank-size curve becomes steeper. In this case, urban hierarchy is more contrasted than in Zipf's case and cities at the top of the hierarchy are larger. Here we obtain a more heterogeneous distribution of city sizes. In the limiting case of $\alpha \to 0$, there would be just one city in the urban system.

---

[8] Another formulation is that of Lotka (1924), which is given by the following equation: $P_{it} = BR_{it}^{-\beta}$ or, in logarithmic form, $\log P_{it} = \log B - \beta \log R_{it}$ where B is a constant and β is the inverse of the Pareto exponent. The two formulations can further be related to as $B = A^{\beta}$.

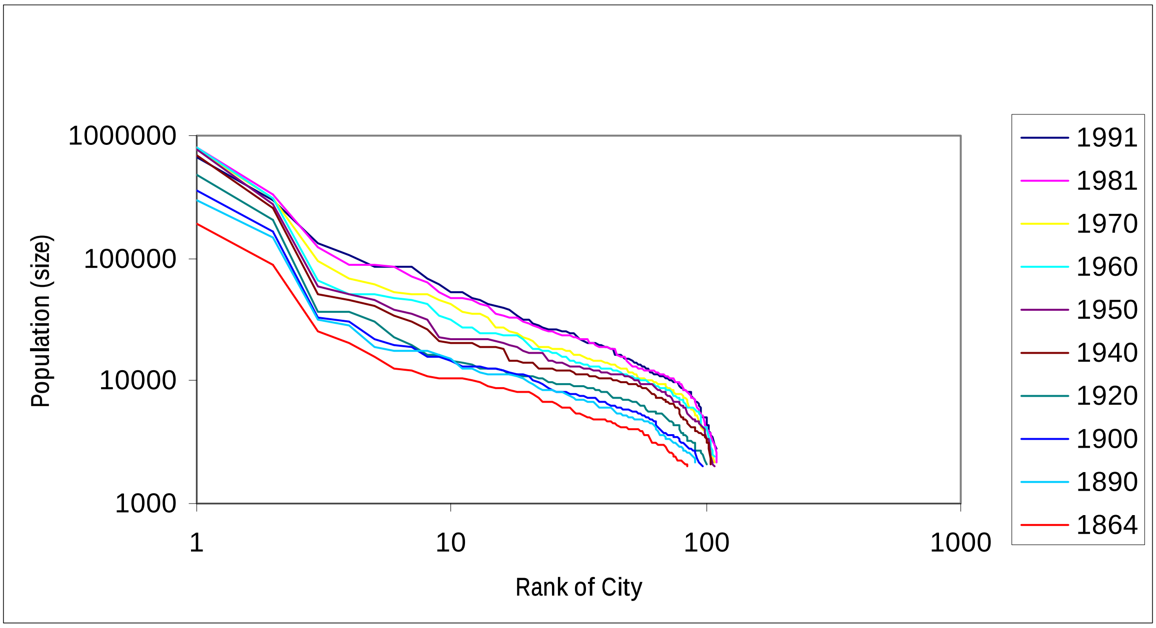
3.3.The long-term evolution of city size distribution: 1864-2001

To study the long term evolution of city size distribution, we began by constructing a rank-size graph, observing how the shape of that distribution evolved over time. Next, we estimated the rank size model by ordinary least squares (OLS) and analysed the long term evolution of slope estimates and the sensitivity of these estimates to sample threshold. Then we studied the deviations from rank-size linearity, following the Rosen and Resnick approach (1980).

The rank-size graph[9]

From Figure 1 we can conclude that, on the whole, the shape of the rank-size distribution remained stable until the eighties, shifting up in the course of time, as a result of urban growth. This does not mean that individual city ranking has remained unchanged; in fact, excluding Lisboa and Porto, cities' relative positions in urban hierarchy have changed. The rank-size graph shows a significant increase in its height and a slight enlargement at the bottom. This result points to an urban growth process characterised by considerable growth in the size of the largest city and a slow increase in the number of cities. Generally, the rank-size line shows an upward concavity between the 3rd and the 20th city, as a consequence of the under-dimension of middle size cities. It also presents a downward concavity in the lower tail of the distribution, translating the excess of small cities.

Figure 1 – Rank-size distributions of Portuguese cities, 1864-1991: Full sample

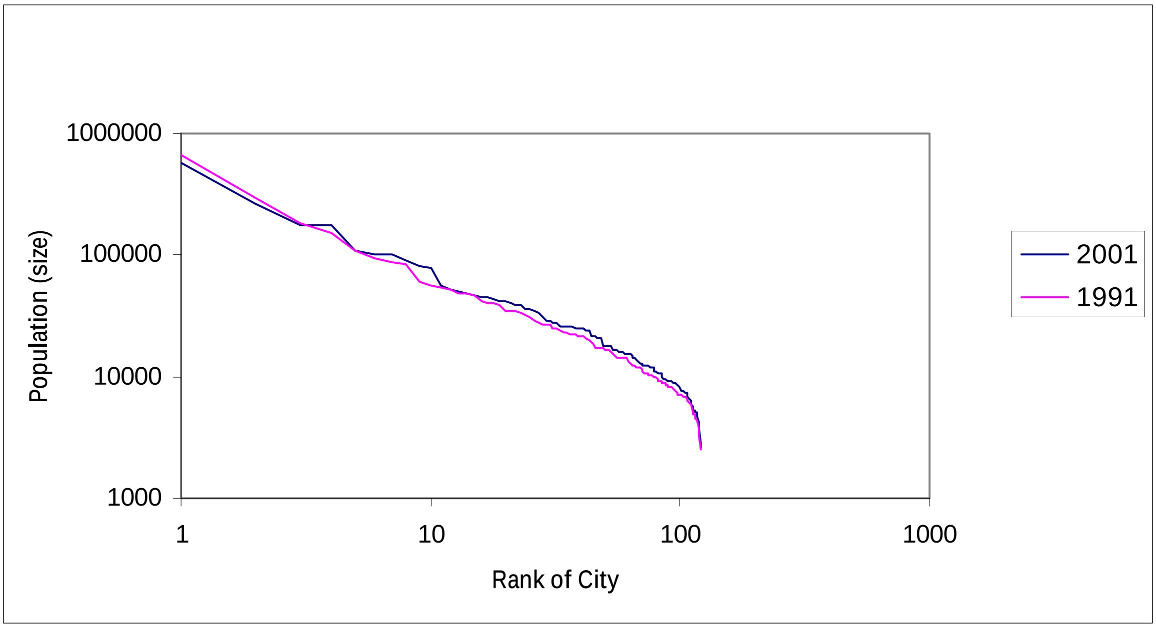


Source: Delgado and Godinho, 2004:11

[9] We must note that the graph refers to Lotka's formulation.

In the last decade, we denote a downward counter clockwise movement of the rank-size line, due to the decline in the size of the two largest cities. There is a more even distribution of city sizes, as the top two cities have lost population, whereas middle size cities have experienced population gains and the dimension of the smallest cities in our sample has remained roughly stable (Figure 2).

Figure 2 - Rank-size distributions of Portuguese cities, 1991-2001: Full sample (Atlas)



Source: Delgado and Godinho, 2004:11

Results from the estimation of the model

The estimation of the rank-size model requires the ordering of cities from the largest down to the smallest. We applied OLS to equation 1'. In order to examine the sensitivity of the slope estimates to the choice of sample threshold we defined several sample cut-offs, chosen taking into account the dimension of the Portuguese city system[10] and current cut-offs for urban definition in the Portuguese statistical system. The estimates of rank-size parameters are all statistically significant at 5% significance level. The quality of the adjustment is quite good, since R^2 are high and close to unity (Table 3).

When the entire distribution is used, from 1864 till 1960, the slope is higher than one and decreasing, indicating that city size distribution is, at the beginning of the period, more evenly distributed than predicted by Zipf's law, becoming increasingly divergent and resulting in a

[10] We did not consider sample thresholds of at least 50 000 inhabitants or higher because the number of cities obeying that criteria is too small.

more contrasted urban hierarchy. From 1970 onwards α is less than one and tends to decrease. However, in the last two decades, we observe a reverse in that tendency, reflecting a process of decreasing inequality.

Table 3 - Results of OLS estimation, 1864-2001

Database	Census Date	Cities with 2000 inhabs. or more			Cities with 5000 inhabs. or more		
		Number of cities	Slope	R^2	Number of cities	Slope	R^2
Albergaria	1864	85	1.189	0.946	35	1.081	0.895
	1890	91	1.120	0.947	51	1.105	0.911
	1900	97	1.098	0.936	58	1.127	0.901
	1920	101	1.082	0.907	69	1.140	0.878
	1940	105	1.061	0.908	80	1.158	0.897
	1950	108	1.022	0.899	88	1.155	0.915
	1960	109	1.026	0.921	95	1.144	0.946
	1970	108	0.963	0.927	92	1.113	0.966
	1981	110	0.937	0.931	97	1.073	0.973
	1991	110	0.953	0.947	100	1.054	0.979
Atlas	1991	122	0.970	0.961	112	1.050	0.988
	2001	122	0.977	0.950	115	1.051	0.977
Database	Census Date	Cities with 10000 inhabs. or more			Cities with 20000 inhabs. or more		
		Number of cities	Slope	R^2	Number of cities	Slope	R^2
Albergaria	1864	12	0.761	0.919	4	0.557	0.949
	1890	19	0.839	0.875	4	0.495	0.917
	1900	21	0.842	0.862	5	0.500	0.932
	1920	23	0.806	0.881	6	0.517	0.928
	1940	45	1.006	0.874	11	0.632	0.924
	1950	53	1.069	0.879	16	0.726	0.896
	1960	57	1.089	0.917	19	0.813	0.882
	1970	59	1.094	0.955	21	0.894	0.931
	1981	75	1.132	0.978	36	1.066	0.959
	1991	73	1.130	0.989	36	1.107	0.979
Atlas	1991	77	1.120	0.993	42	1.138	0.989
	2001	85	1.152	0.991	48	1.221	0.989

These results must be interpreted with caution as Portugal has an urban system with primatial characteristics. For instance, if we take the 1991 city size distribution in the Albergaria database and compare the observed sizes with the expected size of equivalent rank for a top city of 663394 inhabitants and $\alpha = 1$, all the cities from the 2nd to the 25th rank are under-dimensioned. In particular, population deficit is more notorious for cities ranking from the 3rd to 10th position. The opposite situation occurs from the 26th to the 87th position, where cities are bigger than expected. Finally, for all the remaining positions at the bottom of the distribution, cities are smaller than predicted by rank-size rule – some of them with less than 50% of their expected population.

When smaller cities are excluded (sample thresholds of 10000 inhabitants or more), slope estimates tend to increase over time, starting from values less than one, indicating a reduction in city size inequality. This distinct evolution, in comparison with the full sample, mirrors the changes in growth behaviour of middle sized cities vis-à-vis the first city. At the beginning of our study period, intermediate cities, in the class size of 30000-100000 inhabitants, developed more slowly than Lisboa, growing at a faster rate, after the fifties.

Since we are studying the long term evolution of the urban system, an absolute cut-off does not account for the change in typical city size with the urbanisation process. So we consider an upper tail distribution which includes cities in the top one third of size distribution, on each census date, and re-estimate the model (Table 4). Slope estimates exhibit a long term U-shaped pattern, with a minimum value in 1920-1940, indicating an urbanisation process characterised by increasing city size inequality, for the upper tail distribution, until the middle of the last century. Afterwards, the reversing of the tendency points to a diminishing inequality, reinforced in the last decades.

Table 4 - Results of OLS estimation: upper 1/3 of the cities, 1864-2001

Database	Census Date	Top third - upper 1/3 of cities			
		Number of cities	Minimum size	Slope	R^2
Albergaria	1864	28	6046	1.020	0.882
	1890	30	7156	0.969	0.886
	1900	32	7591	0.962	0.877
	1920	34	8798	0.924	0.857
	1940	35	10802	0.927	0.884
	1950	36	12307	0.949	0.882
	1960	36	13091	0.969	0.909
	1970	36	14837	0.998	0.946
	1981	37	19318	1.071	0.960
	1991	37	19990	1.110	0.980
Atlas	1991	41	21416	1.135	0.989
	2001	41	24481	1.209	0.988

The sensitivity of the slope estimates to sample cut-offs is well illustrated in Figure 3, and is higher at the beginning of the observation period. From 1864 to the middle of the 20th century, as sample threshold increases, slope estimates decrease and differences are more important for higher sample cut-offs. The distribution gets more uneven as we impose higher thresholds. For the last decades, slope estimates tend to increase with the sample threshold.

In conclusion, in the first part of the period smaller cities tend to generate a more even distribution, whereas in the last decades the rise in α values with sample threshold seem to indicate that medium and larger cities are the source of a more equal distribution. This tendency is also evident when we confront the full sample with the upper tail distribution (Figure 4).

Figure 3 – Sensitivity of slope estimates to sample threshold

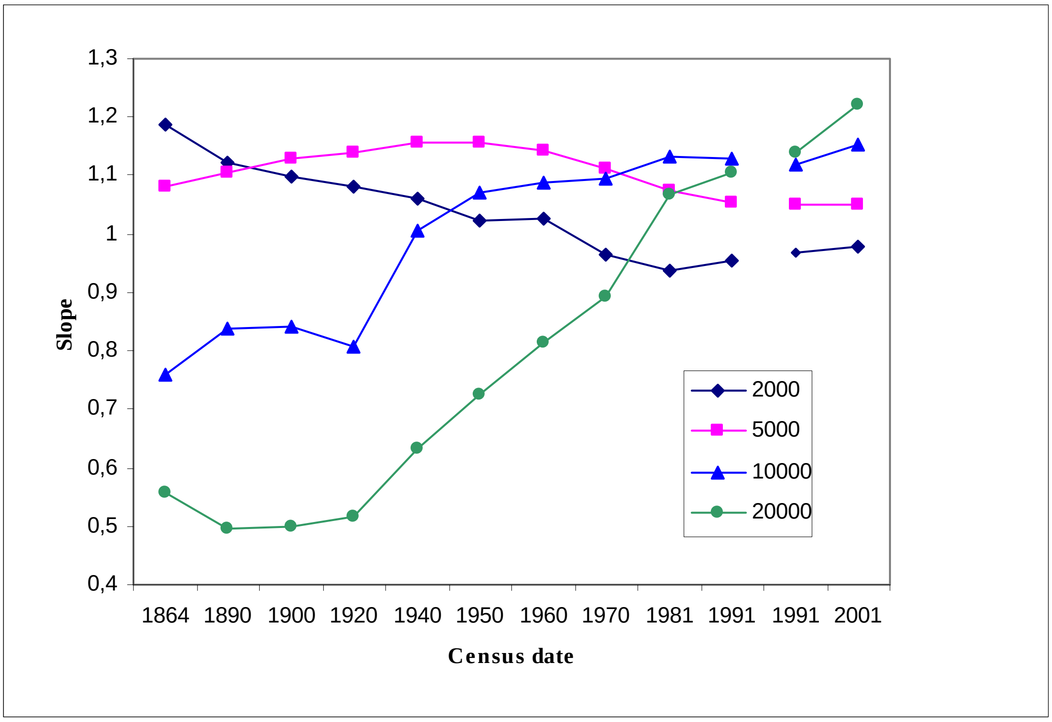
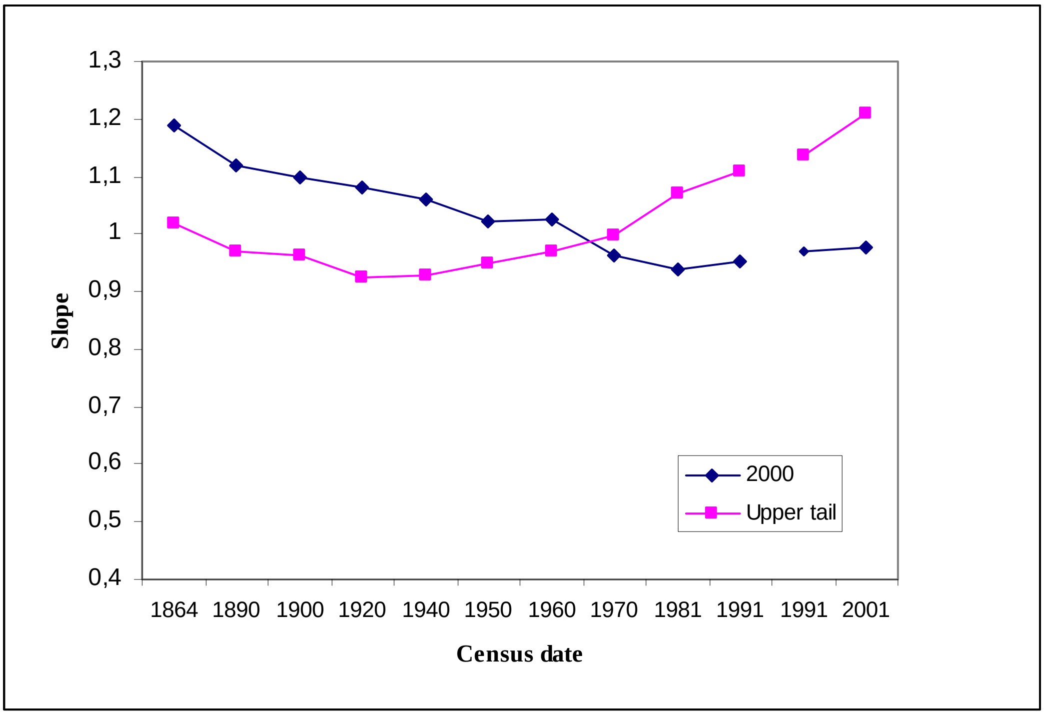


Figure 4- Sensitivity of slope estimates: full sample versus upper 1/3 of cities



Deviations from rank-size regularity

The fact that slope estimates are sensitive to sample size signals a non-Paretian behaviour of the distribution. Therefore, we examine the deviations of the rank-size distribution from linearity by adding a quadratic term to equation 1', following the standard approach in literature. Thus, we estimate the following equation:

$$(2)\quad log\, R_{it} = a + b\, log\, P_{it} + c\left(log\, P_{it}\right)^2 \;.$$

The value of the parameter *c* characterises the curvature: when *c*>0, the rank-size curve is strictly convex (upward concavity) and when *c*<0, it is strictly concave (downward concavity). An upward concavity is obtained when the city size distribution has a smaller number of middle-sized cities than predicted by Zipf's Law. In this case, there is a deficit of intermediate cities in favour of the largest cities' dimension or the number of small cities. A downward concavity means that there is a larger number of middle-sized cities than expected. In this case, there is an excess of intermediate cities relative to the dimension of the largest cities or to the number of small cities. In rank-size distributions with an upward concavity, the largest city will be larger and smaller cities will be more numerous than expected in a linear relationship between the logarithm of city size and the logarithm of its order. On the other hand, in rank-size distributions with a downward concavity, middle-sized cities are larger than expected in a linear relationship between the logarithms of size and order.

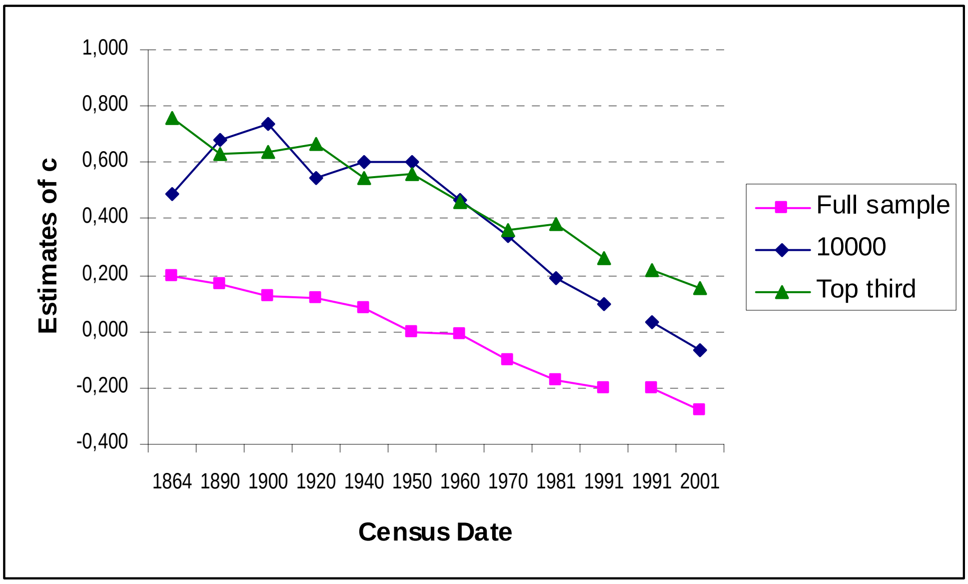
The long term evolution of parameter *c* is depicted in Figure 7.[11] Considering the full sample, until the middle of the 20th century, urban growth favours the largest cities. In 1950 and 1960, the value of *c* is not significantly different from zero meaning that the rank-size distribution tends to conform to linearity. From 1970 onwards, the value of the quadratic parameter is negative reflecting the growth of middle-sized cities, reinforced in the last decades. When we exclude small cities from the sample (10 000 inhabitants and upper tail distributions), the estimates of *c* remain positive for the 1864-1991 period indicating that middle-sized cities are smaller than expected in a linear relationship. Since *c* is decreasing, this characteristic is less accentuated in recent years, signifying that urban growth has been concentrated in cities of that size class.

Our results for the long term evolution of *c* are similar to those of Guérin-Pace's for France in 1831-1990 period and the 2000 inhabitant threshold; but they differ from those of Moriconi-

[11] The estimates of *c* parameter are all statistically significant at 5% significance level, except in 1950 and 1960, for the full sample, and for the 10 000 inhabitant threshold in 1991 (Atlas database).

Ébrard (1993), for 1981, and Soo (2002), for 2001[12], reinforcing the idea that the estimates of *c* are sensitive to city and threshold definition. In fact, Moriconi-Ébrard (1993) uses urban agglomerations with at least 10000 inhabitants, while Soo (2002) uses Brinkhoff's database[13], with a threshold of 15000 inhabitants.

Figure 5 - Long term evolution of *c* estimates



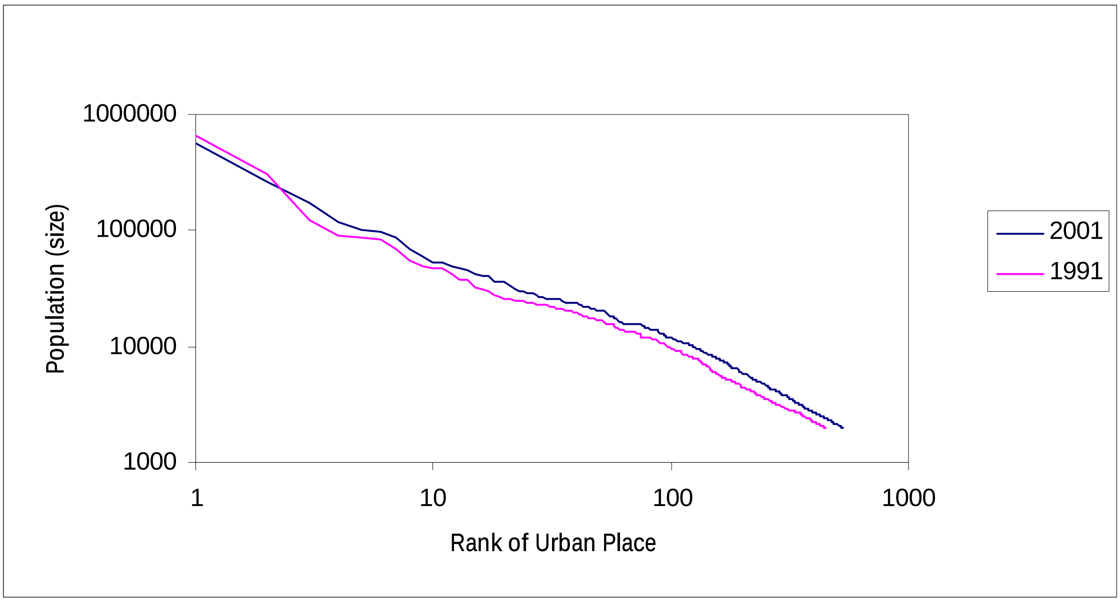
3.4.Comparative analysis of rank size distribution: cities *versus* urban places in the nineties

The proliferation of very small cities in the urban places database, reflected in a median size of about 28% of the corresponding value for the Atlas database (Figure 2), results in a rank-size distribution for urban places that is more scattered than the one we obtain when considering cities (Figure 6). As for the Atlas database, the heavy population losses of Lisboa and Porto produces a downward counter clockwise movement of the rank-size line and a more even distribution.

[12] In both studies c is positive: 0.468 (Moriconi-Ébrard) and 0.124 (Soo).

[13] Comparing the cities in this database with INE's list of legal cities, we conclude that Brinkhoff's definition includes places that are not classified as cities.

Figure 6 – Rank-size distributions of Portuguese urban places, 1991-2001



In order to analyse the sensitivity of parameter estimates to city definition, sample thresholds are the same for both data sets. Table 5 shows the estimates of rank size parameters for both datasets, as well as sample size and R^2 values.

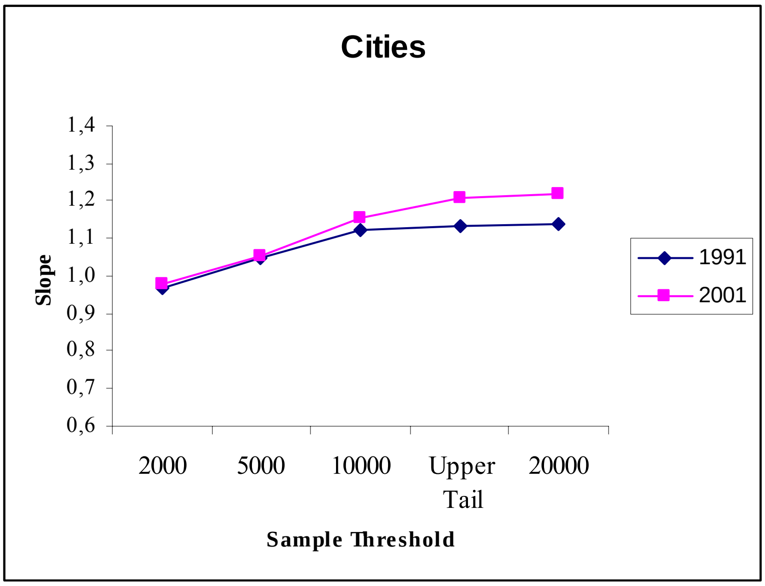
Table 5 - Results of OLS estimation: urban places versus cities (1991-2001)

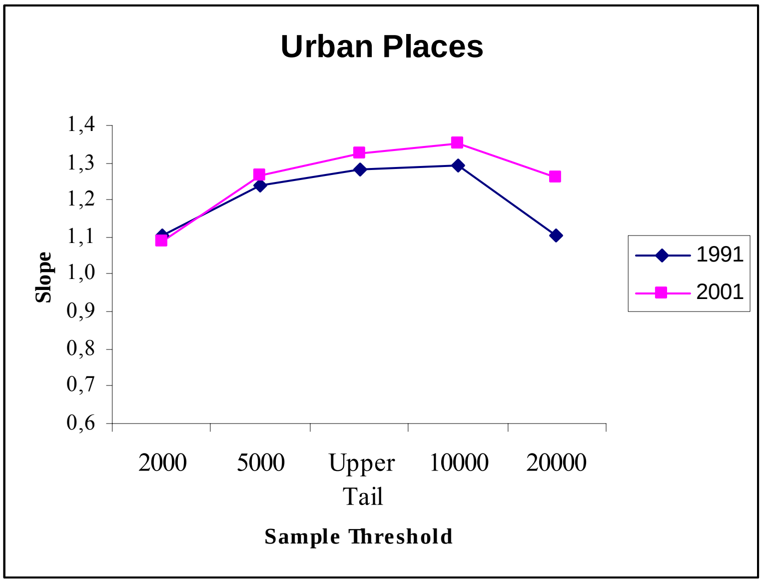
Sample threshold	Census Date	Urban Places			Atlas database		
		Number of cities	Slope	R2	Number of cities	Slope	R2
≥2000	1991	450	1.106	0.985	122	0.970	0.961
	2001	531	1.090	0.983	122	0.977	0.950
≥5000	1991	179	1.241	0.979	112	1.050	0.988
	2001	233	1.264	0.988	115	1.051	0.977
≥10000	1991	97	1.293	0.968	77	1.120	0.993
	2001	123	1.352	0.988	85	1.152	0.991
≥20000	1991	38	1.105	0.957	42	1.138	0.989
	2001	52	1.262	0.979	48	1.221	0.989
Upper tail	1991	150	1.284	0.979	41	1.135	0.989
	2001	178	1.324	0.991	41	1.209	0.988

Looking at the slope estimates for both databases we can observe that, for cities, their values increase with the sample threshold, reflecting a reduction in city size inequality as smaller cities are excluded.

The evolution from 1991 to 2001 shows a decrease in city size inequality which is more important for higher sample cut-offs (Figure 7). For urban places, slope estimates increase with sample threshold until 10000 inhabitants and decrease afterwards. In both samples, the behaviour of the slope for higher cut-offs in 2001, compared with 1991, reflects the growth of middle size urban units and the decrease in the size of the top two units (Lisboa and Porto). Generally, slope estimates for urban places are bigger than those obtained for the city database.

Figure 7 – Sensitivity of slope estimates to urban definition





4. City movements within city size distribution

The precedent analysis of the long-term evolution of city size distribution did not account for the movements that occur within the distribution. By following the position of each city relative to the others, we can examine the movements of cities up or down the city size distribution, over time. For that purpose we use a Markov Chain to describe changes within city size distribution, from 1864 up to the present.

Although the first economic applications of the Markov Chain Process go back to the 1950, urban economists usually refer to the work of Quah (1993) as the keystone reference. In the context of empirical analysis of convergence or divergence between regions or countries, Quah uses a stationary first order Markov Chain to infer about patterns of "inter-temporal evolution of the entire cross section distribution" (Dobkins and Ioannides, 2000, 232). Following this methodology, Eaton and Eckstein (1997) examine the predicted evolution of the size distribution of cities in France and Japan. The same methodology was applied to study the dynamics of the evolution of city size distributions by Dobkins and Ioannides (2000) and Black and Henderson (2003), in the USA, and by Lanaspa *et al.* (2003), in Spain.

4.1 Methodology

Take F_t as the cross section distribution of city sizes at time t. In order to provide a discrete approximation of that distribution we must consider a set of K different size classes or states and calculate the frequency of cities in each state at time t. The evolution of city size distribution is represented by a (K,K) transition probability matrix, M. Each element of this matrix (p_{ij}) indicates the probability that a city belonging to state i in time period t reaches state j in the next period. The transition probabilities are given by:

$$(3)\ p_{ij} = \frac{m_{ij}}{\sum_{j=1}^{n} m_{ij}},\ \text{and} \sum_{j=1}^{n} p_{ij} = 1$$

where m_{ij} is the observed number of cities belonging initially to state i that are in state j in the next period, and n represents the number of possible states. The elements of M are estimated from the relative frequencies of the changing of state between subsequent periods. They are only an approximation of the true probability but, as Anderson and Goodman (1957) show, (3) is the maximum likelihood estimate of the true p_{ij}.

The frequency of cities in each size class in time $t+1$, given by a $(K,1)$ vector F_{t+1}, is described by the following equation:

$$(4) \quad F_{t+1} = M F_t$$

where the $(K,1)$ vector F_t denotes the frequency of cities in each class, at time t.

Admitting that the probabilities between two states are constant over time, then the transition probability matrix is stationary and:

$$(5) \quad F_{t+s} = M^s F_t$$

If the M matrix is regular, the long-term distribution of F_t (or ergodic probability distribution[14]) is obtained taking s to ∞ in equation (5).

$$(6) \quad F_\infty = M^\infty F_t$$

where the resulting $(K,1)$ vector, F_∞, represents the equilibrium distribution of cities obtained under the assumption that the movements observed from t to $t+1$ are repeated as $t \to \infty$.

Considering $M_{t,\, t+1}$ as the transition matrix for the $(t,\ t+1)$ period, we calculate this matrix for all periods in the sample (T) and obtain each element of the estimated average period to period transition matrix ($\overline{M}$), by computing the average of p_{ij} for all the T periods. The ergodic probability distribution is estimated using the ($\overline{M}$) matrix.

4.2. Empirical results

The use of a Markov transition matrix requires the definition of a discrete set of states. Following Eaton and Eckstein (1997) and Lanaspa *et al.* (2003), we defined cell upper points in the size distribution of cities according to their size relative to the average city size on each census date. We obtained seven states, corresponding to the following intervals: more than twice the average (state 1); between the average and twice the average (state 2); between 0.75 and the average (state 3); between 0.50 and 0.75 of the average (state four); between 0.30 and 0.50 of the average (state five); less than 0.30 of the average (state six) and a residual state (state seven) accounting for cities that, on each census date, enter or leave the sample. As our samples were obtained from population census, each period is defined by consecutive census dates and has a variable length.[15]

We estimate the matrix in Table 6 by computing the average of the relative frequency of cities in each state, from the 1864 to 1991[16] inter-census transition matrix. In the average transition matrix, large values in diagonal cells and low values or zeros in the off diagonal cells indicate

[14] Also known as the equilibrium or steady state distribution.
[15] From the middle of the 20th century, inter-census periods correspond to a decennium.
[16] The nature of the data does not allow equal length time periods.

the persistence of the relative position of cities within the distribution; zero values in cells far from the diagonal indicate that there are no drastic movements in the relative position/size of a city from one period to another. In this last case, mobility is a gradual process that occurs between contiguous states.

For Portugal the diagonal terms are higher for larger cities (state 1) and for smaller cities (state 6), that is to say that the probability of moving from the initial state is lower for the cities at the extremes of the distribution. These results indicate that the largest and the smallest cities are less likely to modify their relative position over time. Mobility seems to be higher in intermediate states. In fact, cities with sizes between 0.75 of the average and the average have a 53% probability of remaining in the same state, and cities in the class of 0.50 and 0.75 of the average have a 68% probability of persistence in the same state.

Movements to the adjacent higher state are more probable for small cities (with sizes below 0.50 of the average) whereas cities between 0.75 and twice the average are more likely to move to the next lower state than to climb in the hierarchy. Finally, cities with fewer than 2000 inhabitants (state 7) have a 70% probability of remaining out of the sample and a 29% probability of entering the class of cities with less than 0.30 of the average. On the other hand, for smaller cities (state 6), the probability of dropping out (that is, passing from state 6 to state 7) is about zero[17].

Table 6 - Average transition matrix for Portuguese cities, 1864 to 1991

Cell's upper end points	Cell's upper end points						
	∞	2	1	0.75	0.50	0.30	Out of the sample
∞	**0.986**	0.014	0.000	0.000	0.000	0.000	0.000
2	0.049	**0.798**	0.133	0.020	0.000	0.000	0.000
1	0.012	0.199	**0.530**	0.259	0.000	0.000	0.000
0.75	0.000	0.020	0.118	**0.680**	0.182	0.000	0.000
0.50	0.000	0.004	0.003	0.130	**0.787**	0.076	0.000
0.30	0.000	0.000	0.000	0.000	0.119	**0.877**	0.004
Out of the sample	0.000	0.000	0.000	0.000	0.011	0.291	**0.698**

The transition matrix for the Atlas database, in Table 7, refers to the last inter-census period. The large number of zeros in the off-diagonal cells and the high values in the main diagonal

[17] In fact our sample considers urban places that have the administrative status of "city" and at least 2000 inhabitants. So, the probability of dropping out, taking into account the age of the urban units and the criteria for becoming a city, is necessarily very small.

show a high persistence of the city size distribution of Portuguese cities. As in the previous case, the probability of remaining in the same state is higher for larger and smaller cities than for medium size ones, that is, mobility is more likely to occur in cities with between 0.30 and the average size. Cities changing their relative position within the hierarchy tend to move up to the next state.

Table 7 - Transition matrix for Portuguese cities, 1991 to 2001

	Cell's upper end points						
	∞	2	1	0.75	0.50	0.30	Out of the sample
∞	**1**	0.000	0.000	0.000	0.000	0.000	0.000
2	0.063	**0.938**	0.000	0.000	0.000	0.000	0.000
1	0.000	0.167	**0.750**	0.083	0.000	0.000	0.000
0.75	0.000	0.000	0.333	**0.556**	0.111	0.000	0.000
0.50	0.000	0.000	0.032	0.226	**0.645**	0.097	0.000
0.30	0.000	0.000	0.000	0.000	0.167	**0.833**	0.000
Out of the sample	0.000	0.000	0.000	0.000	0.000	0.000	**1.000**

The ergodic distribution (Table 8) is usually seen as the long run equilibrium distribution of city sizes. It gives an indicator of the tendencies at work within the distribution. The size distribution implied by the ergodic probability is a projection of the distribution of city sizes if the observed pattern of movement continued. The ergodic probability vector calculated from Table 6 shows that the most probable state, in the long-term distribution of Portuguese cities, is the first one ("∞"). Thus, there is a tendency towards the reinforcement of the number of cities that have more than twice the average size. On the other hand, for all the remaining states, except for the residual one, there is a lesser probability of remaining in a given state comparative to the initial distribution. If the above tendencies persist, city-size distribution will be gradually biased towards the relatively larger cities.

Table 8 – Ergodic probabilities

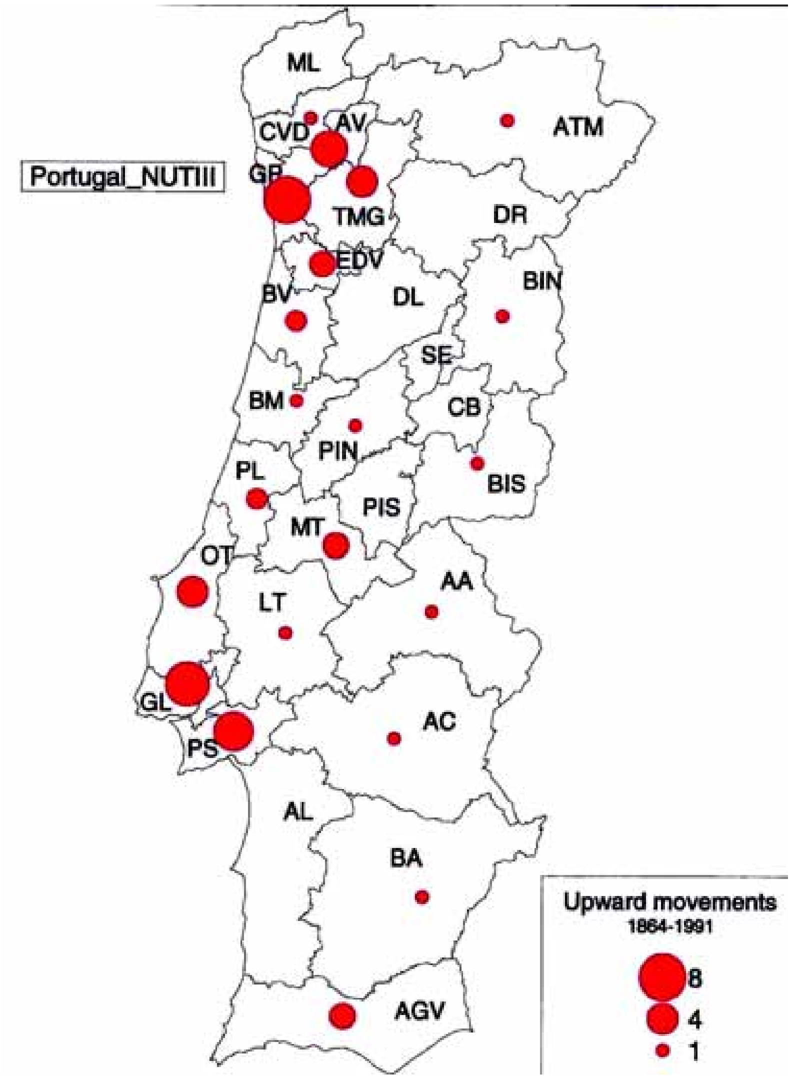
	Cell's upper end points						
	∞	2	1	0.75	0.50	0.30	Out of the sample
Ergodic probability	0,4302	0,1075	0,0629	0,1253	0,1664	0,1062	0,0014

5. Spatial pattern of “winners” and “losers”

In order to analyse the spatial pattern of movements in the Portuguese urban system we confronted the distribution of cities by size classes in 1864 and 1991, identifying cities that move up in their relative position (winners) and cities that move down (losers). From Figure 8[18] we can conclude that the number of upward movements is more than double the number of downward movements. The graph indicates the net gain (or loss) from 1864 to 1991. Some cities have registered a 6 or 5 point gain.

[18] Reproduced in Annex 1

Figure 9 – NUT III distribution of cities registering upward movements (1864-1991)



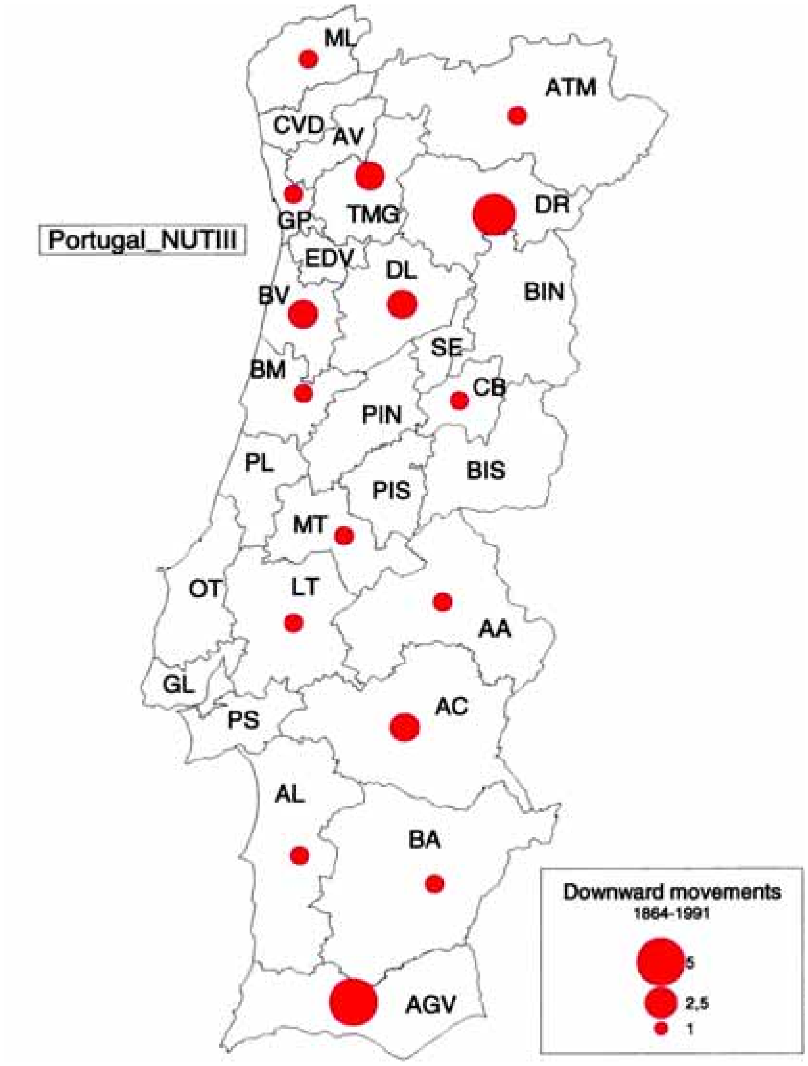
Legend

ML – Minho-Lima
CVD – Cávado
AV- Ave
ATM – Alto Trás-os-Montes
GP – Grande Porto
TMG – Tâmega
DR – Douro
EDV – Entre Douro e Vouga
BV – Baixo Vouga
DL – Dão-Lafões
BIN – Beira Interior Norte
SE – Serra da Estrela
CB – Cova da Beira
BM – Baixo Mondego
PIN – Pinhal Interior Norte
PIS – Pinhal Interior Sul
PL – Pinhal Litoral
MT – Médio tejo
BIS – Beira Interior Sul
OT – Oeste
LT – Lezíria do Tejo
Gl _ Grande Lisboa
PS – Península de Setúbal
AA – Alto Alentejo
AC – Alentejo Central
AL – Alentejo Litoral
BA – Baixo Alentejo
AGV - Algarve

The inspection of the geographical location of cities moving up in the hierarchy suggests the existence of a spatial pattern. Winners seem to concentrate on the coast and especially in and around the two main metropolitan areas (Figure 9). In fact, out of the 57 cities that registered positive changes in their relative position from 1864 to 1991, 40% belong to the metropolitan areas of Lisboa and Porto.

As for cities moving down the size classes (Figure 10), there is a more scattered geographical pattern, although 9 of these are concentrated in just two NUT III (Algarve and Douro).

Figure 10 – NUT III distribution of cities registering downward movements (1864-1991)



Legend
ML – Minho-Lima
CVD – Cávado
AV- Ave
ATM – Alto Trás-os-Montes
GP – Grande Porto
TMG – Tâmega
DR – Douro
EDV – Entre Douro e Vouga
BV – Baixo Vouga
DL – Dão-Lafões
BIN – Beira Interior Norte
SE – Serra da Estrela
CB – Cova da Beira
BM – Baixo Mondego
PIN – Pinhal Interior Norte
PIS – Pinhal Interior Sul
PL – Pinhal Litoral
MT – Médio tejo
BIS – Beira Interior Sul
OT – Oeste
LT – Lezíria do Tejo
Gl _ Grande Lisboa
PS – Península de Setúbal
AA – Alto Alentejo
AC – Alentejo Central
AL – Alentejo Litoral
BA – Baixo Alentejo
AGV - Algarve

The geographical location of winners and losers, suggests the existence of a spatial dependency in the evolution of the urban system. In order to test for spatial dependency we defined X_i as a variable that takes value 1 if the city registers a net gain (net loss) from 1864 to 1991 and 0 otherwise. The contiguity matrix was constructed considering that city *i* and city *j* are contiguous cities if they belong to the same NUT III. Given the nature of the variable, we applied the methodology of Cliff and Ord to test for spatial autocorrelation and computed their *H* statistic for both sets of cities[19]. After normalising, the appropriate values of *H* for upward and downward movements are, respectively, 2.55 and 0.92. In the first case we reject the null hypotheses of spatial independence, at the 5 per cent level of significance. So we can

[19] A more detailed description of the *H* statistics is presented in Annex I.

conclude that the distribution of cities registering upward movements from 1864 to 1991 is not spatially random. As for downward movements, the value of the Cliff and Ord statistic does not allow the rejection of the null hypotheses, for the same level of significance.

6. Conclusions

This paper presents evidence about urban evolution in Portugal over more than a century, focusing on the characteristics and evolving pattern of city size distribution. One limitation of our study relates to the nature of our basic sample. The use of administrative cities leads to the inclusion of very small places and to the exclusion of urban places with considerable population, but lacking the administrative status of city.

The following aspects emerge from our study:

- ✓ The Portuguese urban system is characterised by the proliferation of small cities and two dominant cities, Lisboa and Porto, which are the central cities of Portugal's two metropolitan areas. The long term evolution shows a slow increase in the number of cities. The growth of urban population was faster than that of the number of cities, urban growth resulting mainly from the concentration of population in existing cities. In the last decades, the two top cities have experienced heavy population losses, whereas intermediate cities, especially those on their periphery, have registered significant population gains. As a result, we observe a decrease in the top two primacy index;

- ✓ For the 1991-2001 period, we obtain a different image of the Portuguese urban system if we take the urban place database. The urbanisation rate rises and the number of urban units is substantially higher. The proliferation of very small towns is reflected in a median size of 3934 and 4323 inhabitants, in 1991 and 2001, respectively. At the same time, we observe the emergence of 81 new urban units with 2000 inhabitants or more in a decade;

- ✓ The rank size line shifts up in the course of time as a result of urban growth and becomes smoother, expressing the development of the urban system as a whole, accompanied by a reduction of inequality between city sizes in the upper tail of the distribution;

- ✓ For the basic sample and a 2000 inhabitant cut-off, the Pareto exponent is higher than one but decreasing, which generates a more contrasted urban hierarchy; from 1970 onwards it is less than one and tends to decrease; however, in the last two decades, we detect a reverse in this tendency, pointing to a process of decreasing inequality. When smaller cities are excluded from the sample, slope estimates tends to increase over time, starting from values less than one, reflecting a reduction in city size inequality. In comparison with the full sample, this distinct evolution portrays the changes in growth behaviour of middle-sized cities vis-à-vis the two top cities. The sensitivity of slope estimates to sample threshold is higher at the beginning of our study period; on the other hand, their behaviour permits us to conclude that, in the first part of the period, smaller cities tends to generate a more even distribution, whereas on more recent dates it appears that medium and larger cities are the source of a more equal distribution;

- ✓ For the urban places database, slope estimates are higher than one, indicating that city sizes are more evenly distributed than that predicted by Zipf's law. As smaller places are excluded, we observe a reduction in city size inequality, since the values of the exponent are always superior and increase with sample threshold until 10000 inhabitants;

- ✓ Deviations from rank-size regularity enhance two different processes in the evolution of the urban system: until the middle of the twenty century, urban growth was accompanied by population concentration in the largest cities; afterwards, growth benefits intermediate cities, reinforced in the last decades by heavy population losses in the top two cities;

- ✓ Despite the observed pattern of urban growth - increased concentration in the early phases of the urbanisation process, followed by a tendency of decreased concentration afterwards - we must bear in mind that we are using a city proper database and that the rank size model does not take into account the location of the cities. As a result, the process of decentralisation of urban growth can not be entirely viewed as an inter-urban decentralisation process, since the parameter estimates captures the suburbanisation process of population in the larger cities. In fact, the change in the growth behaviour of the two top cities *vis-à-vis* the middle size cities points to a process of selective growth since it favours mainly cities located closer to the central cities in the metropolitan areas of Lisboa and Porto;

- ✓ The use of a Markov chain process to describe movements within the distribution indicates that mobility is a gradual process that occurs mainly between contiguous size classes; mobility is more likely in intermediate states; there is a persistence of the relative position of cities within the distribution, given that in the average transition matrix we obtain large values in diagonal cells. The long run equilibrium distribution of city sizes reflects a tendency toward the reinforcement of the number of the larger cities;

- ✓ The spatial pattern of "winners" and "losers" between 1864 and 1991 shows that the "winners" tends to concentrate on the coast and especially in and around the two metropolitan areas. The test for spatial autocorrelation leads to the conclusion that the distribution of cities registering upward movements is not spatially random, reinforcing the idea of a selective growth process;

- ✓ On the other hand, the observed pattern of the 81 new urban places with 2000 or more inhabitants that emerge from 1991 to 2001, strengthens that idea. In fact, they are mostly located on the coast and in the two metropolitan areas, as well as in the urban nebula that we can perceive from them.

The evolution of the Portuguese urban system mirrors structural changes that took place mainly in the second half of the 20th century: modern industrialisation, in the fifties, export oriented from the sixties, and economic restructuring in the seventies and the eighties, following severe political changes and the integration in the European Union. It also reflects the evolution from a centralised political regime, administrating vast colonial territories, to a democratic regime, with a more decentralised administrative organisation and confined to its European borders.

References

Albergaria, H. (1999), "A Dinâmica Populacional das Cidades do Continente Português", *Revista de Estatística*, Vol. 2, 47-66

Allen, G.R. (1954), "The 'Courbe des populations', a further analysis", *Bulletin of the Oxford University Institute of Statistics*, Vol. 16, 179-189

Auerbach, F. (1913), "Das gesetz der Bevölkerungskonzentration", Petermans Geographische Mitteilungen, 59, No.1, 74-76

Berry, B.J.L. (1961), "City size distribution and economic development", Economic Development and Cultural Change, Vol. IX, No. 4, 573-587

Bessi-Pietri, P.(2000), "Les formes récentes de la croissance urbaine", Economie et Statistique, No. 336, 35-52

Black, D. and V. Henderson (2003), "Urban Evolution in the USA", *Journal of Economic Geography*, 3, 343-372

Brakman, S., H. Garretsen, C. Van Marrewijk and M. van den Berg (1999), "The Return of Zipf's: Towards a further understanding of the Rank-Size Distribution", *Journal of Regional Science*, Vol. 39, No. 1, 183-213

Carroll, G. (1982), "National city-size distributions: What do we know after 67 years of research?", *Progress in Human Geography*, Vol. 6, 1-43

Carter, H. (1981), *The Study of Urban Geography*, London, Edward Arnold

Cheshire, P. (1999), "Trends in sizes and structures of urban areas", in P. Cheshire and E.S. Mills (eds), *Handbook of Regional and Urban Economics*, Vol. 3, Amsterdam, Elsevier Science B. V., pp. 1339-1373

Chevailler, J.-C.(1978), *Eléments d'Économétrie Spatiale*, Paris, Editions Sirey

Clark, J.S. and J.C. Stabler (1991), "Gibrat's Law and the Growth of Canadian Cities", *Urban Studies*, Vol. 28, No. 4, 635-639

Córdoba, J.-C. (2003), "On the Distribution of City Sizes", *Working Paper*, Rice University

Delgado, A.P. and I.M. Godinho, 2004, The evolution of city size distribution in Portugal: 1864-2001, Working paper No. 151, Faculdade de Economia do Porto

Dobkins, L.H. and Y.M. Ioannides (2000), "The Evolution of Size Distribution: U.S. Cities", in *Economics of Cities: Theoretical Perspectives*, J.M. Huriot and J.-F. Thisse (eds), Cambridge, UK: Cambridge University Press

Duranton, G. (2002), "City Size Distribution As A Consequence of the Growth Process", Centre for Economic Performance, in http://cep.lse. ac.uk/pubs accessed in December 2003

Eaton, J. and Z. Eckstein (1994), "Cities and Growth: Theory and Evidence from France and Japan", *Working Paper* No. 4612, National Bureau of Economic Research

Eaton, J. and Z. Eckstein (1997), "Cities and growth: Theory and evidence from France and Japan", *Regional Science and Urban Economics*, Vol. 27, 443-474

Fan, C.C. and E. Casetti (1994), "*The spatial and temporal dynamics of US regional income inequality*", Annals of Regional Science, Vol. 28, 177-196

Ferreira, L., M.C. Cardoso and S.A.Silva (2003), *Evolução do Sistema Urbano Português*, Faculdade de Economia do Porto, unpublished

Fujita, M., P. Krugman and A. Venables (1999), *The Spatial Economy. Cities, Regions and International Trade*, Cambridge, Mass.:The M.I.T. Press

Gabaix, X. (1999a), "Zipf's Law and the growth of cities", *American Economic Review* (AEA Papers and Proceedings), Vol. 89, 129-132

Gabaix, X. (1999b), "Zipf's Law for Cities: An Explanation", *The Quarterly Journal of Economics*, Vol. CXIV, No. 3, 739-767

Gabaix, X. and Y.M. Ioannides (2003), "The Evolution of City Size Distributions", *Working Paper* 2003, Tufts University, in http://ase.tufts.edu/econ/papers , accessed in December 2003

Gabaix, X. and Y.M. Ioannides (2003), "The Evolution of City Size Distributions", in J.V-Henderson and J.-F. Thisse (eds) *Handbook of Urban and Regional Economics*, in http://www.core.ucl.ac.be/staff/ThisseHandbook/Handbook/html, accessed in December 2003

Gibrat, R. (1931), *Les inégalités économiques*, Paris : Librairie du Recueil Sirey

Girão, A.A. and F.O.L. Velho (1944), *Evolução Demográfica e Ocupação do Solo Continental (1890-1940)*, Centro de Estudos Geográficos, Faculdade de Letras, Universidade de Coimbra, Coimbra

Goodrich, E.P., (1926), *The Statistical Relationship Between Population and the City Plan*, Chicago: The Urban Community.

Guérin-Pace, F. (1995), "Rank-Size Distributions and the Process of Urban Growth", *Urban Studies*, Vol. 32, No. 3, pp. 551-562

Henderson, J.V. (1974), "The sizes and types of cities", The American Economic Review, Vol. 64, No.4, 640-656

INE (2002), *Atlas das Cidades,* Lisboa, Instituto Nacional de Estatística

Ioannides, Y.M. and H.G. Overman (1999), "Dynamic Evolution of the U.S. City Size Distribution", *Working Paper* No. 99-16, Tufts University, in http://ase.tufts.edu/econ/papers accessed in December 2003

Ioannides, Y.M. and H.G. Overman (2000), "Spatial Evolution of the US Urban System", *Working Paper* 2000-18, Tufts University, in http://ase.tufts.edu/econ/papers accessed in December 2003

Ioannides, Y.M. and H.G. Overman (2003), "Zipf's law for cities: an empirical examination", *Regional Science and Urban Economics*, 33, 127-137

Ioannides, Y.M. and L.H. Dobkins (1999), "Dynamic Evolution of the U.S. City Size Distribution", *Working Paper* 1999-16, Tufts University, in http://ase.tufts.edu/econ/papers accessed in December 2003

Jayet, H. (1993), *Analyse Spatiale Quantitative. Une Introduction*, Paris, Economica

Knudsen, T. (2001), "Zipf's law for cities and beyond: The case of Denmark", ″, Jan. 2001

Krugman, P. (1995), "Innovation and Agglomeration: Two parables suggested by city-size distributions", *Japan and World Economy*, Vol. 7, 371-390

Krugman, P. (1996), "Confronting the Mystery of Urban Hierarchy", *Journal of the Japanese and International Economies*, Vol. 10, 399-418

Lanaspa, L., A.M. Perdiguero and F. Sanz (2004), "La distribución del tamaño de las ciudades en España, 1900-1999", *Revista de Economia Aplicada*, No. 34 (vol. XII), 5-16

Lanaspa, L., F. Pueyo and F. Sanz (2003), "The evolution of Spanish urban structure during the twentieth century", *Urban Studies*, Vol. 40, No. 3, 567-580

Lasuén, J.R., A. Lorca and J. Oria (1967), "City-size distributions and economic growth", *Ekistics*, Vol. 24, 221-226

Le Gallo, J. (2001), "Space-time Analysis of GDP Disparities among European Regions: A Markov chains approach", 2001-06, LATEC, Laboratoire d'Analyse des Techniques Economiques, Université de Bourgogne

Lembo, A.J. Jr., "Spatial Autocorrelation", in http://www.css.cornell.edu/courses/620/lecture10, accessed on 2005/04/25

Lopes, A.S. (1971), *As Funções Económicas dos Pequenos Centros*, Lisboa: Instituto Gulbenkian de Ciências.

Lopes, A.S. (1979), *Desenvolvimento Regional*, Lisboa: Fundação Calouste Gulbenkian

Lotka, A. (1925), *Elements of Physical Biology*, Baltimore: Williams & Wilkins

Mills, E.S. and B.W. Hamilton (1994), *Urban Economics*, New York: HarperCollins College Publishers

Moriconi-Ebrard, F. (1993), *L'Urbanisation du Monde depuis 1950*, Paris: Anthropos

Overman, H.G. and Y.M. Ioannides (2000), "Cross-Sectional Evolution of the US City Size Distribution", in http://cep.lse.ac.uk/pubs/ accessed in December 2003

Parr, J. (1976), "A Class of Deviations from Rank-Size Regularity: Three Interpretations", *Regional Studies*, Vol. 10, 285-292

Parr, J. (1985), "A Note on the Size Distribution of Cities over Time", *Journal of Urban Economics*, Vol. 19, 535-546

Quah, D. (1993), "Empirical cross-section dynamics in economic growth", *European Economic Review*, 37, 426-434

Reed, W. J. (2002), "On the Rank-Size Distribution for Human Settlements", *Journal of Regional Science*, Vol. 42, No. 1, 1-17

Reed, W.J. (2001), "The Pareto, Zipf and other power laws", *Economics Letters*, 74, 15-19

Richardson, H. W. (1973), "Theory of the distribution of city sizes: Review and prospects2, *Regional Studies*, Vol. 7, 239-251

Rosen, K.T. and M. Resnick (1980), "The Size Distribution of Cities: An Examination of the Pareto Law and Primacy", *Journal of Urban Economics*, Vol. 8, 165-186

Salgueiro, T.B. (1992), *A Cidade em Portugal. Uma Geografia Urbana*, Porto: Edições Afrontamento

Simon, H. (1955), "On a class of skew distribution functions", *Biometrika*, Vol. 42, 425-440

Singer, H.W. (1936), "The 'courbe des populations.' A parallel to Pareto's Law", *The Economic Journal*, Vol. XLVI, No. 182, 254-263

Soo, K. T. (2002), "Zipf's Law for Cities: A Cross Country Investigation", *Working Paper*, Centre for Economic Performance, London School of Economics

Suarez-Villa, L. (1988), "Metropolitan evolution, sectoral economic change and the city size distribution", *Urban Studies*, Vol.25, 1-20

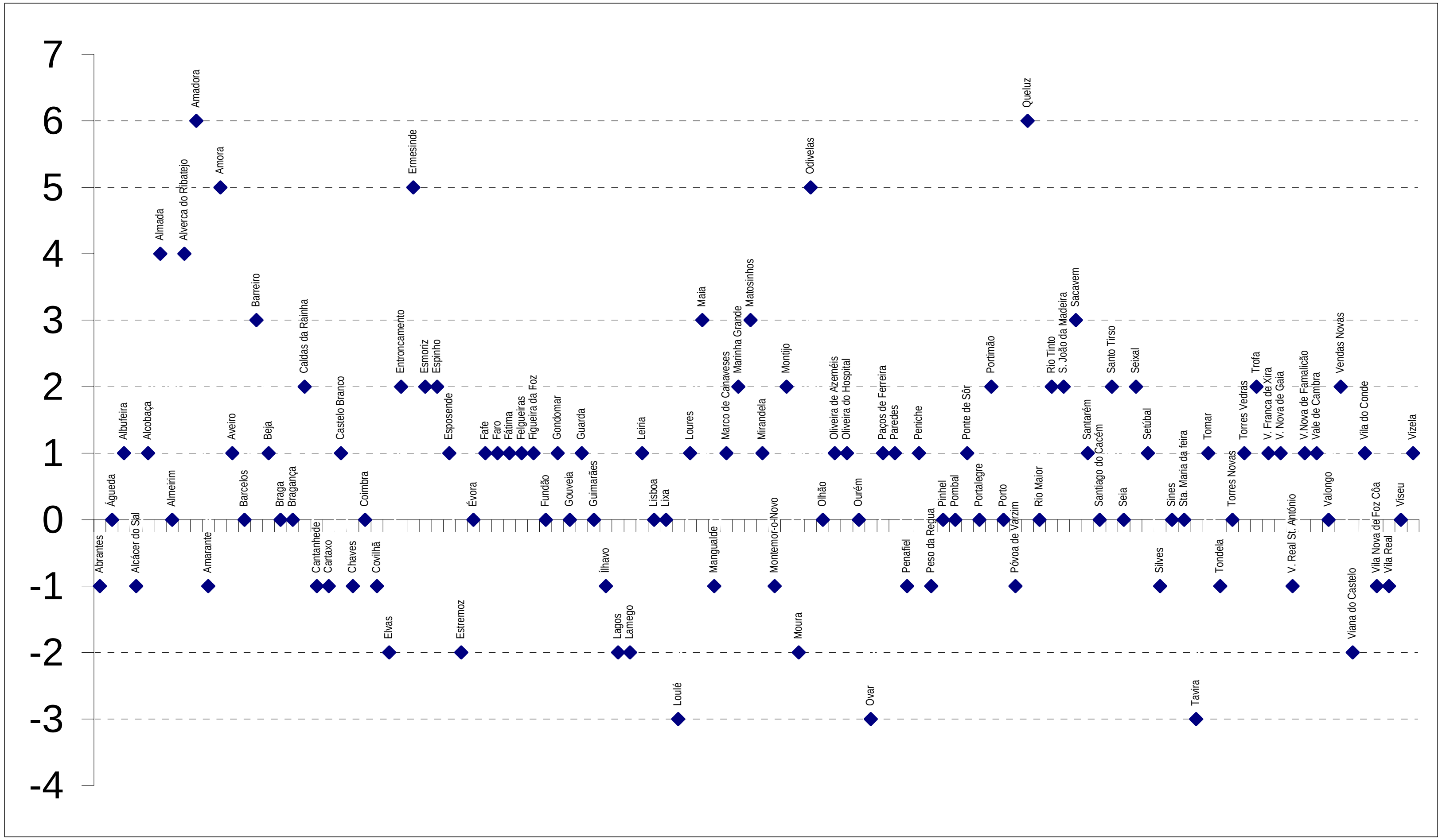
Tabuchi, T., J.-F. Thisse and D.-Z. Zeng (2003), "On the Number and Size of Cities", *Discussion papers*, Centre for Economic Policy Research, in http://www.core.ucl.ac.be/staff/thissePapers/NR61.pdf, accessed in December 2003

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*, Reading, M.A.: Addison-Wesley

Annex I

Figure 8 – Net movements, 1864-1991



Annex 2

Testing for spatial autocorrelation

Consider a set of spatial units, cities, characterised by a binary variable X_i, that takes the value 1 if city i registers a net gain (a net loss) in its relative position from period t to period $t+n$, and zero if not.

Let $C=\{c_{ij}\}$ be the contiguity matrix whose elements are defined as follows: if city i and city j belong to the same *NUT III* then $c_{ij} = 1$, if they belong to different *NUT III*, $c_{ij} = 0$. In this matrix $c_{ij}=c_{ji}$ and $c_{ii}=0, \forall\ i$.

We can compute $\sum_{j=1}^{n} c_{ij} = L_i$ as the total number of cities that belong to the same *NUT III* as city i (total number of contiguities for city i), and $\sum_{i=1}^{n} L_i = \mathrm{A}$. We defined $L=A/2$ as the total number of contiguous cities or join-counts in the set of cities.

If we have n cities, with n_1 cities registering a net gain (net loss) and n_2 registering a loss (gain) or maintaining their relative position ($n=n_1 + n_2$), the probability of a winning (losing) city is:

$$p(n_1) = {n_1}/{n}$$

and the probability of a loss (gain) or non change in relative position of a city is:

$$p(n_2) = {n_2}/{n} \text{, with } p(n_1) + p(n_2) = 1$$

The probability of having two contiguous cities registering a net gain (net loss) is:

$$p(n_1) * p(n_1) \text{ or } [p(n_1)]^2$$

With L contiguous cities, the expected number of winning (losing) cities that are contiguous is given by

$$E(WW) = [p(n_1)]^2 L$$

and the correspondent standard deviation is:

$$\sigma(WW) = \sqrt{[p(n_1)]^2 L + [p(n_1)]^3 K - [p(n_1)]^4 (L+K)} \text{, where } K = {1}/{2} \sum_{i=1}^{n} c_i (c_i - 1)$$

In order to test for spatial autocorrelation, given the nature of X_i, we must calculate the Cliff and Ord (1981) statistic H:

$$H = \frac{1}{2}\sum_{i \neq j}^{n}\sum_{j=1}^{n} c_{ij} x_i x_j ,$$

where x_i and x_j are cities and c_{ij} is the correspondent value of the contiguity matrix. H is the number of times that two winning (losing) cities are located in the same *NUT III*.

The appropriate test statistic, the normalised value of H, is:

$$t = \frac{H - E(WW)}{\sigma(WW)}$$

The relevant null hypothesis (H_0) is the existence of no spatial structure. If $| t | > |t_\alpha|$, where t_α is given by the table of the standardised normal, we can reject H_0 at the α per cent level of significance. In this case we can conclude that the distribution of winning (losing) cities is not spatially random.

Recent FEP Working Papers

Nº 229	Sandra Tavares Silva and Aurora A. C. Teixeira, *"On the divergence of evolutionary research paths in the past fifty years: a comprehensive bibliometric account"*, September 2006
Nº 228	Argentino Pessoa, *"Public-Private Sector Partnerships in Developing Countries: Prospects and Drawbacks"*, September 2006
Nº 227	Sandra Tavares Silva and Aurora A. C. Teixeira, *"An evolutionary model of firms' institutional behavior focusing on labor decisions"*, August 2006
Nº 226	Aurora A. C. Teixeira and Natércia Fortuna, *"Human capital, trade and long-run productivity. Testing the technological absorption hypothesis for the Portuguese economy, 1960-2001"*, August 2006
Nº 225	Catarina Monteiro and Aurora A. C. Teixeira, *"Local sustainable mobility management. Are Portuguese municipalities aware?"*, August 2006
Nº 224	Filipe J. Sousa and Luís M. de Castro, *"Of the significance of business relationships"*, July 2006
Nº 223	Pedro Cosme da Costa Vieira, *"Nuclear high-radioactive residues: a new economic solution based on the emergence of a global competitive market"*, July 2006
Nº 222	Paulo Santos, Aurora A. C. Teixeira and Ana Oliveira-Brochado, *"The 'de-territorialisation of closeness' - a typology of international successful R&D projects involving cultural and geographic proximity"*, July 2006
Nº 221	Manuel M. F. Martins, *"Dilemas macroeconómicos e política monetária: o caso da Zona Euro"*, July 2006
Nº 220	Ana Oliveira-Brochado and F. Vitorino Martins, *"Examining the segment retention problem for the "Group Satellite" case"*, July 2006
Nº 219	Óscar Afonso Rui and Henrique Alves, *"To Deficit or Not to Deficit": Should European Fiscal Rules Differ Among Countries?"*, July 2006
Nº 218	Rui Henrique Alves and Óscar Afonso, *"The "New" Stability and Growth Pact: More Flexible, Less Stupid?"*, July 2006
Nº 217	J Maciej Cieślukowski and Rui Henrique Alves, *"Financial Autonomy of the European Union after Enlargement"*, July 2006
Nº 216	Joao Correia-da-Silva and Carlos Hervés-Beloso, *"Prudent Expectations Equilibrium in Economies with Uncertain Delivery"*, June 2006
Nº 215	Maria Rosário Moreira and Rui Alves, *"How far from Just-in-time are Portuguese firms? A survey of its progress and perception"*, June 2006
Nº 214	Maria Fátima Rocha and Aurora A.C. Teixeira, *"A cross-country evaluation of cheating in academia: is it related to 'real world' business ethics?"*, June 2006
Nº 213	Maria Rosário Moreira and Rui Alves, *"Does Order Negotiation Improve The Job-Shop Workload Control?"*, June 2006
Nº 212	Pedro Cosme da Costa Vieira and Aurora A. C. Teixeira, *"Human capital and corruption: a microeconomic model of the bribes market with democratic contestability"*, May 2006
Nº 211	Ana Teresa Tavares and Aurora A. C. Teixeira, *"Is human capital a significant determinant of Portugal's FDI attractiveness?"*, May 2006
Nº 210	Maria Rosário Moreira and Rui Alves, *A new input-output control order release mechanism: how workload control improves manufacturing operations in a job shop*, April 2006
Nº 209	Patrícia Teixeira Lopes and Lúcia Lima Rodrigues, *Accounting for Financial Instruments: An Analysis of the Determinants of Disclosure in the Portuguese Stock Exchange*, April 2006
Nº 208	Pedro Cosme Costa Vieira, *Energia nuclear: Uma solução para*

	Portugal?, April 2006
Nº 207	Aurora A.C. Teixeira and Joana Costa, *What type of firm forges closer innovation linkages with Portuguese Universities?*, March 2006
Nº 206	Joao Correia-da-Silva and Carlos Hervés-Beloso, *Rational Expectations Equilibrium in Economies with Uncertain Delivery*, March 2006
Nº 205	Luís Delfim Santos and José Varejão, *Employment, Pay and Discrimination in the Tourism Industry*, February 2006
Nº 204	Carlos F. Alves and Victor Mendes, *Mutual fund flows' performance reaction: does convexity apply to small markets?*, February 2006
Nº 203	Carlos F. Alves and Victor Mendes, *Are mutual fund investors in jail?*, February 2006
Nº 202	Óscar Afonso and Paulo B. Vasconcelos, *Numerical computation for initial value problems in economics*, February 2006
Nº 201	Manuel Portugal Ferreira, Ana Teresa Tavares, William Hesterly and Sungu Armagan, *Network and firm antecedents of spin-offs: Motherhooding spin-offs*, February 2006
Nº 200	Aurora A.C. Teixeira, *Vinte anos (1985-2005) de FEP Working Papers: um estudo sobre a respectiva probabilidade de publicação nacional e internacional*, January 2006
Nº 199	Samuel Cruz Alves Pereira, *Aggregation in activity-based costing and the short run activity cost function*, January 2006
Nº 198	Samuel Cruz Alves Pereira and Pedro Cosme Costa Vieira, *How to control market power of activity centres? A theoretical model showing the advantages of implementing competition within organizations*, January 2006
Nº 197	Maria de Fátima Rocha and Aurora A.C. Teixeira, *College cheating in Portugal: results from a large scale survey*, December 2005
Nº 196	Stephen G. Donald, Natércia Fortuna and Vladas Pipiras, *Local and global rank tests for multivariate varying-coefficient models*, December 2005
Nº 195	Pedro Rui Mazeda Gil, *The Firm's Perception of Demand Shocks and the Expected Profitability of Capital under Uncertainty*, December 2005
Nº 194	Ana Oliveira-Brochado and Francisco Vitorino Martins, *Assessing the Number of Components in Mixture Models: a Review.*, November 2005
Nº 193	Lúcia Paiva Martins de Sousa and Pedro Cosme da Costa Vieira, *Um ranking das revistas científicas especializadas em economia regional e urbana*, November 2005
Nº 192	António Almodovar and Maria de Fátima Brandão, *Is there any progress in Economics? Some answers from the historians of economic thought*, October 2005
Nº 191	Maria de Fátima Rocha and Aurora A.C. Teixeira, *Crime without punishment: An update review of the determinants of cheating among university students*, October 2005
Nº 190	Joao Correia-da-Silva and Carlos Hervés-Beloso, *Subjective Expectations Equilibrium in Economies with Uncertain Delivery*, October 2005
Nº 189	Pedro Cosme da Costa Vieira, *A new economic journals' ranking that takes into account the number of pages and co-authors*, October 2005

Editor: Prof. Sandra Silva (sandras@fep.up.pt)
Download available at:
http://www.fep.up.pt/investigacao/workingpapers/workingpapers.htm
also in http://ideas.repec.org/PaperSeries.html

www.fep.up.pt

Faculdade de Economia da Universidade do Porto
Rua Dr. Roberto Frias, 4200-464 Porto | Tel. 225 571 100
Tel. 225571100 | www.fep.up.pt