Patterns in U.S. urban growth (1790–2000)

González-Val, Rafael and Lanaspa, Luis

20 May 2011

Patterns in U.S. Urban Growth (1790–2000)

Rafael González-Val[a]

Luis Lanaspa[b]

[a] Universitat de Barcelona & Institut d'Economia de Barcelona

[b] Universidad de Zaragoza

Abstract: This paper reconsiders the evolution of the growth of American cities since 1790 in the light of new theories of urban growth. Our null hypothesis for long-term growth is random growth. We obtain evidence supporting random growth against the alternative of mean reversion (convergence) in city sizes using panel unit root tests. We also examine mobility within the distribution to try to extract growth patterns different from the general unit root trend detected. We find evidence of high mobility when we model growth as a first-order Markov process. Finally, using a cluster procedure we find strong evidence in favour of conditional convergence in city growth rates within convergence clubs, which we can interpret as "local" mean-reverting behaviours. Both the high mobility and the results of the clustering analysis seem to indicate a sequential city growth pattern.

Keywords: city size, urban growth, random growth, sequential city growth, transition matrices, club convergence

JEL: C12, 018, R11, R12

1. Introduction

This paper reconsiders the evolution of the growth of American cities since 1790 in the light of new theories of urban growth, paying special attention to sequential city growth theories. The urban system of the United States (US) has often been studied, because of its special characteristics. First, it is a relatively young system (the first census by the US Census Bureau dates from 1790) characterized by the entry of new cities (Dobkins and Ioannides, 2000). Also, its inhabitants present very high mobility; Cheshire and Magrini (2006) estimate that mobility in the US is 15 times higher than that in Europe. Both characteristics, high mobility and the entry of new cities, should reduce the time transition to spatial equilibrium between cities. In line with this, González-Val (2010) finds that the last decades of the twentieth century are characterized by stability in the number of cities and the percentage of the US total population they represent, indicating a shift to a stable city size distribution and a more consolidated urban landscape. Finally, industry cycles have an important effect on the growth rates of American cities (Duranton, 2007). Thus, in the second half of the nineteenth century and the early twentieth century the growing urban population was concentrated in the north-eastern region known as the manufacturing belt, while in the second half of the twentieth century the rise of the Sun Belt (a phenomenon known as regional inversion; Lanaspa-Santolaria et al., 2002) attracted population to the West Coast area.

Many papers study the long-term evolution of American city growth. These include Dobkins and Ioannides (2000, 2001), Kim (2000), Beeson et al. (2001), Overman and Ioannides (2001), Black and Henderson (2003), Ioannides and Overman (2003), Kim and Margo (2004), González-Val (2010) and Michaels et al. (2010). The spatial units (states, counties, minor civil divisions, metropolitan areas, incorporated places, etc.) and time periods studied and the statistical and econometric methods used in the literature vary widely.

Our aim is to analyse the evolution of the largest American cities from the beginning of the urban system in 1790. Such a wide time horizon enables us, first, to consider the effect of the entry of new cities (most of them during the nineteenth century), and second, to look for different patterns of city growth. New theories have recently emerged that examine both aspects, concluding that historically, city growth may be sequential. Sequential city growth means that cities have early periods of fast growth (from their date of entry as a city) followed by slow growth and/or stagnation. The idea is that during some periods, the largest cities that entered the distribution first are the ones that grow most. Later their growth slows, and the smaller cities that entered later are the ones that grow most. When these reach a certain size their growth rate slows again and other smaller cities are the ones that grow fastest, and so on. It should be noted that the final result is convergence among cities. This convergence is not in size, as the final city size is determined by other factors such as amenities, city productivity, land availability, etc., but in the growth rates in steady state.

Only two papers model sequential city growth: Henderson and Venables (2009) and Cuberes (2009). The model developed by Henderson and Venables (2009) examines city formation in a country whose urban population is growing steadily over time, with new cities required to accommodate this growth. It yields sequential formation of cities, where new cities grow from scratch to a stationary size. The basic assumptions are that city formation requires investment in fixed capital in the form of housing and urban infrastructure and that agents are forward-looking. Cuberes (2009) presents another model of sequential city growth; the key to generating sequential growth is the assumption of irreversible investment in physical capital. The predictions of this second model are empirically tested by Cuberes (2011), who finds strong support for sequential city growth using two comprehensive data sets on populations of cities and metropolitan areas for a large set of countries.

The next section presents the data used. Our basic hypothesis for long-term growth is random growth. We use random growth as a benchmark because the effect of other factors (locational fundamentals or increasing returns) may change over time when such a long period is considered due to the decrease in transport costs (Davis and Weinstein, 2002). Moreover, Ioannides and Overman (2003) and González-Val (2010) find that random growth is a good description of city size growth in the US during the twentieth century. Therefore, in Section 3 we test random growth versus mean reversion (convergence) in US cities using panel unit root tests. We obtain evidence supporting random growth against the alternative of mean reversion in city sizes. In Section 4 we examine mobility within the distribution to try to extract growth patterns different from the general unit root trend. We use two different techniques. First (Section 4.1), we calculate transition matrices, which tell us the degree of mobility in terms of probability, applying a generalized equation to enable cities to enter and leave the sample. Second (Section 4.2), we apply a cluster algorithm to identify different groups of cities that converge with each other. The results point to a certain type of sequential growth, at least within groups. We discuss the different empirical results in Section 5, and conclude in Section 6.

2. Data

There are various ways of defining a "city". The evolution of the American urban structure has been analysed using different geographical units: counties (Beeson et al., 2001), minor civil divisions (Michaels et al., 2010), metropolitan areas (Dobkins and Ioannides, 2000, 2001; Black and Henderson, 2003; Ioannides and Overman, 2003), urbanized areas (Garmestani et al., 2008) or the economic areas recently defined by Rozenfeld et al. (2011) using the city clustering algorithm (CCA). However, since our aim is to study the evolution of the urban system from its origin, we must use data from the "legal" cities, which are those reported since the first census in 1790.[1] Units such as metropolitan areas were introduced later.[2] Thus, we identify cities as what the US Census Bureau denominates incorporated places. These places have also been used recently in empirical analyses of American city size distribution (Eeckhout, 2004, 2009; Levy, 2009; Giesen et al., 2010; González-Val, 2010).

The US Census Bureau uses the generic term "incorporated place" to refer to a type of governmental unit incorporated under state law as a city, town (except New England states, New York and Wisconsin), borough (except in Alaska and New York) or village, with legally established limits, powers and functions. We take our data from the US Census Bureau (2004);[3] the sample consists of all the incorporated places with 100,000 inhabitants or more in 2000.[4]

---

[1] We talk about the "origin" of the urban system because the 1790 census is the first one, and provides data for the first 16 cities. However, these cities existed earlier. Kim (2000) gives data for 4 and 5 cities in 1690 and 1720, respectively. His data come from Bridenbaugh (1938) and the Historical Statistics of the United States. However, we prefer to use a single source of data, the US Census Bureau. Also, the periodicity of these data would not be the same as the rest of the sample (decennial census).

[2] The standard definitions of the metropolitan areas were first issued in 1949 by the then Bureau of the Budget, the predecessor of the present Office of Management and Budget (OMB).

[3] Source: Table 32. Only 16 of all the cities (8.42%) show a significant change in their boundaries (the case of annexed areas): Anchorage, Boston, Columbus, Hampton, Honolulu CDP, Indianapolis, Jacksonville, Lexington-Fayette, Nashville-Davidson, Newport News, New York, Philadelphia, Pittsburgh, Virginia Beach, Washington and Winston-Salem. Information about entities whose name and/or boundary have changed, entities that no longer exist, newly established entities (both legal and statistical) and changes in geographic relationships is given in the "geographic change notes" section.

[4] Imposing a minimum population threshold is relevant for the analysis of city size distribution (Eeckhout, 2004). However, it seems to be less decisive in the study of city growth. González-Val (2010) obtains the same conclusion, using data from all incorporated places without any size restriction, as do Ioannides and Overman (2003) with their sample of MSAs: the validity of random growth in the US city

Unincorporated places (concentrations of population that form no part of an incorporated place but that are locally identified with a name) are excluded, because they began to be counted after 1950 (they were renamed census designated places (CDPs) in 1980). Although some of them are consolidated as incorporated places and are reported in the 2000 census as cities, we also exclude them. The only exception is Honolulu CDP, because due to a Hawaiian state law there are no incorporated places there; they are all unincorporated.

Therefore, our final sample in 2000 is the 190 largest cities. This sample size is similar to that of other studies using MSAs. Black and Henderson (2003) use data from 194 (1900) to 282 (1990) MSAs, while the sample of Ioannides and Overman (2003) ranges from 112 (1900) to 334 (1990). Their samples are slightly larger because in the US to qualify as an MSA a central city of 50,000 or more inhabitants is needed (a lower minimum population threshold than ours). In fact, most of these incorporated places are the central city of an MSA.

Table 1 shows the sample sizes for each decade and the descriptive statistics. For the first decades and until the mid-nineteenth century, the number of cities is low and grows very slowly; however, these few cities represent about two-thirds of the total urban population of the period. From 1850 to 1900 the number of cities doubles (from 73 to 157). The last major entry of new cities takes place from 1900 to 1930, and from that date the number of cities remains stable. In 2000 the percentage of the urban population represented by this upper-tail distribution is much lower (31%), due to the appearance of many small and mid-sized cities (there were 19,296 incorporated places in the 2000 census, with an average population of 8,968.44 inhabitants) and the change that had taken place to a more consolidated urban landscape.

The size of our sample is an advantage from the methodological point of view, as the techniques we apply are specially designed for small samples. However, the sample is defined according to the largest cities in the latest period, which might imply a slight bias, as these are the "winning" cities, cities that have presented the highest growth rates over time. We deal with this problem in Sections 3 and 4.2 where this possible bias could have an influence.

---

growth during the twentieth century. Cuberes (2011) carries out several robustness checks and his results for sequential city growth do not vary much with different cut-offs for selected cities.

3. Testing long-term trends: random growth versus mean reversion

Description

Random growth theories are based on stochastic growth processes and probabilistic models. The most important models are those of Champernowne (1953), Simon (1955) and more recently Gabaix (1999) or Córdoba (2008). In the case of population growth these models are able to reproduce two empirical regularities that are well known in urban economics: Zipf's and Gibrat's laws (or the rank-size rule and the law of proportionate growth).

Random growth theory is especially important from our long-term perspective, because the influence of other factors such as locational fundamentals or increasing returns may change (or even disappear) over time. Locational fundamentals are exogenous factors linked to the physical landscape, such as temperature, rainfall, access to the sea, the presence of natural resources or the availability of arable land. These characteristics are randomly distributed across space, and although they may have played a crucial role in early settlements, one would expect their influence to decrease over time.[5] On the other hand, urban increasing returns, also known as agglomeration economies, appear later as a consequence of industrial development. The empirical literature on agglomeration economies and their positive effects on urban growth is wide, although there is a great deal of variability in the results reported in the literature; see the meta-analysis by Melo et al. (2009).

Therefore, our basic hypothesis for long-term growth is random growth (or Gibrat's law[6]). We will follow the methodology proposed by Clark and Stabler (1991), who suggested that testing for random growth is equivalent to testing for the presence of a unit root. They build up in the Vining model of city growth with autocorrelated errors (Vining, 1976). Let S_{it} be the size (population) of city i at time t. Also, assume that the relationship between the size of a city in time period t and $t-1$ is

$$S_{it} = \gamma_{it} S_{it-1}, \qquad (1)$$

[5] However, empirical studies demonstrate that in some cases their influence in determining agglomeration still remains important; see Ellison and Glaeser (1999) or Davis and Weinstein (2002).

[6] According to Gabaix and Ioannides (2004), *"Gibrat's Law states that the growth rate of an economic entity (firm, mutual fund, city) of size S has a distribution function with mean and variance that are independent of S."*

where γ_{it} is the growth rate of city i over the period $t-1$ to t. Now suppose that this growth rate can be decomposed into three[7] components: a random component ε_{it}, a non-stochastic component relating the current growth rate to a (possibly time-varying) constant and past growth rates, and initial city size:

$$\gamma_{it} = K_{it} S_{it-1}^{\delta_i} \prod_{j=1}^{p} \gamma_{it-j}^{\beta_{ij}} \left(1+\varepsilon_{it}\right), \tag{2}$$

where K_{it} is a possibly time-varying constant, and δ_i and β_{ij} are parameters measuring the relative importance of the initial city size and past growth rates on current city growth, respectively, and ε_{it} is a random error term. Random growth would imply $\delta_i = 0$, meaning that the growth of a particular city does not depend on the initial city size. Substituting equation (2) into equation (1), taking logs and subtracting $\ln S_{it-1}$ from both sides of the equation, one obtains:

$$\Delta \ln S_{it} = k_{it} + \rho_i \ln S_{it-1} + \sum_{j=1}^{p} \beta_{ij} \Delta \ln S_{it-j} + \varepsilon_{it}, \tag{3}$$

where $k_{it} = \ln K_{it}$, $\rho_i = \delta_i$ and the following approximate equality is used: $\ln\left(1+\varepsilon_{it}\right) \approx \varepsilon_{it}$ for small values of ε_{it}. This shows that testing for random growth (Gibrat's law) is equivalent to testing for a unit root in city sizes. If we find evidence in favour of a unit root (ρ_i is not significantly different from zero), this means that city i's growth rate is independent of city i's initial size. On the other hand, when $\rho_i < 0$ the evolution of city i will be a stationary process and city i's growth rate declines with the initial city size (there is mean reversion in the stochastic growth process).[8] Starting from equation (3) Clark and Stabler (1991) estimate the standard Dickey–Fuller regressions, not rejecting random growth for the seven largest cities in Canada from 1975 to 1984.

Results

Gabaix and Ioannides (2004) emphasize "*that the next generation of city evolution empirics could draw from the sophisticated econometric literature on unit roots*". In

[7] We apply the extended version of the model by Bosker et al. (2008).

[8] A consequence of an estimated $\rho_i < 0$ is that any shock will dissipate over time; see Davis and Weinstein (2002).

line with this suggestion, most of the recent studies apply unit root tests: Black and Henderson (2003), Sharma (2003), Resende (2004), Henderson and Wang (2007) and Bosker et al. (2008).

Some of these authors (Black and Henderson, 2003; Henderson and Wang, 2007; Soo, 2007) test the presence of a unit root by proposing a growth equation, which they estimate using panel data. Nevertheless, as pointed out by Gabaix and Ioannides (2004) and Bosker et al. (2008), this methodology presents some drawbacks. First, the periodicity of our data is by decades, and we have only 22 temporal observations (decade-by-decade city sizes over a total period of 210 years), when the ideal would be to have at least annual data. Most studies use data from the decennial census, so this limitation is a common problem in the literature. Second, the presence of cross-sectional dependence across the cities in the panel can give rise to estimations that are not very robust. It has been well established in the literature that panel unit root and stationarity tests that do not explicitly allow for this feature among individuals present size distortions (Banerjee et al., 2005).

For this reason, we use one of the tests especially created to deal with this question; Pesaran's (2007) test for unit roots in heterogeneous panels with cross-section dependence is calculated based on the CADF statistic (cross-sectional ADF statistic, see below). To eliminate cross-dependence, the standard Dickey–Fuller (or augmented Dickey–Fuller (ADF)) regressions are augmented with the cross-section averages of lagged levels and first differences of the individual series, such that the influence of the unobservable common factor is asymptotically filtered.

The test of the unit root hypothesis is based on the t-ratio of the OLS estimate of b_i in the following cross-sectional augmented DF (CADF) regression:

$$\Delta y_{it} = a_i + b_i y_{i,t-1} + c_i \bar{y}_{t-1} + d_i \Delta \bar{y}_t + e_{it}, \tag{4}$$

where $y_{it} = \ln S_{it}$, a_i is the individual city-specific average growth rate and $\bar{y}_t$ is the cross-section mean of y_{it}, $\bar{y}_t = N^{-1} \sum_{j=1}^{N} y_{jt}$. The null hypothesis assumes that all series are non-stationary, and Pesaran's CADF is consistent under the alternative that only a fraction of the series is stationary.

Another advantage of Pesaran's CADF test over other recently developed unit root tests (Levin et al., 2002; Im et al., 2003) is that it is suitable for unbalanced panels, as is the case with our city sample. New cities appear over time, from 16 in 1790 to 190 in 2000.

However, due to limitations in the data (the CADF test works with unbalanced panels but if we consider the complete sample it is a strongly unbalanced panel; there is an excessive amount of missing data) we must restrict our analysis to a maximum of 150 cities. These 150 cities are a fixed sample for the entire 1790–2000 period, and correspond to the largest cities (upper-tail distribution) in 1900. This way we can control the possible bias mentioned in Section 2, as not all the largest cities of 1900 would maintain their position a century later. Therefore, the sample defined according to 1900 ranks contains "winning" and "losing" cities.[9]

Table 2 shows the results of the standardized Ztbar statistic of the CADF test, $Z[\bar{t}]$, and the corresponding p-value for three sample groups (top 75, 100 and 150 largest cities in 1900) and different specifications: $AR(p)$ with $p = 1,2,3$ including a constant or constant and trend.[10] The results are similar for the three sample sizes. When only one lag is included the null hypothesis of a unit root is rejected for any specification. However, as the number of lags in the model increases we soon find evidence in favour of our null hypothesis: in the model with two lags when a trend is included, and in the model with three lags with any specification. This last result is especially relevant, as Said and Dickey's (1984) $T^{1/3}$ rule would establish the lag choice $p = 3$ $\left(22^{1/3} = 2.8\right)$.

This evidence in favour of a unit root indicates that city growth during the 1790–2000 period was independent of the initial size, supporting our hypothesis of random growth. We carried out several robustness checks[11]. First, we defined the sample according to the largest cities in 2000, the latest period for which we have data. The results of the test, when it could be carried out,[12] were similar: with two lags or more, we could not reject the unit root for any specification of the model. We also tried defining the group of cities randomly, and again we obtained the result that the null hypothesis of a unit root could not be rejected (in this case the only model with which it could be rejected was with $p = 1$ and without trend). Finally, we estimated separately a panel for the sample of 16 cities that are present in all periods. In this case, as we considered a balanced panel we were also able to run the tests of Levin et al. (2002) and Im et al. (2003) (IPS test). The results for this group of the oldest cities were similar; we could

[9] Moreover, 1900 is when our sample exceeds 150 cities (see Table 1).
[10] The estimations were made with the pescadf Stata package, developed by Piotr Lewandowski.
[11] The specific values of the tests are available from the authors on request.
[12] In this case, due to data limitations, we could only carry out the test for the top 75 cities.

not reject the null hypothesis from two lags onward with any specification of the model and with any of the three tests.

4. What lies beneath the random growth: intra-distribution mobility

In the above section we found evidence supporting random growth against the alternative of mean reversion (convergence) in American cities during the 1790–2000 period. This type of growth pattern implies that cities evolve according to a stochastic process in which the growth rate does not depend on the initial size, so that the differences in the final size of the cities depend on exogenously distributed characteristics (locational fundamentals theory) or random shocks. In this case, the limit distribution of city size must converge to a Pareto distribution that obeys Zipf's law (Gabaix, 1999).

In this section we take a different perspective. Our intention is to examine mobility within the distribution, trying to extract growth patterns different from the general unit root trend detected in the previous section. To do this we use two different techniques. First, we calculate transition matrices, which tell us the degree of mobility in terms of probability. Second, we apply a cluster algorithm to identify different groups of cities that converge with each other. Both approaches are complementary; while the transition matrices define some groups in relative terms and the movements of cities between these groups are examined, with the second method we use the algorithm to identify endogenously the groups of cities that converge over time, looking for evidence of some type of "local" mean-reverting behaviour.

4.1 Transition matrices

Description

Eaton and Eckstein (1997) were the first to apply Quah's (1993) transition matrices to city size evolution. Let F_t be the vector representing the city size distribution at instant t, relative to the average size. We can say that this distribution follows a stochastic process defined by a Markov chain if the transition from one period to the next is given by:

$$F_{t+1} = M_t F_t \tag{5}$$

where M_t is the movement matrix or transition matrix, defining the law of movement from one period to the next. A Markov chain requires discrete time and a finite space of states E, which represents a discrete approximation to population distribution. Implicit in (5) is also what is known as the Markov property, i.e., that the future of the process depends only on its most immediate past (a homogeneous first-order stationary Markov process). Element p_{ijt} of the matrix M_t represents the probability that a city in state i in t moves to state j in t+1, i, $j \in E$. It is evident that $p_{ijt} \geq 0$ and that $\sum_{j \in E} p_{ijt} = 1, \forall i \in E$.

The elements of the matrix M_t can be estimated by maximum likelihood (see Hamilton (1994) and more recently Bosker et al. (2008)) applying:

$$\hat{p}_{ijt} = \frac{\sum_{t=1}^{T} n_{it,jt+1}}{\sum_{t=1}^{T} n_{it}}, \tag{6}$$

where $n_{it,jt+1}$ is the number of cities moving from state i in year t to state j in year $t+1$ and n_{it} the number of cities in state i in year t.

The general expression (5) is valid for the case in which no cities enter or leave the sample from one year to the next. This is not our case, and thus we need to deduce the correct equation, which describes the evolution of a distribution that allows cities to enter or leave (Lanaspa et al., 2011).

In the case of a sample that grows over time, in which from one period to the next cities only enter, Black and Henderson (2003) show that the correct equation is:

$$F_{t+1} = (1 - i_t) M_t F_t + i_t Z_t \tag{7}$$

where i_t is a scalar denoting the percentage of new cities in t+1 over the total existing cities in t+1 and Z_t is the vector of relative frequencies of the cities that enter.

In our case, where cities enter and leave the sample from one period to the next, let N be the number of cities, constant, in each period; let N_t be the number of cities entering or leaving from t to t+1; let $n_t = (N_t/N)$; let Z_t (X_t) be the vector of relative frequencies of the cities that enter (leave); finally, let M_t be the transition matrix from t to t+1 but only of the ($N-N_t$) cities that are in the sample both in t and in t+1.

Under these assumptions, the number of cities in the s-th state in t+1, $F_{t+1}^{S} N$, is:

$$F_{t+1}^{S} N = M_t^{S} (NF_t - N_t X_t) + N_t Z_t^{S}$$

where M_t^{S} is the s-th column of the matrix M_t. By definition the left-hand side of the above expression represents the number of cities in state S in the period t+1. Where they come from is shown on the right-hand side: the first term takes into account the cities that move to S among those that remain in the sample, discounting those that have left the sample; the second term shows those that move to S from among the new entrants. From the above equation we reach:

$$F_{t+1}^{S} = M_t^{S} F_t - n_t M_t^{S} X_t + n_t Z_t^{S} .$$

Generalizing for all the states, not only the s-th, we obtain the equation we were looking for:

$$F_{t+1} = M_t F_t - n_t M_t X_t + n_t Z_t . \qquad (8)$$

The difference between equation (8) and Black and Henderson's (2003) expression (equation 7) is the term $n_t M_t X_t$, which represents the distribution of cities that leave the sample.

Results

The three matrices M_T are given in Table 3, for the three sample sizes considered, 75, 100 and 150 cities. This methodology, with the theoretical extension introduced above, always takes into account the largest cities at each moment in time, allowing these largest cities to change, enter or leave the sample, or remain in it from one period to the next. Five states are considered; a larger number would increase the mobility artificially, and a smaller number would provide little information on intradistribution mobility. The upper limits for each state are: 0.4, 0.7, 1, 2 and ∞ times the average for each year.[13] The thresholds of the different categories are not exactly the same, but they are very similar to those used by Eaton and Eckstein (1997), Dobkins and Ioannides (2000) and Bosker et al. (2008), and in any case one of the criteria used to define them is that the number of cities in each of the categories should not be very different. As is already known, the major problem with this approach is that any choice of states inevitably involves a certain amount of arbitrariness. With this in mind, we have explored alternative cut-off points, although they are not very different from the states finally chosen, and the qualitative results remain the same. The relative frequencies are also shown of the cities that enter (Z_t) and leave the sample (X_t) throughout the period, as defined above.

Several conclusions emerge from Table 3. The first and most important is that we find intense mobility in the distribution of cities; persistence is not high. In fact, the elements

[13] The average is not calculated for all the cities, but for those that remain in the sample for two consecutive periods (see the definition of the matrix M_t).

of the diagonal of the matrices, which correspond to the cities that belong to the same state for two consecutive periods, are significantly different from one. Of the fifteen elements in the diagonals, only three are higher than 0.9, while six values are between 0.7 and 0.8, and one is below 0.7.

It is usual in the literature to find little mobility, as detected for the US by Black and Henderson (1999, 2003) and by Beeson et al. (2001), but there the samples cover a considerably smaller time horizon than the one we are considering. Our sample covers more than two centuries; by studying the urban structure from its beginning the conclusions may be different, as over these centuries, the late eighteenth, the nineteeth and the twentieth, the American urban structure was formed and built through demographic expansion (waves of immigration throughout the nineteeth century) and territorial expansion (the so-called conquest of the West and the founding of the cities of the West and Mid-West). Other works that consider the same time horizon (1790–2000) also find evidence of high mobility within the distribution (Batty, 2006; Cuberes, 2011). Thus, Batty (2006) develops rank-clocks that show how, with the exception of New York, the cities of the original 13 colonies gradually lost their positions with the entrance of new cities. Our data show the same behaviour, as a consequence of the mobility noted above and the entry of new cities. Table 4 shows ranks in 2000[14] corresponding to the cities that existed in the first period (1790); only New York and Philadelphia are still among the top 10 cities, while the rest have lost their positions and have been overtaken by other cities that entered the system later.

Cuberes (2011) finds that the average-rank of the fastest-growing cities (not just American cities, as his sample includes data for cities in other countries) tends to increase over time, a result that he interprets as evidence in favour of sequential urban growth. If cities grow sequentially, the cities that are initially the largest must represent a large share of the total urban population of the country in the initial periods and a relatively smaller one later on. As Table 1 shows, the behaviour of our sample of cities is consistent with this affirmation.

The second conclusion refers to the cities that enter and leave the sample. Those that leave the sample do so almost exclusively from the fifth state, that of the smallest cities. It makes sense that large cities do not disappear suddenly. In Cuberes (2009) and Henderson and Venables (2009) the explanation is that there is irreversible investment.

[14] Calculated using the entire city size distribution (19,296 incorporated places).

In Glaeser and Gyourko (2005) it happens because housing is a durable good that depreciates slowly over time. This fact is not the same for cities entering the sample, as they enter in all the states, except for that of the largest cities. This result indicates that cities enter the sample with a considerable size (most of them cities created in the West) and grow very quickly until they reach the size of existing cities (leapfrogging).

4.2 Convergence clubs

Description

The results in Section 3 show that we cannot reject the random growth (unit root) hypothesis for most of the proposed specifications, against the alternative hypothesis of convergence (mean reversion). However, in the previous section we find evidence of high mobility when we model growth as a first-order Markov process. Therefore, in this section we apply a cluster algorithm to try to identify different groups of cities that converge with each other, looking for evidence of some type of "local" mean-reverting behaviour.

The cluster procedure is based on the log $t-$test (Phillips and Sul, 2007, 2009), which focuses on how idiosyncratic transitions behave over time in relation to the common growth component. This new approach is different from that of previous empirical studies on growth convergence clubs, such as Durlauf and Johnson (1995) and Canova (2004). The regression model is:

$$\log\frac{H_1}{H_t}-2\log(\log t)=\beta_0+\beta_1\log t+u_t, \qquad \text{for } t=T_0,...,T \qquad (9)$$

where $\frac{H_1}{H_t}$ is the cross-sectional variance ratio, H_t is the transition distance, $H_t=N^{-1}\sum_{i=1}^{N}(h_{it}-1)^2$ and h_{it} is the relative transition coefficient, defined as $h_{it}=\frac{\log S_{it}}{N^{-1}\sum_{i=1}^{N}\log S_{it}}$. A relative transition coefficient eliminates the common growth component (μ_t) by scaling and measures the transition element for city i relative to the cross-section average. The variable h_{it} traces out an individual trajectory for each i relative to the average, so Phillips and Sul (2009) call h_{it} the "relative transition path". At the same time, h_{it} measures city i's relative departure from the common steady-state growth path μ_t. Equation (9) is obtained from a neoclassical growth model (see Phillips and Sul, 2007).

The test is based on a simple time series regression and involves a one-sided $t-$test of the null hypothesis of growth convergence against alternatives that include no convergence and partial convergence among subgroups. The test is called the '$\log t$' convergence test because the t-statistic refers to the coefficient of the $\log t$ regressor in the regression equation (9). We are interested not only in the sign of the coefficient β_1 of $\log t$ but also in its magnitude, which measures the speed of convergence. If $\beta_1 \geq 2$ and the common growth component μ_t follows a random walk with drift or a trend stationary process,[15] then values of β_1 that are this large will imply convergence in level city populations. However, if $2 > \beta_1 \geq 0$ this speed of convergence corresponds to conditional convergence, in which population growth rates converge over time across the cities within the club.[16]

The cluster procedure performs the $\log t$ test for each of the groups, and stops when the group of remaining cities does not satisfy the convergence test. First of all it defines an initial core primary group, and other groups are formed according to certain criteria that maximize the value of the t-statistic. A much more detailed explanation of the constructive steps of the procedure can be found in Phillips and Sul (2007, 2009).

Results

Table 5 shows the results of applying the cluster algorithm to our sample of cities.[17] Again, the results are reported for three sample sizes: the top 75, 100 and 150 largest cities in 1900.[18] In this case, the choice of the reference period can be relevant, as the largest cities in 2000 are a sample of "winning" cities, cities that over time and since

---

[15] Note that the hypothesis of random growth in the common growth component has been previously tested in Section 3.

[16] Note that this terminology is slightly different from the classical definition of conditional convergence, which depends on individuals' structural characteristics and initial conditions (Galor, 1996). An analysis of the general characteristics of the various convergence clubs as well as the many possible determining factors and initial conditions in each case is beyond the scope of this paper.

[17] The estimations were performed with the Gauss code kindly provided by Donggyu Sul on his web page. As Phillips and Sul (2007) recommend, we set $r = 0.3$ (r is the initiating sample fraction).

[18] To apply the algorithm we must have a balanced panel data. Given that most of the cities appear in the sample after 1790, we must carry out a little data transformation, assigning a population of 1 to the cities that did not exist in each period. This transformation means that these cities have a zero log-population in the periods in which they did not exist. If this change would have any effect on the cluster procedure the cities that appear in the same period would be grouped in the same club; however, Figure 1 shows how the groups are formed by cities that appear in different periods.

they first appeared have presented the highest growth rates.[19] However, some of the cities that were among the largest in 1900 have lost their positions in the ranking and have been overtaken by other cities, so that if we consider this sample of cities we capture more heterogeneous behaviours.[20]

The “club” column shows the number of cities that are members of each convergence group. The results are consistent for the three sample sizes, as despite enlarging the sample the cities do not usually change group. Only with the top 150 sample is there a small redistribution of cities, as one less convergence club is detected. The distribution of cities within groups can be consulted in the Appendix.

Given that the city distribution is fairly consistent regardless of the sample size, for clarity we will show only the graphs for the top 75. Figure 1 shows the evolution over time of the log-population of the cities in each convergence club. Our analysis will focus on these results. The first graph shows the evolution of the top 75 cities, and it is very difficult to deduce any specific type of pattern in it. However, some of the groups represented in the rest of the graphs present a clear sequential pattern, especially in the entry of new cities. These cities start with a zero log-population, but grow at a faster rate than the rest of the cities in their club until they reach similar growth rates.[21] This behaviour is consistent with a pattern of sequential city growth, at least within groups.

The algorithm classifies cities into 12 groups, all of them convergence clubs because $\beta_1 > 0$. There are four remaining cities that are not classified within any club and for which the convergence hypothesis is rejected. In each group the coefficient is significantly positive, revealing strong empirical support for the club classification. Also, only one of the estimated coefficients is significantly greater than 2 (club 2), indicating that there is strong evidence of conditional convergence within each of these clubs, while the evidence in favour of level convergence is much smaller. Of the four cities belonging to club 2, three are in the South Region, although the geographical distribution of cities does not show any specific spatial pattern in any of the groups. Only club 11 consists of cities belonging to the same region (Northeast), although

---

[19] In fact, with the largest cities in 2000 we find only 4 convergence clubs, as all of them are cities characterized by high growth rates. The results are available from the authors on request.

[20] In the top 150 cities in 2000 there are 31 cities (20.67%) that are not in the top 150 cities in 1900. The differences are still greater in the top 75 and 100, as there are 36 different cities that represent 48% and 36% of the sample, respectively.

[21] Some of the graphs are very similar to Figure 4 (a) in Henderson and Venables (2009), obtained by simulations of their theoretical model of city formation.

another common characteristic of these cities is that they are among the oldest. The cities that have existed since 1790 are classified into groups 10 to 12, indicating that while they present a different growth pattern from the cities that appeared later, they also differ from each other.

It should be noted that of the 12 clubs, only clubs 1 and 2 correspond to cities that rise in the ranking (on average) from 1900 to 2000. The cities within the other clubs lose positions in the ranking (on average), especially the cities in clubs 7, 9 and 12, confirming our idea that our sample captures more heterogeneous behaviours than the sample of "winning" cities in 2000, as we also include "failing" cities that performed poorly in terms of growth over the entire time interval.

5. Discussion

In the sections above we have found mixed evidence regarding city growth in the long term. First, we cannot reject the random growth (unit root) hypothesis for most of the proposed specifications, against the alternative hypothesis of convergence (mean reversion). However, we find evidence of high mobility when we model growth as a first-order Markov process; this mobility is consistent with the results of other studies that consider the same 1790–2000 period (Batty, 2006; Cuberes, 2011). Finally, using a cluster procedure we find strong evidence supporting conditional convergence in city growth rates within convergence clubs, which we can interpret as "local" mean-reverting behaviours. Both the high mobility and the results of clustering analysis seem to indicate a sequential city growth pattern.

These results raise two questions: first, whether these different empirical results are compatible and second, whether the city size distribution has evolved according to the random growth pattern (if Zipf's law holds) or whether, on the contrary, the trend has been convergence among cities.

The first question asks whether a random growth result is compatible with a degree of convergence in the evolution of city growth rates; in other words, whether a unit root is compatible with some kind of mean-reverting component. Gabaix and Ioannides (2004) answer this question by putting forward what they call *"deviations from Gibrat's Law* (random growth) *that do not affect the distribution"*, starting from

$$\ln S_{it} - \ln S_{it-1} = \mu(X_{it}, t) + \varepsilon_{it}, \tag{10}$$

where X_{it} is a possibly time-varying vector of characteristics of city i; $\mu(X_{it}, t)$ is the expectation of city i's growth rate as a function of economic conditions at time t; and ε_{it} is white noise. In the simplest random growth model, ε_{it} is independently and identically distributed over time (this means that ε_{it} has a zero mean and a constant variance that is uncorrelated with ε_{is} for $t \neq s$) and $\mu(X_{it}, t)$ is constant.

Gabaix and Ioannides (2004) consider two types of deviations, relaxing both assumptions. Rossi-Hansberg and Wright (2007) discuss economic interpretations of deviations from Zipf's and Gibrat's laws. We are interested in the consequences of relaxing the assumption of an i.i.d. ε_{it}, assuming constant $\mu(X_{it}, t) = \mu$. In its place the following stochastic structure is assumed: $\varepsilon_{it} = b_{it} + \eta_{it} - \eta_{it-1}$, where b_{it} is i.i.d. and η_{it} follows a stationary process. Replacing in (10) we obtain:

$$\ln S_{it} - \ln S_{i0} = \mu t + \sum_{s=1}^{t} b_{is} + \eta_{it} - \eta_{i0}. \tag{11}$$

The term $\sum_{s=1}^{t} b_{is}$ gives a unit root in the growth process, while the term η_{it} can have any stationarity. According to Gabaix and Ioannides (2004), this means that *"for Zipf's law to hold, the city evolution process can contain a mean reversion component, as long as it contains a non-zero unit root component."* Therefore, our mixed empirical evidence is not contradictory, but compatible. Also, our conclusion leads us directly to our second question, the behaviour of city size distribution over the 1790–2000 period (and whether Zipf's law holds).

Let us denote S as the size and R as its corresponding rank (1 for the largest, 2 for the second largest and so on). A power law (Pareto distribution) links city size and rank as follows: $R(S) = AS^{-a}$. This expression has been used extensively in urban economics to study city size distribution (see, for example, Eeckhout (2004) and Ioannides and Overman (2003) for the US case). It is usually specified and estimated in its logarithmic version:

$$\ln R = b - a \ln S + \xi, \tag{12}$$

where ξ is the error term and b and a are the parameters that characterize the distribution. The latter is known as the Pareto exponent, and Zipf's law is considered to hold when $a = 1$. This means that when ordered from largest to smallest, the size of the second city is half that of the first one, the size of the third is a third of the first one and so on. The greater the coefficient, the more homogeneous are the city sizes. Also, an increase of the coefficient over time would mean a process of convergence in city sizes. Similarly, the smaller the coefficient, the less homogeneous are city sizes, and a decreasing evolution would mean a process of divergence.

Gabaix and Ibragimov (2011) proposed specifying equation (12) subtracting $1/2$ to the rank to obtain an unbiased estimation of a:

$$\ln\left(R-\frac{1}{2}\right)=b-a\ln S+\varepsilon\,. \tag{13}$$

Equation (13) was estimated by OLS for our sample of cities in the different decades during the 1790–2000 period. Figure 2 shows the results. We estimated using all the cities available in each decade (from 16 in 1790 to 190 in 2000). The results show that the distribution remained stable until 1950, so the entry of new cities did not have significant effects, although the estimated coefficients are less than one, indicating a high degree of inequality among city sizes. Therefore, during this period the stable evolution of the city size distribution reflects the random growth process, even though the resulting Pareto exponent of the distribution is lower than one, rejecting Zipf's law for this group of the largest cities.[22]

From 1950 the estimated Pareto coefficient grows to reach (and exceed) the value of one. Note that from 1950 to 2000 only 11 cities enter the sample, so that the evolution of the exponent reacts only to the city growth process. The increasing trend of the exponents indicates a process of convergence among cities. We also estimated the Gini coefficients for each period. The Gini coefficients have the advantage of not imposing a specific size distribution (Pareto for rank-size coefficients). The results are similar; from 1790 to 1950 the Gini coefficient rose from 0.65 to 0.68,[23] while in the year 2000 it was 0.50. Therefore, during this period the evolution of the distribution clearly corresponds to a convergence phase. The explanation for this convergence process is well known in the literature (post-war suburbanization). During the second half of the twentieth century mid-sized and small American cities grew much more than the largest cities in

[22] Except in 1830 and 1840, for which the confidence intervals indicate that we cannot reject that the coefficient is significantly different from 1.
[23] However, the evolution of the Gini coefficient is not as stable as that of the Pareto exponent, as within this period it does reflect changes in the inequality of the distribution in some decades.

the same metropolitan area.[24] Glaeser et al. (2011) claim that some of the impact of sprawl and the role that the automobile played in dispersing the American population can explain some of these patterns. The effect that we capture from this process is that the cities of the upper-tail distribution became more homogeneous in size, due to the larger growth of mid-sized cities bringing them closer to the largest ones.

6. Conclusions

In this paper we study the growth pattern of the system of cities in the United States from its origin. We obtain several conclusions. First, we find evidence supporting random growth in American cities during the 1790–2000 period, indicating that the growth rate does not depend on initial size. Second, we find evidence of high intra-distribution mobility when we consider growth as a first-order Markov process. Third, using a cluster procedure we find evidence in favour of the conditional convergence of city growth rates within convergence clubs, allowing us to conclude that “local” mean-reverting behaviours exist. Our results lend support to recent theories of sequential city growth.

Acknowledgements

The authors acknowledge financial support from the Spanish Ministerio de Educación y Ciencia (ECO2009-09332 and ECO2010-16934 projects), the DGA (ADETRE research group) and FEDER. We have benefited from the helpful comments of David Cuberes. Suggestions from Fernando Sanz also contributed to improving the paper. Comments made by seminar participants at the IEB seminar (Barcelona, 2011) were also highly appreciated. All remaining errors are ours.

References

[1] Banerjee, A., M. Massimiliano, and C. Osbat, (2005). Testing for PPP: should we use panel methods? Empirical Economics, 30: 77–91.

[2] Batty, M., (2006). Rank clocks. Nature, Vol. 444, 30 November 2006, 592-596.

[3] Beeson, P.E., D. N. DeJong, and W. Troesken, (2001). Population Growth in US Counties, 1840-1990. Regional Science and Urban Economics, 31: 669-699.

---

[24] There are several works studying the causes of this process. For example, Margo (1992) examines the role of incomes.

[4] Black, D., and V. Henderson, (1999). Spatial Evolution of Population and Industry in the United States. The American Economic Review, Vol. 89(2), Papers and Proceedings of the One Hundred Eleventh Annual Meeting of the American Economic Association (May, 1999), 321-327.

[5] Black, D., and V. Henderson, (2003). Urban evolution in the USA. Journal of Economic Geography, Vol. 3(4): 343-372.

[6] Bosker, E. M., S. Brakman, H. Garretsen and M. Schramm, (2008). A century of shocks: the evolution of the German city size distribution 1925 – 1999. Regional Science and Urban Economics 38: 330–347.

[7] Bridenbaugh, C., (1938). Cities in the Wilderness. The Ronald Press.

[8] Canova, F., (2004). Testing for convergence clubs in income per capita: a predictive density approach. International Economic Review, 45: 49–77.

[10] Champernowne, D., (1953). A model of income distribution. Economic Journal, LXIII: 318-351.

[11] Cheshire, P. C., and S. Magrini, (2006). Population Growth in European Cities: Weather Matters- but only Nationally. Regional Studies, 40(1): 23-37.

[12] Clark, J. S., and J. C. Stabler, (1991). Gibrat's Law and the Growth of Canadian Cities. Urban Studies, 28(4): 635-639.

[13] Córdoba, J. C., (2008). A generalized Gibrat's law. International Economic Review, Vol. 49(4): 1463-1468.

[14] Cuberes, D., (2009). A Model of Sequential City Growth. The B.E. Journal of Macroeconomics: Vol. 9: Iss. 1 (Contributions), Article 18.

[15] Cuberes, D., (2011). Sequential City Growth: Empirical Evidence. Journal of Urban Economics, 69: 229–239.

[16] Davis, D. R., and D. E. Weinstein, (2002). Bones, bombs, and break points: the geography of economic activity. American Economic Review, 92(5): 1269–1289.

[17] Dobkins, L. H., and Y. M. Ioannides, (2000). Dynamic evolution of the US city size distribution. Included in Huriot, J. M. and J. F. Thisse (Eds.), The economics of cities. Cambridge University Press, Cambridge, pp. 217-260.

[18] Dobkins, L. H., and Y. M. Ioannides, (2001). Spatial interactions among U.S. cities: 1900–1990. Regional Science and Urban Economics 31: 701–731.

[19] Duranton, G., (2007). Urban Evolutions: The Fast, the Slow, and the Still. American Economic Review, 97(1): 197-221.

[20] Durlauf, S. N., and P. A. Johnson, (1995). Multiple regimes and cross-country growth behavior. Journal of Applied Econometrics, 10: 365–384.

[21] Eaton, J., and Z. Eckstein, (1997). Cities and Growth: Theory and Evidence from France and Japan. Regional Science and Urban Economics, 1997, 27(4 –5), 443–474.

[22] Eeckhout, J., (2004). Gibrat's Law for (All) Cities. American Economic Review, 94(5): 1429–1451.

[23] Eeckhout, J., (2009). Gibrat’s Law for (all) Cities: Reply. American Economic Review, 99(4): 1676–1683.

[24] Ellison, G., and E. L. Glaeser, (1999). The geographic concentration of industry: Does natural advantage explain agglomeration? American Economic, Review Papers and Proceedings, 89(2), 311-316.

[25] Gabaix, X., (1999). Zipf’s law for cities: An explanation. Quarterly Journal of Economics, 114(3):739-767.

[26] Gabaix, X., and R. Ibragimov, (2011). Rank-1/2: A simple way to improve the OLS estimation of tail exponents. Journal of Business & Economic Statistics, 29(1): 24-39.

[27] Gabaix, X., and Y. M. Ioannides, (2004). The evolution of city size distributions. Handbook of urban and regional economics, Vol. 4, J. V. Henderson and J. F. Thisse, eds. Amsterdam: Elsevier Science, North-Holland, 2341-2378.

[28] Galor, O., (1996). Convergence? Inferences from Theoretical Models. The Economic Journal, Vol. 106(437): 1056-1069.

[29] Garmestani, A. S., C. R. Allen, and C. M. Gallagher, (2008). Power laws, discontinuities and regional city size distributions. Journal of Economic Behavior & Organization, 68: 209–216.

[30] Giesen, K., A. Zimmermann, and J. Suedekum, (2010). The size distribution across all cities – double Pareto lognormal strikes. Journal of Urban Economics, 68: 129–137.

[31] Glaeser, E. L., and J. Gyourko, (2005). Urban Decline and Durable Housing. Journal of Political Economy, 113(2): 345-375.

[32] Glaeser, E. L., G. A. M. Ponzetto, and K. Tobio, (2011). Cities, Skills, and Regional Change. Forthcoming in Regional Studies.

[33] González-Val, R., (2010). The Evolution of the US City Size Distribution from a Long-run Perspective (1900–2000). Journal of Regional Science, 50(5): 952–972.

[34] Hamilton, J. D., (1994). Time Series Analysis. Princeton, NJ: Princeton University Press.

[35] Henderson, J. V., and A. Venables, (2009). The Dynamics of City Formation. Review of Economic Dynamics, 12: 233–254.

[36] Henderson, J. V., and H. G. Wang, (2007). Urbanization and city growth: The role of institutions. Regional Science and Urban Economics, 37(3): 283–313.

[37] Im, K. S., M. H. Pesaran, and Y. Shin, (2003). Testing for Unit Roots in Heterogeneous Panels. Journal of Econometrics, 115: 53–74.

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

[39] Kim, S., (2000). Urban development in the United States. Southern Economic Journal 66, 855-880.

[40] Kim, S., and R. A. Margo, (2004). Historical perspectives on U.S. Economic Geography. Handbook of urban and regional economics, vol. 4, J. V. Henderson and J. F. Thisse, eds. Amsterdam: Elsevier Science, North-Holland, Chapter 66, pp. 2982-3019.

[41] Lanaspa-Santolaria, L. F., A. Montañes, L. I. Olloqui-Cuartero, and F. Sanz-Gracia, (2002). The Phenomenon of Regional Inversion in the US manufacturing sector. Papers in Regional Science, 81 (4): 461-482

[42] Lanaspa, L. F., F. Pueyo, and F. Sanz, (2011). Urban dynamics during the twentieth century. A tale of five European countries. Mimeo, Universidad de Zaragoza.

[43] Levin, A., C.-F. Lin, and C.-S. J. Chu, (2002). Unit Root Tests in Panel Data: Asymptotic and Finite Sample Properties. Journal of Econometrics, 108: 1–24.

[44] Levy, M., (2009). Gibrat's Law for (all) Cities: A Comment. American Economic Review, 99(4): 1672–1675.

[45] Margo, R. A., (1992). Explaining the Postwar Suburbanization of Population in the United States: The Role of income. Journal of Urban Economics, 31: 301-310.

[46] Melo, P. C., D. J. Graham, and R. B. Noland, (2009). A Meta-analysis of Estimates of Urban Agglomeration Economies. Regional Science and Urban Economics, 39: 332-342.

[47] Michaels, G., F. Rauch, and S. J. Redding, (2010). Urbanization and Structural Transformation. Unpublished manuscript, London School of Economics.

[48] Overman, H. G., and Y. M. Ioannides, (2001). Cross-Sectional Evolution of the U.S. City Size Distribution. Journal of Urban Economics 49, 543-566.

[49] Quah, D., (1993). Empirical cross-section dynamics in economic growth. European Economic Review, 31: 426-434.

[50] Pesaran, M. H., (2007). A simple panel unit root test in the presence of cross-section dependence. Journal of Applied Econometrics, 22: 265–312.

[51] Phillips, P. C. B., and D. Sul, (2007). Transition Modeling and Econometric Convergence Tests. Econometrica, Vol. 75, 1771-1855.

[52] Phillips, P. C. B., and D. Sul, (2009). Economic Transition and Growth. Journal of Applied Econometrics, 24, 1153-1185.

[53] Resende, M., (2004). Gibrat's Law and the Growth of Cities in Brazil: A Panel Data Investigation. Urban Studies, Vol. 41(8): 1537-1549.

[54] Rossi-Hansberg, E., and M. L. J. Wright, (2007). Urban structure and growth. Review of Economic Studies, 74: 597–624.

[55] Rozenfeld, H. D., D. Rybski, X. Gabaix, and H. A. Makse, (2011). The Area and Population of Cities: New Insights from a Different Perspective on Cities. American Economic Review, forthcoming.

[56] Said, S. E., and D. A. Dickey, (1984). Testing for Unit Roots in Autoregressive-Moving Average Models of Unknown Order. Biometrika, 71(3): 599-607.

[57] Sharma, S., (2003). Persistence and Stability in City Growth. Journal of Urban Economics 53, 300-320.

[58] Simon, H., (1955). On a class of skew distribution functions. Biometrika, 42: 425-440.

[59] Soo, K.T., (2007) Zipf's Law and Urban Growth in Malaysia. Urban Studies 44(1):1-14

[60] U.S. Census Bureau, (2004). 2000 Census of Population and Housing, Population and Housing Unit Counts PHC-3-1, United States Summary. Washington, DC. Available at: http://www.census.gov/prod/cen2000/phc3-us-pt1.pdf

[61] Vining, D. R., (1976). Autocorrelated Growth Rates and the Pareto Law: A Further Analysis. The Journal of Political Economy, 84(2): 369-380.

Table 1. Number of Cities and Descriptive Statistics by Year

Year	Cities	Mean	Standard deviation	Minimum	Maximum	US urban population (UP)	Percentage of UP in our sample
1790	16	8,746.50	13,313.13	200	49,401	201,655	69.40%
1800	22	10,255.00	18,565.84	81	79,216	322,371	69.98%
1810	25	14,278.04	26,052.55	383	119,734	525,459	67.93%
1820	28	16,832.07	31,499.38	606	152,056	693,255	67.98%
1830	36	20,631.19	43,079.73	877	242,278	1,127,247	65.89%
1840	50	24,502.46	58,753.40	1,222	391,114	1,845,055	66.40%
1850	73	30,220.67	85,663.40	415	696,115	3,574,496	61.72%
1860	94	44,193.24	136,697.40	175	1,174,779	6,216,518	66.82%
1870	110	55,417.75	160,729.66	155	1,478,103	9,902,361	61.56%
1880	125	65,037.17	197,482.93	556	1,911,698	14,129,735	57.54%
1890	149	77,799.07	232,080.75	273	2,507,414	22,106,265	52.44%
1900	157	108,432.39	329,863.51	202	3,437,202	30,214,832	56.34%
1910	165	142,935.56	433,335.63	297	4,766,883	42,064,001	56.07%
1920	171	176,340.04	509,938.16	326	5,620,048	54,253,282	55.58%
1930	179	211,572.36	614,701.55	515	6,930,446	69,160,599	54.76%
1940	179	224,762.88	651,013.99	582	7,454,995	74,705,338	53.85%
1950	179	260,994.59	695,986.21	727	7,891,957	96,846,817	48.24%
1960	182	290,794.10	683,649.24	3,695	7,781,984	125,268,750	42.25%
1970	187	308,875.27	679,828.20	14,089	7,895,563	149,646,617	38.60%
1980	188	311,706.85	617,176.35	62,134	7,071,639	167,050,992	35.08%
1990	190	332,701.32	635,704.55	95,802	7,322,564	187,053,487	33.79%
2000	190	364,890.56	690,433.95	100,565	8,008,278	222,360,539	31.18%

Note: US urban population data are taken from the US Census Bureau. Source: http://www.census.gov/population/censusdata/table-4.pdf.

Table 2. Panel unit root tests, 1790–2000. Pesaran's CADF statistic

Model	Sample size		
	Top 75	Top 100	Top 150
Augmenting lag (1)			
Constant	-8.039 (0.000)	-5.548 (0.000)	-8.139 (0.000)
Constant & trend	-7.855 (0.000)	-5.711 (0.000)	-1.922 (0.027)
Augmenting lags (2)			
Constant	-3.416 (0.000)	-0.290 (0.386)	-3.120 (0.001)
Constant & trend	-0.706 (0.240)	1.529 (0.937)	10.135 (1.000)
Augmenting lags (3)			
Constant	4.384 (1.000)	11.569 (1.000)	21.564 (1.000)
Constant & trend	12.917 (1.000)	20.465 (1.000)	29.447 (1.000)

Note: test-statistic (p-value). Top cities according to ranks in 1900.

Table 3. Average 10-year transition matrices, 1790–2000

Sample Size: 75					
	∞	2	1	0.7	0.4
∞	0.928	0.072	0	0	0
2	0.050	0.820	0.129	0	0
1	0	0.162	0.676	0.162	0
0.7	0	0.006	0.077	0.792	0.125
0.4	0	0.002	0.012	0.107	0.880
X_t	0	0	0	0.00073	0.06506
Z_t	0	0.00073	0.00073	0.00512	0.10234

Sample Size: 100					
	∞	2	1	0.7	0.4
∞	0.915	0.085	0	0	0
2	0.071	0.820	0.104	0.005	0
1	0	0.114	0.710	0.176	0
0.7	0	0.018	0.095	0.742	0.145
0.4	0	0.001	0.007	0.105	0.887
X_t	0	0	0	0	0.049711
Z_t	0	0.000578	0.000578	0.003468	0.093642

Sample Size: 150					
	∞	2	1	0.7	0.4
∞	0.908	0.092	0	0	0
2	0.085	0.797	0.118	0	0
1	0.005	0.123	0.731	0.142	0
0.7	0	0.023	0.085	0.771	0.120
0.4	0	0.001	0.003	0.109	0.887
X_t	0	0	0	0	0.026083
Z_t	0	0.000421	0.000421	0.003786	0.077829

Table 4. Ranks in 1790 and 2000

City	Rank in 1790	Rank in 2000
New York	1	1
Philadelphia	2	5
Boston	3	20
Baltimore	4	17
Providence	5	123
New Haven	6	184
Richmond	7	96
Norfolk	8	74
Alexandria	9	174
Hartford	10	188
Cambridge	11	243
Worcester	12	125
Springfield	13	138
Lexington-Fayette	14	65
Manchester	15	224
Louisville	16	67

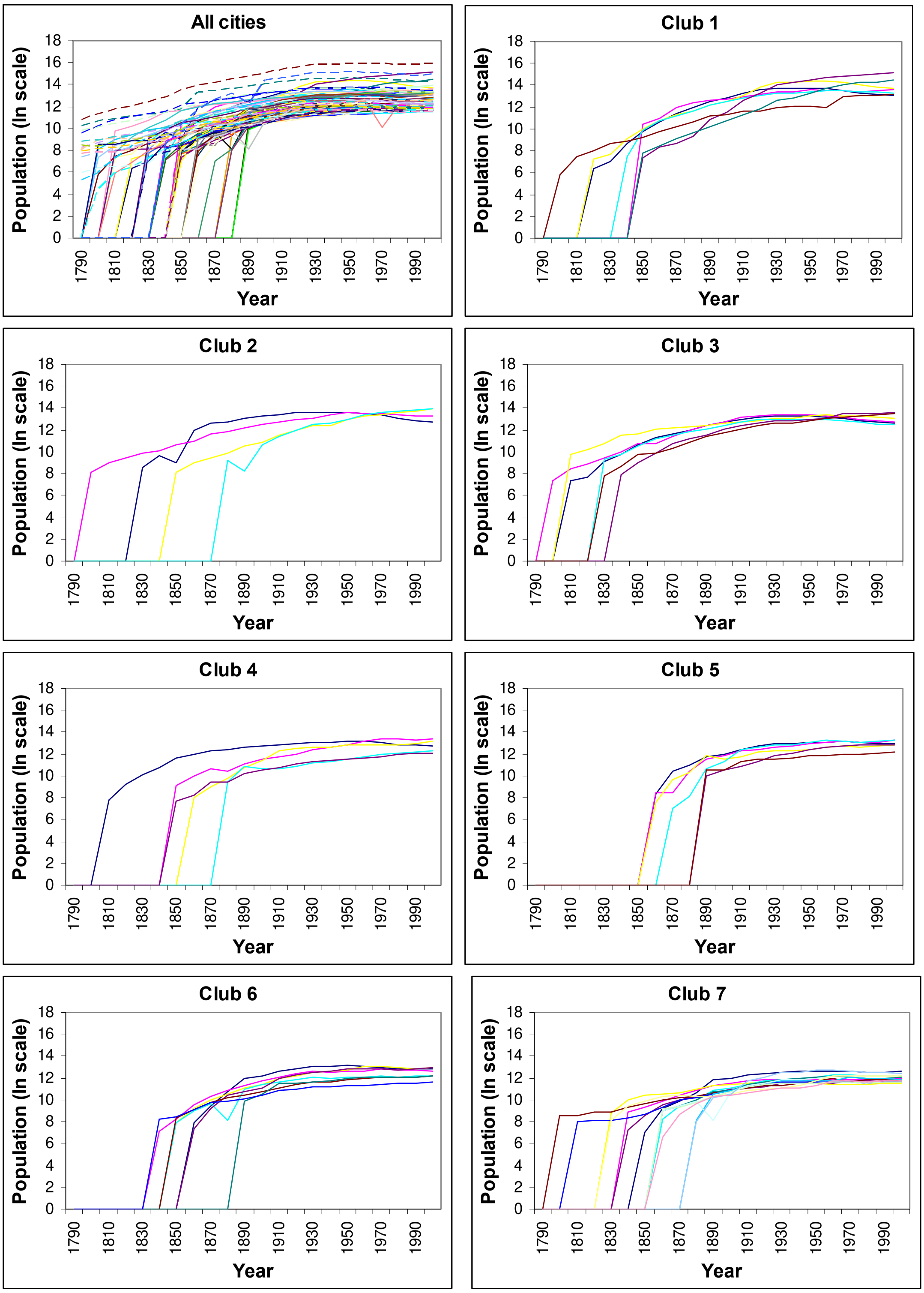
Note: Ranks in 2000 calculated using data from all incorporated places (19,296).

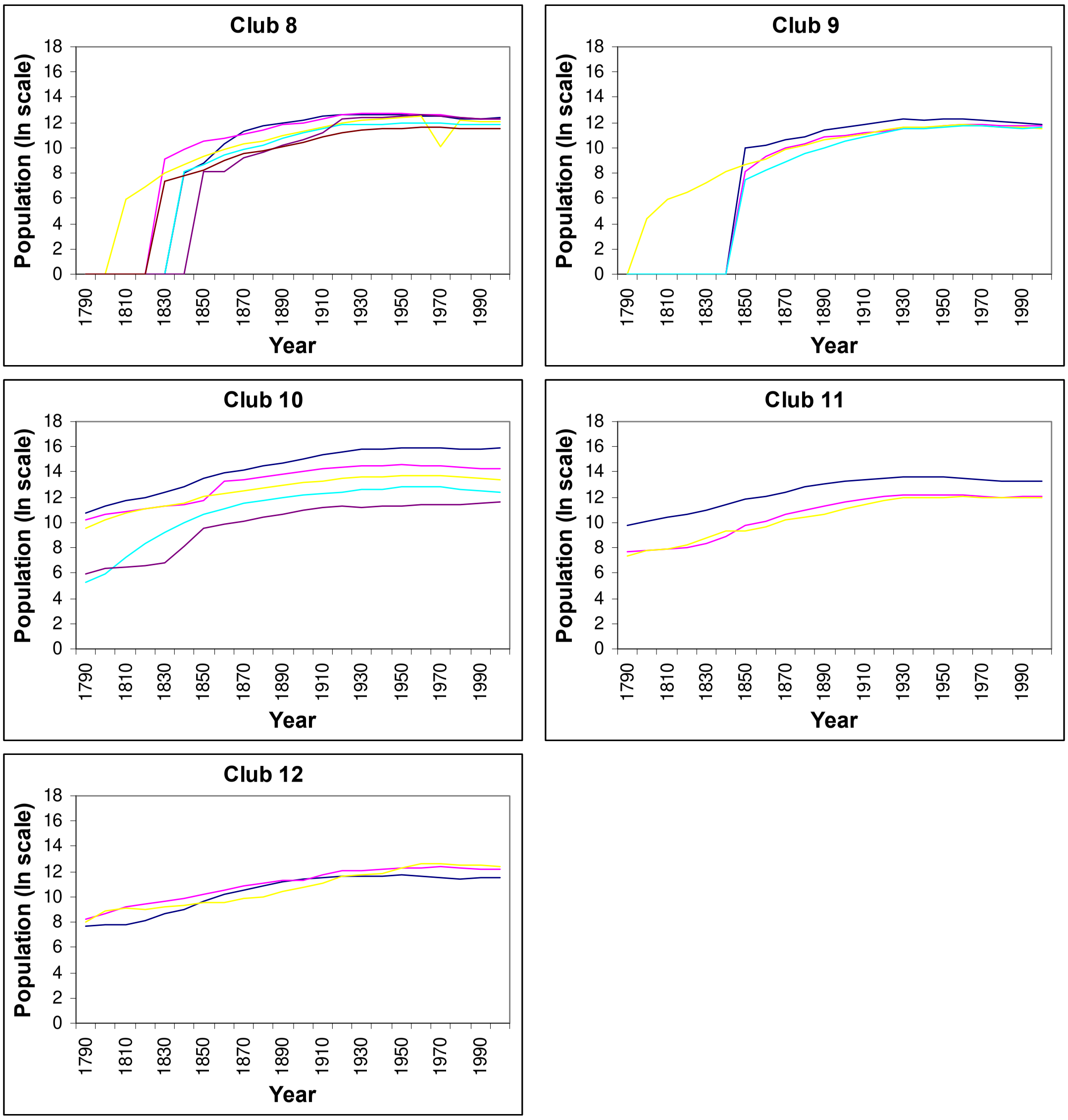
Table 5. Convergence clubs, 1790–2000

Club	β_1 (t-statistic)	Club	β_1 (t-statistic)	Club	β_1 (t-statistic)
1 [7]	0.105 (0.146)	1 [12]	0.744 (2.386)	1 [26]	1.217 (6.979)
2 [4]	2.507 (3.844)	2 [7]	0.671 (4.686)	2 [17]	0.254 (3.720)
3 [6]	0.893 (2.326)	3 [6]	0.893 (2.326)	3 [9]	0.225 (2.674)
4 [5]	0.256 (3.225)	4 [7]	0.142 (0.910)	4 [15]	0.141 (1.634)
5 [6]	0.294 (1.885)	5 [12]	0.560 (2.119)	5 [20]	0.400 (1.462)
6 [8]	0.435 (5.784)	6 [12]	0.010 (0.087)	6 [23]	0.064 (0.502)
7 [14]	0.224 (2.389)	7 [18]	0.370 (4.367)	7 [21]	0.539 (4.215)
8 [6]	1.970 (1.188)	8 [6]	1.970 (1.188)	8 [3]	2.405 (2.303)
9 [4]	0.353 (0.985)	9 [5]	0.700 (2.794)	9 [6]	0.011 (0.396)
10 [5]	0.224 (4.673)	10 [5]	0.224 (4.673)	10 [3]	0.842 (6.385)
11 [3]	0.842 (6.385)	11 [3]	0.842 (6.385)	11 [3]	0.347 (0.711)
12 [3]	0.347 (0.711)	12 [3]	0.347 (0.711)	Sample Size: Top 150	
Sample Size: Top 75		Sample Size: Top 100			

Notes: The numbers in brackets are the number of cities. Top cities are defined according to the ranks in 1900. The corresponding t-statistic in the regression is constructed in the usual way using HAC standard errors. At the 5% level, for example, the null hypothesis of convergence is rejected if the t-statistic < -1.65. All of the t-statistics reported are positive, indicating that we cannot reject the null at 5% in any case.

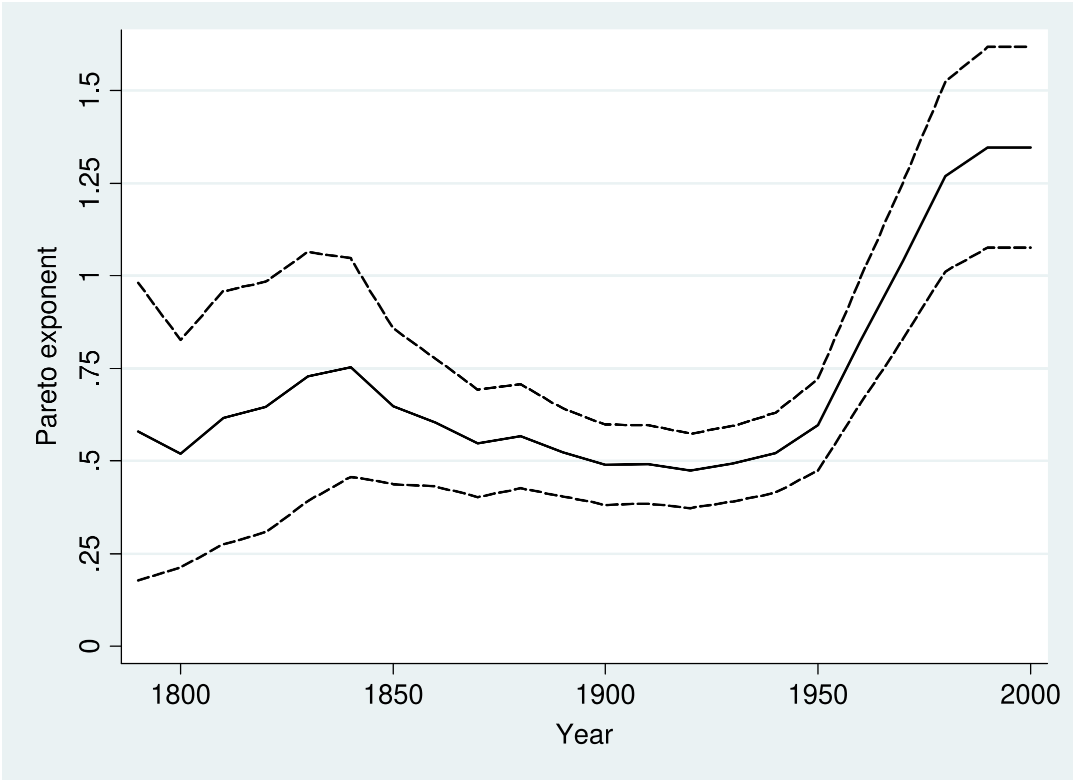
Figure 1. Cities' log population evolution, Top 75, 1790–2000





Note: Top 75 according to the ranks in 1900.

Figure 2. Evolution of the estimated Pareto exponents



Notes: The Pareto exponent is estimated using Gabaix and Ibragimov's Rank-$1/2$ estimator. Dashed lines represent the standard errors calculated applying Gabaix and Ioannides's (2004) corrected standard errors: $\mathrm{GI\,s.e.} = \hat{a} \cdot (2/N)^{1/2}$, where N is the sample size.

Appendix: Cities within clubs

Rank in 1900	Name	Club (Sample Size: Top 75)	Club (Sample Size: Top 100)	Club (Sample Size: Top 150)
1	New York	10	10	9
2	Chicago			
3	Philadelphia	10	10	9
4	St. Louis	2	2	2
5	Boston	11	11	10
6	Baltimore	10	10	9
7	Cleveland	1	1	1
8	Buffalo	3	3	3
9	San Francisco	1	1	1
10	Cincinnati	4	4	4
11	Pittsburgh	3	3	3
12	New Orleans	3	3	3
13	Detroit	1	1	1
14	Milwaukee	1	1	1
15	Washington	2	2	2
16	Newark	3	3	3
17	Jersey	8	8	7
18	Louisville	10	10	9
19	Minneapolis	6	6	6
20	Providence			
21	Indianapolis	3	3	3
22	Kansas	5	5	5
23	St. Paul	7	7	6
24	Rochester	8	8	7
25	Denver	5	5	5
26	Toledo	6	6	6
27	Columbus	3	3	3
28	Worcester	11	11	10
29	Syracuse	9	9	7
30	New Haven			
31	Paterson	7	7	7
32	Omaha	5	5	5
33	Los Angeles	1	1	1
34	Memphis	4	4	4
35	Lowell	7	7	7
36	Cambridge	12	12	11
37	Portland	4	4	4
38	Atlanta	6	6	6
39	Grand Rapids	6	6	6
40	Dayton	8	8	6
41	Richmond	12	12	11
42	Nashville-Davidson	1	1	1
43	Seattle	5	5	5
44	Hartford			
45	Bridgeport	8	8	7
46	Oakland	6	6	6
47	Des Moines	7	7	7
48	Springfield	11	11	10
49	Evansville	9	9	7
50	Manchester	10	10	9

51	Peoria	7	7	7
52	Savannah	7	7	7
53	Salt Lake	7	7	7
54	San Antonio	2	2	2
55	Erie	9	9	8
56	Elizabeth	7	7	7
57	Kansas	7	7	6
58	Yonkers	7	7	6
59	Norfolk	12	12	11
60	Waterbury	7	7	7
61	Fort Wayne	6	6	6
62	Houston	1	1	1
63	Akron	8	8	7
64	Dallas	2	2	2
65	Lincoln	4	4	4
66	Honolulu CDP	5	5	5
67	Mobile	7	7	6
68	Birmingham	7	7	7
69	Little Rock	4	4	4
70	Tacoma	5	5	5
71	Spokane	6	6	6
72	South Bend	9	9	8
73	Allentown	8	8	7
74	Springfield	6	6	6
75	Topeka	7	7	7
76	Knoxville		5	5
77	Rockford		7	6
78	Montgomery		4	4
79	Chattanooga		6	6
80	Sacramento		2	2
81	Jacksonville		1	1
82	Fort Worth		4	4
83	Cedar Rapids		6	6
84	Lexington-Fayette		6	6
85	Wichita		5	5
86	Springfield		5	5
87	Austin		1	1
88	San Jose		1	1
89	Colorado Springs		1	1
90	Waco		7	6
91	Newport News		5	5
92	Madison		5	5
93	Charlotte		2	2
94	San Diego		1	1
95	Columbus		5	5
96	Stockton		2	2
97	Portsmouth		9	7
98	Lansing		7	7
99	Shreveport		6	6
100	Stamford		7	6
101	El Paso			2
102	Tampa			6
103	Alexandria			9
104	Ann Arbor			5

105	Winston-Salem	5
106	Raleigh	2
107	Laredo	2
108	Berkeley	7
109	Flint	8
110	Fresno	1
111	Baton Rouge	4
112	Oklahoma	4
113	Greensboro	4
114	Beaumont	7
115	Pasadena	6
116	Huntsville	3
117	Riverside	1
118	Vallejo	4
119	Jackson	5
120	Tucson	2
121	Independence	5
122	Durham	2
123	Santa Rosa	1
124	Albuquerque	2
125	San Bernardino	3
126	Boise City	1
127	Phoenix	1
128	Pomona	3
129	Santa Ana	1
130	Bakersfield	1
131	Corpus Christi	4
132	Reno	1
133	Salem	2
134	Abilene	6
135	Salinas	1
136	Eugene	2
137	Tallahassee	2
138	Hampton	5
139	Orlando	2
140	Long Beach	4
141	Modesto	1
142	Hayward	4
143	Miami	5
144	St. Petersburg	5
145	Anaheim	1
146	Amarillo	5
147	Tulsa	4
148	Plano	1
149	Orange	1
150	Arlington	1