

Sánchez-Vidal, María; González-Val, Rafael; Viladecans-Marsal, Elisabet

**Conference Paper**

## Sequential city growth in the US: Does age matter?

53rd Congress of the European Regional Science Association: "Regional Integration: Europe, the Mediterranean and the World Economy", 27-31 August 2013, Palermo, Italy

**Provided in Cooperation with:**

European Regional Science Association (ERSA)

*Suggested Citation:* Sánchez-Vidal, María; González-Val, Rafael; Viladecans-Marsal, Elisabet (2013) : Sequential city growth in the US: Does age matter?, 53rd Congress of the European Regional Science Association: "Regional Integration: Europe, the Mediterranean and the World Economy", 27-31 August 2013, Palermo, Italy, European Regional Science Association (ERSA), Louvain-la-Neuve

This Version is available at:

<https://hdl.handle.net/10419/123834>

**Standard-Nutzungsbedingungen:**

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

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

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

**Terms of use:**

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

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

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

# Sequential city growth in the US: Does age matter?\*

**Rafael González-Val**

*Universidad de Zaragoza – Institut d’Economia de Barcelona (IEB)*

**María Sánchez-Vidal**

*Universitat de Barcelona – Institut d’Economia de Barcelona (IEB)*

**Elisabet Viladecans-Marsal**

*Universitat de Barcelona – Institut d’Economia de Barcelona (IEB)*

**ABSTRACT:** We provide empirical evidence of the dynamics of city size distribution for the whole of the twentieth century in U.S. cities and metropolitan areas. We focus our analysis on the new cities that were created during the period of analysis. The main contribution of this paper, therefore, is the parametric and nonparametric analysis of the population growth experienced by these new-born cities. Our results enable us to confirm that, when cities appear, they grow very rapidly and, as the decades pass, their growth slows or even falls into decline. Moreover, the nonparametric analysis shows that the most of the growth differential is driven by the cities’ first decade of existence. This is consistent with the theoretical framework regarding mean reversion (convergence) in the steady state and with the theories of sequential city growth.

Keywords: cities, sequential city growth, city size distribution, city age

JEL classification: O18, R11, R12

---

\* Financial support from ECO2009-09332 and ECO2010-16934 (Ministerio de Educación, Cultura y Deporte) and 2009SGR102 (Generalitat de Catalunya) projects, ADETRE research group (DGA) and FEDER is gratefully acknowledged.

## 1. Introduction

The dynamics of city size distribution and, in particular, the analysis of Gibrat's law – that is, that a city's population growth rate is independent of its initial size –has attracted the attention of researchers for many years. In fact, there are plenty of studies evaluating the performance of Gibrat's law for different countries and periods. Ioannides and Overman (2003) find that Gibrat's law holds for the US, Eeckhout (2004) points out the same when including all the cities without restriction and so does Giesen and Südekum (2011) for the case of Germany. Others such as Black and Henderson (2003) or Bosker et al. (2008) find that this is not the case both for the US and West Germany respectively. Despite this amount of literature quantifying the size effect on growth, there is little evidence on analysing the effect of city's age on its growth. In this context, this paper adopts parametric and nonparametric techniques to evaluate the age-dependent patterns of urban growth using data on cities and Metropolitan Statistical Areas (MSAs) of the US for the period 1900 to 2000. Moreover, the non parametric analysis provides additional empirical evidence to the above mentioned theories regarding the acceptance or rejection of Gibrat's law focusing on the role of new-born cities.

The inclusion of new cities is of special relevance in the case of the US which saw its cities grow in number from 10,496 to 19,211 over the 20<sup>th</sup> century. Besides, at the same time, these cities increased in population and size. Figure 1 shows the evolution of the total number of cities throughout the twentieth century in the US. At first glance we can observe that the number of cities grows over time but this growth is not the same trough all the period. In fact, the graph shows that this growth is concave, being higher during the first third of the century<sup>1</sup> and becoming more stable the years after. There are many examples of cities appearing during the 20<sup>th</sup> century. For instance, Long Beach in the State of New York, was created in 1922, and today is the 15<sup>th</sup> biggest city in the State (the 18<sup>th</sup> in 2000). With a population of 35,462 inhabitants (2000), it enjoyed an annual growth rate of between 4.5 and 5.5% during its first three decades of existence, though this rate slowed down to 0.5% in the 1990s. The second half of the twentieth century is characterized by a suburbanization process and the proliferation of cities in the south of the country. Good examples of this phenomenon are provided by Carson City and San Marcos, two cities in California. The former is, in fact, a suburb of Los Angeles and the latter of San Diego. Both cities were created during the 1960s as a consequence of the aforementioned process and the creation of the Sun Belt. Carson City was born in 1968, grew at an annual rate of 1.3% during its first decade of existence and then at a slower rate up to 2000. The case of San Marcos differs slightly. The decline in its growth rate with the passing decades is the same as in the previous case, but its annual growth rates have been much higher: ranging from 15% on average for the first decade of its existence to 3% over the last decade, growing from a settlement of just 3,896 inhabitants

---

<sup>1</sup> In fact, 62.26% of the new cities created in the whole century were born in the first three decades, while the average rate of new creations for the rest of the period stands at around 5% per decade.

in 1970 to 54,977 in 2000. These are three examples from our dataset but we can find almost 9,000 similar cases.

[FIGURE 1]

However, we are not the first focusing the analysis on new cities. Previous works by Dobkins and Ioannides (2000) and Henderson and Wang (2007) also allow new cities to enter the sample when crossing a particular threshold. However, the inclusion of all new cities without any threshold restriction is only considered in the works by Giesen and Südekum (2012) who uses data on the foundation dates of 7,000 American cities for the period 1790 to 2000 and Desmet and Rappaport (2013) whose data consist on the entire set of counties and MSAs for the US from 1800 to 2000. Our work closely relates to both studies. Giesen and Südekum (2012), by means of a theoretical model, find that the distribution of city sizes is systematically related to the country's city age distribution. They point out that young cities initially grow faster but in the long run all the cities grow at the same rate (Gibrat's law). Desmet and Rappaport (2013) point out that in earlier periods smaller counties converge and larger ones diverge but, taking into account the changes in age composition over time, both convergence and divergence dissipate and Gibrat's law gradually emerges. Our results are very much in line with theirs. We find that young small cities tend to grow at higher rates but, as decades pass, their growth stabilizes or even declines. Moreover, this high level of growth rates is spread across ages but is especially important in the first years of existence. After that, Gibrat's law tends to hold better.

Our work shows then, a sequential growth pattern of cities according to their age. To grow sequentially means that, within a country, a few cities initially grow much faster than the rest, but at some point their growth slows and other cities start to grow, and so on. This fact has been theoretically documented by Cuberes (2009) and Henderson and Venables (2009) with some models in which cities grow sequentially, allowing for the entrance of new cities in the sample. The only empirical approach to these theories is Cuberes (2011) who, drawing on data for cities from 54 countries and on data for metropolitan areas from 115 countries, shows that urban agglomerations have followed a sequential growth pattern. However, these set of studies focus on the sequential pattern driven by the size of the city while our work does so tracing the age-dependent patterns.

Moreover, we reproduce the analysis for metropolitan areas as they represent more accurate economic areas than cities. However, our results do not confirm our earlier findings for cities. This could reflect the fact that a metropolitan area is an aggregation of different cities; even if the area is new, the cities within it might not be. Moreover, it is not possible to know how old the area is since it does not enter the sample until it reaches the minimum population threshold of 50,000 inhabitants. As such, larger - and, therefore, more

mature - cities within the area, have lower growth rates than smaller cities within the same area and the aggregate effects may disappear.

The rest of the paper is structured as follows. Section 2 presents the data. Section 3 explains the parametric empirical methodology and section 4 discuss its main results. Section 5 provides the nonparametric analysis and its results. Section 6 concludes.

## 2. Data

We use data for US cities and Metropolitan Statistical Areas (MSAs) for the whole of the 20<sup>th</sup> century. The database is the same as that employed by González-Val (2010) with the addition of extra periods for the MSA dataset. The information for both geographical units was obtained from the annual census published by the US Census Bureau. From the outset, it should be borne in mind that a city can be defined in many ways. Here, for our analysis, we use that of the ‘incorporated place’. According to the census, an incorporated place is *a type of governmental unit incorporated under state law as a city, a town (except in New England, New York and Wisconsin), a borough (except in Alaska and New York city), or a village and having legally prescribed limits, powers and functions*. The Census Bureau recognizes incorporated places in all states except Hawaii, for which reason it is excluded from our sample. In addition, the states of Puerto Rico and Alaska are excluded as they (together with Hawaii) were not annexed until the second half of the 20<sup>th</sup> century. As Eeckhout (2004) stresses, the whole sample of cities in each state without restriction of size needs to be considered since otherwise a truncated distribution can produce biased results.

Data for MSAs are also used so as to take into account that part of the population that lives outside cities and so as to be able to compare the results provided by both geographical units. In line with Ioannides and Overman (2003), for the period from 1900 to 1950, we use data from Bogue’s Standard Metropolitan Areas (1953). These are based on the definition of Standard Metropolitan Areas (SMAs)<sup>2</sup> for 1950, used to reconstruct the population for the period 1900 to 1940. This means, however, that in 1900 some of the SMAs were below the 50,000-inhabitant threshold, and these are excluded until they reach that cutoff. For the period 1950 to 2000 our MSA data are taken from the Census Bureau.

As Glaeser and Shapiro (2003) point out, MSAs are multi-county units that capture labor markets and, as such, might serve as more effective economic units than incorporated places. Yet, the use of MSAs gives rise to a problem that is directly related to their definition: as an MSA usually comprises a group of counties that requires a central city with a minimum of 50,000 inhabitants (a criterion that has changed over the period of analysis), only larger cities are considered. Using MSAs gives rise to another more specific

---

<sup>2</sup> The definition of a metropolitan area was first issued in 1949 under the name of Standard Metropolitan Area (SMA). It changed to Standard Metropolitan Statistical Area (SMSA) in 1959 and in 1983 was replaced by Metropolitan Statistical Area (MSA).

problem for the analysis we conduct here. As Dobkins and Ioannides (2001) show, the US system is characterized by the entry of new cities that can have an impact on its city size distribution. As we are particularly interested in these cities, the data on incorporated places provide more information than those on the MSAs. However, MSAs are larger geographical areas and include a large proportion of the population living in rural areas. Yet, despite the fact that the sample of incorporated places accounts for a lower percentage of the total population, it is considerably more urban (94.18% in 2000) than that of the MSAs (88.35%).

Table 1 shows the descriptive statistics for the population of incorporated places in each decade of the twentieth century, while Table 2 presents the same statistics for the MSAs, the minimum threshold being 50,000 inhabitants. An initial inspection shows that the number of cities and MSAs increases over time as does their size. In fact, new-born cities represent 42.52% of the total sample of incorporated places while the number of new MSAs amounts to 180, which represents 49.85% of the sample. What these tables illustrate, therefore, is the urbanization process that the US experienced over the last century. The number of cities in 2000 is almost twice that in 1900; the number of MSAs has increased more than threefold. This is clearly indicative of the importance of taking into consideration the appearance of new units (cities or MSAs) when studying the US population growth process.

[TABLE 1]

[TABLE 2]

### 3. Empirical analysis

In the context of studies of city size distribution and, in particular, related to the sequential city growth literature, using a panel dataset we seek to test which US cities grew the most during each decade of the 20<sup>th</sup> century. In line with this literature, we expect the new-born cities to grow rapidly during the first decades of their life before stabilizing (and even declining) in the decades that follow. In order to test this hypothesis, we estimate the following model:

$$g_{i,t} = \sum_{k \geq 1} \beta_k d_{k,i,t} + \gamma city\ size_{i,t-1} + \delta_t + \varphi_s + \eta_r + \mu_s + \varepsilon_{i,t} \quad (1)$$

where the dependent variable  $g_{it}$  is the growth rate for each city (or MSA)  $i$  at time  $t$  calculated as  $g_{it} = \ln p_{it} - \ln p_{it-1}$ , being  $p$  the population. The variable  $d_k$  is a dummy capturing the age of the cities. The sub index  $k$  represents the number of decades<sup>3</sup> that a city is present in our sample. Therefore,  $d_1$

---

<sup>3</sup> As our data are divided by decades (not years), in 1910 we would have cities created from 1901 to 1910. In 1920 we find cities from 1911 to 1920. The same holds for the other eight nine decades of the century.

( $d_k$  when  $k = 1$ ) is equal to one when the city is new (first decade of existence) and zero otherwise. A city is considered new when it records a positive population in one decade while having no population in the previous one(s). Additionally,  $d_2$  ( $d_k$  when  $k = 2$ ) is equal to one if the city has existed for two decades and zero otherwise,  $d_3$  ( $d_k$  when  $k = 3$ ) is equal to one if the city has existed for three decades and zero otherwise and so on. Therefore, this dummy variable represents cities of all ages, from new-born ( $d_1$ ) to nine decades old ( $d_9$ ), for all the period of analysis. The variable *city size* controls for the initial size of the city,  $\delta_t$  is a time fixed effect,  $\varphi_s$  is a state fixed effect (in cities' estimations we also add county fixed effects),  $\eta_r$  is a region fixed effect and  $\mu_s$  is a dummy capturing other location fixed effects.  $\varepsilon_{i,t}$  is the error term.

Table 3 shows the evolution of the nine age dummies over the 20<sup>th</sup> century. For each decade,  $d_1$  is the number of new cities created in that decade so that in 1910 a total of 3,291 new cities were born; in 1920 the number was of 1,747 new cities, and so on. For each decade,  $d_2$  is the number of cities with one decade of existence. For instance, in 1950 there were 489 cities with one decade of existence in 1960 there were 627 and so on. Column  $d_3$  shows the cities with two decades of existence, column  $d_4$  the ones with three decades of existence, and so on. The total number of cities by age (independently of the year of their creation) is displayed at the bottom of each column of Table 3. This number is the sample size used in the nonparametric analysis conducted in Section 5. Moreover, we can trace the cities' evolution from the decade they first appeared until the end of the period by observing the diagonals in table 3. In fact, if  $d_1$  shows the number of new cities per decade,  $d_2$  the ones with one decade of existence,  $d_3$  those with three decades each decade and so on, then, we can trace the 3,291 newborn cities in 1910 by observing the number of cities corresponding  $d_2$  in 1920,  $d_3$  in 1930 and so on. Thus, it becomes apparent that some cities disappeared during the century because the numbers in the diagonals are not always the same. This fact can be explained by a variety of causes including hurricanes, the death of the town's benefactor or the fact that some cities expanded their borders and absorbed others. This phenomenon was concentrated in some western states and in the American Great Plains, especially during the first half of the 20<sup>th</sup> century<sup>4</sup>. However, the number of cities disappearing from the sample always represents less than a 3% of the US total number of cities (1,667 disappearances throughout the whole period), even in the first half of the century. Consequently, if we calculate the average net and gross creation rates of cities over the 20<sup>th</sup>, we find that they do not vary much between them, being the average net creation rate of cities a 6.51% and the average gross creation rate of cities a 7.46%.

[TABLE 3]

---

<sup>4</sup> See Blanchard (1960) for a fuller discussion of ghost towns in the US.

Based on the hypothesis we seek to test here, we expect  $\beta_k$  to be positive and significant during the first decades following the birth of the city but, as the decades pass, we expect this coefficient to decrease, even acquiring a negative value. However, in order to avoid any bias in these estimations, we need to add a number of controls that capture the time or space effects that might influence these results. Thus, we incorporate time and state fixed effects in our estimation. On top of the state fixed effects, for the cities' estimations, we also include county fixed effects in order to control for a smaller geographical area.

Additionally, Black and Henderson (2003) find that US cities with coastal locations grow faster and they incorporate regional variables in their analysis so as to capture their market potential. Other studies, including Rappaport and Sachs (2001), Mitchener and McLean (2003) and Bleakley and Lin (2012), also point out that having access to navigable waters plays an important role in accounting for population distribution and growth. Thus, to control for these characteristics, we also include a dummy variable that captures the access to navigable waters (including access to rivers, lakes and oceans) at the state level, and four dummy variables, one for each of the major US regions: the Northeast, the Midwest, the South and the West.

Moreover, Duranton (2007) points out that cities grow or decline following gains or losses of the industries. Therefore, we include one more control variable capturing changes in industrial composition in the US over the course of the 20<sup>th</sup> century. As Kim and Margo (2004) explain, during the first half of the twentieth century, the rise of the industrial economy and the manufacturing (or 'rust') belt saw people move westwards. Since 1950, thanks to the diffusion of air conditioning and milder winters, the population has grown in the southern part of the country, leading to the creation of the Sun Belt<sup>5</sup>. Thus, we include two dummies at the state level, one for each of the rust and sun belts respectively, in order to control for these regional and industrial impacts on the population growth rate.

Furthermore, in order to account better for the magnitude of the city age effects on its growth and distinguish them from those of the city size on its population growth, we also include a variable capturing the initial city size of the cities ( $\ln p_{it-1}$ ). In line with the literature, by including this variable we are able to test the mean reversion hypothesis. When the coefficient of this variable is negative, we can assume mean reversion (convergence) in the steady state while a positive one may indicate divergence. A non-significant coefficient can be interpreted as being indicative of independence between growth and initial size, supporting Gibrat's law and, therefore, rejecting the mean reversion hypothesis. In the previous literature, some authors like Black and Henderson (2003) and Henderson and Wang (2007) find that the smallest cities grow faster, supporting the mean reversion hypothesis. In our analysis, we have also introduced the age of the city, which is correlated with the size (Giesen and Südekum, 2013). Therefore, it may be difficult to

---

<sup>5</sup> Other studies like Rappaport (2007) also study the population mobility according to the weather conditions.



distinguish between the net effect of city age on growth from the one of size on growth in the parametric analysis. As a consequence, in order to examine the exact relationship between the temporal dimension of growth (the age of the city) and its initial size on growth, we perform a nonparametric analysis in Section 5. More specifically, the nonparametric analysis aims to evaluate which is the exact size effect on growth for every city's decade of existence. Therefore, we can also examine whether we can accept the mean reversion hypothesis or reject it (being Gibrat's law the one holding) and if there are relevant differences on the impact of city size on growth across different ages. Moreover, the city size may, in some cases, be a source of possible endogeneity. However, our results regarding the effect of city's age on its growth are robust to the inclusion or not of the city size variable.

We reproduce the analysis for the MSAs in order to test whether the growth pattern of cities still applies when aggregating the geographical units. Table 4 shows the evolution of the nine age-dummies for the MSAs during the 20<sup>th</sup> century. Two main differences can be seen between Tables 3 and 4: first, no MSAs disappear from the sample (once an MSA reaches the minimum population threshold it never falls below it) and, second, the falling trend in the appearance of new MSAs is not as clear as that for the cities. The former relates to the MSA definition: to become an MSA a minimum population of 50,000 in the central city is required. Then, when working with MSAs we are taking into account mainly larger cities with high levels of capital stock and scale externalities that make them to remain big and not disappear<sup>6</sup>. The second distinctive characteristic is attributable to the change in the criterion used to define an MSA in 1960 (47.2% of the MSAs were created that decade).

[TABLE 4]

#### 4. Results

In this section we present the results of the estimation of Eq. (1). Table 5 shows the results for cities while Table 6 presents those for the MSAs. All regressions include the nine age dummy variables. The control variables are sequentially introduced from regression (1) to (6). For both geographical units (cities and MSAs), the regressions corresponding to each column represent the same specification with only the unit of analysis being changed from cities to MSA.

The coefficients can be interpreted as the average impact, measured in logarithmic points, on the growth rate of a specific city  $i$  (or MSA) depending on the age of that city (or MSA) compared to the other pre-existing cities in the sample. As explained above,  $d_1$  represents the city when it is newly born,  $d_2$  when it has

---

<sup>6</sup> See Henderson and Wang (2007) for further explanations of how larger cities do not lose a big proportion of their population over the period 1960-2000.

existed for one decade,  $d_3$  two decades and so on, meaning that  $d_9$  represents more mature cities than  $d_1$ . Therefore, the coefficient associated with  $d_1$  represents the average impact on growth of being a new-born city with respect to the rest of pre-existing cities in the sample (base category), the one associated with  $d_2$  represents the average impact of one-decade old cities with respect to the rest of the cities i.e. the older ones, and so on so forth. For that reason, we are interested in the trend presented by the coefficients from  $d_1$  to  $d_9$ , as this represents the dynamic effects of cities' age on their growth.

Table 5 presents the results for cities where the total number of observations corresponds to the summation of all the cities that grow (the newborn and the incumbents) over the 20<sup>th</sup> century. Column (1) presents the results of estimating Eq. (1) by OLS without any control variable (not even the city size). Column (2) shows the same estimation but including the city size variable. At first glance, we see that the coefficients of the nine dummies in both specifications follow the expected pattern: they are significantly positive for  $d_1$  and become smaller until they record negative values. However, note that the results from (1) and (2) might lack precision as there may well be a considerable amount of missing and uncontrolled information in these specifications. In order to solve any problem of bias, we estimate equations (3), (4), (5) and (6) using different control variables. In column (3) we estimate the same equation but taking into account the possibility that time effects might be driving part of the results. However, the coefficients are similar to those estimated in the previous regressions as is the overall trend. As before, the coefficient associated with  $d_1$  is significantly positive and it decreases with the increase in city age, becoming negative when the city is mature.

Column (4) presents the results for the city fixed effect estimation. Here, the interpretation of the coefficients is different from those of the other five regressions. Now, the estimated parameters show how new-born city  $i$  grows in decade  $t > I$  in comparison with how new-born city  $i$  grew in decade  $t$ . An analysis of the coefficients reveals that the trend followed is the same as that in the previous estimations (the coefficient associated with  $d_1$  being higher than that associated with  $d_2$  and so on), indicating that the growth of a new-born city is greater than that of a mature city. However, the overall size of the coefficients is smaller than before. In fact, the first two dummies are not significant because they are indeed the base category<sup>7</sup> but from  $d_3$  to  $d_9$  they become significantly negative. In column (5) we estimate the same model but we include a state fixed effect and a county fixed effect to control for a spatial dimension. The results, again, present the same pattern with significantly positive coefficients associated with  $d_1$  and a decreasing trend until  $d_9$ . It is not, in fact, a perfectly decreasing trend because with the passing decades growth tends to stabilize and only declines at the end of the period. This trend can be observed in Figure 2 that plots the estimated parameters for column (5).

---

<sup>7</sup> We estimated the same regression without the incumbent cities and the results were robust.

[FIGURE 2]

Finally, column (6) shows the results when estimating Eq. (1) including all the control variables: size, time, state, county and regional fixed effects. We also include the other geographical dummy variables: access to navigable waters and belonging to the Sun or Rust belts. As in all the previous cases, the coefficients follow the same decreasing trend allowing us to demonstrate that when a city is born, its growth is high and as the decades pass, the growth becomes more moderate and even declines. The average impact on growth (in logarithmic points) of a city in the first decade of its creation is about 0.106 more with respect to the pre-existing cities (the older incumbent ones). One decade later, the coefficient falls significantly (from 0.106 to 0.023), although the impact on growth remains positive. Thus, the higher growth occurs during the first decade of a city's existence. However, if we focus on the coefficient associated with the last decade (-0.024), we see that the older the city becomes the lower is its average impact on growth. These results are consistent with the theories of sequential city growth.

However, despite the results for the city age impact on its growth are consistent across the different estimations, the coefficients associated with the city size variable do not present the same pattern. As mentioned before, this initial city size variable seeks to capture the existence of mean reversion. In line with the preceding discussion, it should be significantly negative in order to accept the mean reversion hypothesis or non significant to accept Gibrat's law. By observing the coefficients associated with this variable from columns (2) to (6), we see that all of them apart from the one corresponding to column (4) are significantly positive. As mentioned before, regressions (2), (3), (5) and (6) correspond to OLS estimations while column (4) presents the results for a city-fixed effects estimation. The OLS procedure is a between estimator, which typically uses just the cross-sectional variation to estimate the parameters while the fixed effects is a within estimation which evaluates changes in time within subjects.

Although the difference across within and between estimators is not significant when evaluating the nine dummy variables capturing the age of a city<sup>8</sup>, it is so for the city size variable. We argue that, when estimating by OLS, we are comparing across cities (big and small) while already controlling for age (but only for newborn cities which are typically the smallest). However, in the within estimation, the time variation within the city size is the exploit dimension. As a consequence, the coefficient of the OLS procedure, although very small, is significantly positive and so nothing can be said about mean reversion while the one in the fixed effects one is significantly negative confirming the existence of mean reversion.

---

<sup>8</sup> The interpretation of the coefficients varies between OLS and fixed effects but the results are the same. The coefficients associated to the nine dummy variables in the between estimation represent the difference in the impact on growth rates between the newborn cities and the pre-existing ones ( $d_1$ ), the one decade old cities and the older ones ( $d_2$ ) and so on. On the other hand, the coefficients of the within estimation can be interpreted as being  $d_1$  and  $d_2$  the base category and then, we can observe the cities' age evolution compared to the base category (the city being newborn). In both cases indeed, the results lead to the same conclusion.

Moreover, we are not the first to empirically document this moderately positive relation between initial city size and growth. Desmet and Rappaport (2013) also find it across intermediate and large locations during the late nineteenth and early twentieth centuries.

However, to differentiate between the direct effects of city size on its growth from this of city's age on its growth, we conduct a non parametric analysis in Section 5. More precisely, we are interested in whether there are systematic deviations from Gibrat's law for cities of all ages or they are focused on any specific city's age. Thus, we try to see if there are different impacts of size on growth across different city ages.

[TABLE 5]

Table 6 presents the results for the MSAs, its six columns being the same as those in Table 5, being the total number of observations, as in the case of the cities, the summation of all the MSAs that grow (the newborn and the incumbents) over the 20<sup>th</sup> century. In the first two columns (1) and (2) we cannot identify the same decreasing trend as the one we found in the cities' estimation (Table 5). These specifications might lack precision, as those first two identified above for the cities. For this reason, we also estimate the model incorporating time fixed effects, city fixed effects and state fixed effects in columns (3), (4) and (5) respectively. None of these three regressions presents the same results' pattern as in those for the cities in terms of a declining growth trend.

Finally, column (6) includes state, time and region fixed effects and the geographical controls. As with the previous columns, almost none of the coefficients are statistically significant and the expected decreasing trend is not seen. Thus, we can conclude that the MSAs do not present the same trend as the one presented by cities and that the aggregation of geographical units does not provide the same results. Moreover, the results of the MSA size in some regressions presents a significantly negative value but, when adding all the control variables (column 6) it becomes non significant. These results seem to point out that there is no relation between the initial size of a city and its growth when dealing with MSAs data and then, Gibrat's law may be confirmed for bigger units of analysis than cities.

Nevertheless, some studies like Cuberes (2011) and Desmet and Rappaport (2013) find that both cities and MSA's grow sequentially which is not happening in our estimations as our results point out that cities are the only ones following this sequential pattern. However, these results are sensitive to the unit and period of analysis and, although they use data for MSA's, their MSA definitions and the periods analysed differ from ours. Cuberes (2011) uses a worldwide dataset for many different periods. In fact, for the case of the US, he uses MSAs typically above a threshold which is not the same as ours (50,000 inhabitants). In the case of Desmet and Rappaport (2013), they work with a hybrid of metro areas and the remaining US counties for

the period 1800 to 2000. Additionally, our study starts in 1900; almost one hundred years after theirs, and then, MSA growth patterns may be different than those shown by both analysis.

Thus, a plausible explanation for our results lies in the definition of an MSA. A metropolitan area typically comprises a group of counties with a central city with a minimum of 50,000 inhabitants and a number of other smaller places located at points in the orbit of this central city. According to the sequential growth literature, the central city (assumed to be older and therefore larger than most surrounding places) will present different growth patterns over the time period to those of other cities within the same MSA. More specifically, the central city will be more mature than the rest and its growth rate is therefore not expected to be as high. By contrast, there will be other smaller and younger cities that will grow more rapidly during the same period. As such, the final growth rate of the MSA is the average of many rates of different cities weighted by city size. Another plausible explanation is that in order to become an MSA a city with more than 50,000 inhabitants is needed. Therefore, a new MSA is nothing but the evolution of the cities within it and it might be the case that the definition of a new MSA is not as accurate as the one of a newborn city.

[TABLE 6]

## 5. Nonparametric analysis

A number of studies employ a nonparametric approach to evaluate the relationship between growth and city size and growth to examine whether Gibrat's law and mean reversion in the steady state holds. Ioannides and Overman (2003), for example, undertake such an analysis with a time-series dataset for metropolitan areas. This same methodology is adopted by Eeckhout (2004) and González-Val (2010). The former uses it to evaluate the impact of city size on its growth for all the cities in the US for two specific years: 1990 and 2000. González-Val (2010) uses the same database as the one described here which includes all cities without restriction. All three studies find that Gibrat's law holds (at least for means) for their data and periods analysed. On the other hand, Michaels *et al.* (2012) regress population growth on a full set of fixed effects for initial population density using their self-made dataset of county subdivisions finding an increasing relationship between population growth and initial population density in intermediate population densities.

However, our study is much in line with Desmet and Rappaport (2013). Using data on counties and MSAs, they empirically document the relationship between the level of population and the growth rate of a city for every twenty-year period over the nineteenth and twentieth centuries. They find that, although Gibrat's law emerges gradually, it never fully attains. We perform a similar analysis, consisting on the nonparametric estimation of city growth against city size for every decade of existence of the US cities. By

doing so, we may be able to disentangle between the city age and city size effects on growth and extract some conclusions about the acceptance or rejection of the mean reversion hypothesis. Moreover, we also examine whether this conclusion holds independently of the age of the city or, on the contrary, it differs across ages.

Our nonparametric approach is the same as the methodology developed by Ioannides and Overman (2003) and used in Eeckhout (2004) and González-Val (2010), but differs in terms of the data we use. Thus, we include only the cities identified as being new-born in each decade and estimate a pool for any possible city age, from one to nine<sup>9</sup>. This means that in decade one, we include the total number of cities with one decade of existence; no matter the year in which they were created (the last row in Table 3 is the sample size for each estimation)

The regression we estimate is the following:

$$g_i = m(s_i) + \varepsilon_i$$

where  $g_i$  is the normalized growth rate, i.e., the difference between growth and the contemporary sample mean divided by the contemporary standard deviation and  $s_i$  is the logarithm of the population size of a city.  $\varepsilon_i$  is the error term. The aim of this approach is to provide an estimation of  $m(s_i)$  without imposing any specific parametric functional form. The estimation of  $m(s_i)$  is a local average that smooths the value around the point  $s$ . The smoothing is conducted using a kernel which is a symmetrical, weighted and continuous function around  $s$ . The Nadaraya-Watson method<sup>10</sup> is used to calculate the estimate of  $m$ , based on the following expression:

$$\hat{m}(s) = \frac{n^{-1} \sum_{i=1}^n K_h(s - s_i) g_i}{n^{-1} \sum_{i=1}^n K_h(s - s_i)}$$

where  $K_h$  denotes the dependence of  $K$  on the bandwidth  $h$ , and where  $K$  is an Epanechnikov kernel<sup>11</sup>. Figure 2 shows the results for  $\hat{m}(s)$  calculated for a bandwidth of  $h = 0.5$ <sup>12</sup> for every decade of the twentieth century including only the new-born cities. Bootstrapped 95 percent confidence bands, calculated using 500 random samples with replacement, are also displayed.

<sup>9</sup> We consider a city age up to nine decades (i.e., over the course of the twentieth century). To be able to consider a city with an age of ten decades, data for 1890 are required.

<sup>10</sup> Employed here as used in Härdle (1990).

<sup>11</sup> The results are robust to the use of a Gaussian kernel as well as the local polynomial fit technique.

<sup>12</sup> The results are robust to different bandwidths including the optimal one for each decade

### [FIGURE 3]

This type of analysis allows us to visually compute the temporal evolution of cities by their size. If random growth does not exist, the average growth of the smallest cities would differ from that of the largest ones. If this were not the case, the figures would only present horizontal lines on the zero value of the growth axis and there would be no deviation from the mean. In Figure 2, it is immediately apparent that smaller cities of all ages present higher growth rates, and that the larger the city the lower its growth rate tends to be. However, as a city becomes bigger (city size increases), the average growth stabilizes in the mean. Therefore, it seems that, for every decade of the 20<sup>th</sup> century, smaller cities tend to grow more.

However, if we plot all the decades together in the same graph, this conclusion can be narrower. Figure 3 shows the nine different estimations on figure 2 in the same plot. Then, although we can argue that there are still some differences in growth rates between the smaller cities and the rest, those differences are much higher for the younger ones (those which are one-decade old). In fact, the dashed line corresponds to the youngest cities in the sample while the others are those corresponding to the cities between two and nine decades old. By examining them, we clearly see that, apart from the one corresponding to the young cities, the others almost look flat around the zero value of growth. Therefore, despite we can assume that there are deviations from Gibrat's law for cities of all ages, they are especially important when regarding young cities. In fact, as cities become older, Gibrat's law may fit better<sup>13</sup>. Moreover, these results do not contradict our parametric findings in which we find that the higher impact of age on growth is mainly driven by the first decade of existence.

### [FIGURE 4]

Our results are in line with those of Giesen and Südekum (2012) that, by means of a theoretical model, find that cities grow with the same expected rate in the long run (Gibrat's law), but young cities initially grow faster. We are also in line with the results of Desmet and Rappaport (2013) who, using different data, find that city size and growth are negatively correlated across small locations for the 19<sup>th</sup> and early 20<sup>th</sup> centuries but Gibrat's law gradually emerges while time passes though without completely hold.

## 6. Conclusions

In this paper we have drawn on data for cities and MSAs in the United States in order to study the

---

<sup>13</sup> From the parametric results of MSA's, which, almost by definition, represent more mature (and bigger) cities, we can clearly see in regresión (6) that the MSA size coefficient is not statistically significant, showing that Gibrat's law fits better as cities become older (and bigger).

evolution of city growth throughout the twentieth century. More specifically, we have focused our attention on the role played by the new-born cities that have been created during the decades of our period of analysis. Applying parametric and nonparametric methods we have obtained two main results. Our first finding is that differences exist in city growth rates according to the age of the city. In general, when a city is born it records very high rates of growth but as the decades pass it matures and its growth stabilizes or even declines. These results are consistent with those of the sequential city growth literature, which reports that in each decade a few cities will grow at a faster rate than the others.

Our second finding is related to the analysis of the dynamics of the city size distribution, i.e. the study of Gibrat's law. We perform nonparametric regressions to examine the relationship between the temporal dimension of growth (the age of the city) and a city's initial size. Our results confirm that there are deviations from Gibrat's law for cities of all ages but they are especially important when regarding young cities. In fact, as cities become older, Gibrat's law may fit better. Therefore, these results point out that most of the growth differential is driven by the first decade of existence, which is pretty much in line with our parametric results as well as with the recent papers analyzing the age impact of cities on their growth.

Our results are very much in line with those presented by the city growth literature and, in particular, with those in studies of sequential city growth. Furthermore, our findings could provide interesting input for policy makers in developing countries such as China and India, which are now experiencing their own processes of urbanization. In recent decades, both countries have experienced a change from a rural to an urban society, i.e., the same pattern followed by the US and many other developed countries. As urban policies slowly adjust to the dynamics of growth, and given the huge populations of both India and China, it must surely be in the best interests of these countries' policy makers to learn lessons from experiences such as that of the US. In fact, if there is a statistical regularity driving some of the population growth of cities, dependent on their initial size or age, some investment (especially in public infrastructure) can be performed strategically.



## REFERENCES

- Black, D., Henderson, V., 2003. Urban Evolution in the USA. *Journal of Economic Geography*. 3, 343-372.
- Blanchard, F.S., 1960. *Ghost towns of New England: their ups and downs*, ed. Dodd, Mead & Company, New York.
- Bleakley, H., Lin, J., 2012. Portage and path dependence. *Quarterly Journal of Economics*. 127(2), 587-644.
- Bosker, M., Brakman S., Garretsen H., Schramm, M., 2008. A century of shocks: the evolution of the German city size distribution 1925–1999. *Regional Science and Urban Economics*. 38, 330–347.
- Bogue, D., 1953. *Population Growth in Standard Metropolitan Areas: 1900-1950*. Washington, DC: Housing and Home Finance Agency.
- Cuberes, D., 2009. A model of sequential city growth. *The BE Journal of Macroeconomics*. 9 (1) (Contributions), Article 18.
- Cuberes, D., 2011. Sequential City Growth: Empirical evidence. *Journal of Urban Economics*. 69, 229-239.
- Desmet, K., Rappaport, J., 2013. The settlement of the United States, 1800 to 2000: The Long Transition to Gibrat's Law. CEPR Discussion Paper #9353.
- Dobkins, L., Ioannides, Y.M., 2000. Dynamic Evolution of the US City size Distribution, in: Huriot, J.M., Thisse, J.F. (Eds.), *The Economics of Cities*. Cambridge University Press, Cambridge, pp. 217-260.
- Dobkins, L., Ioannides, Y.M., 2001. Spatial Interactions among US cities: 1900-1990, *Regional Science and Urban Economics*. 31, 701-731.
- Duranton, G., 2007. Urban evolutions: the fast, the slow and the still. *American Economic Review*. 97(1), 197-221.
- Eeckhout, J., 2004. Gibrat's law for (All) Cities. *American Economic Review*. 94 (5), 1429-1451.
- Giesen K., Südekum J., 2012. City Age and City Size. CESifo Working Paper Series No. 3730
- Giesen K., Südekum J., 2011. Zipf's Law for cities in the regions and the country. *Journal of Economic Geography*. 11(4), 667–686.
- Glaeser, E.L., Shapiro, J., 2003. Urban Growth in the 1990s: Is City Living Back?. *Journal of Regional Science*. 43 (1), 139-165.
- González-Val, R., 2010. The Evolution of the US City Size Distribution from a long term perspective (1900-2000). *Journal of Regional Science*. 50 (5), 952-972.
- Härdle, W., 1990. *Applied Nonparametric Regression*. Econometric Society Monographs. Cambridge, New York and Melbourne: Cambridge University Press.
- Henderson, J.V., Venables, A.J., 2009. Dynamics of city formation. *Review of Economic Dynamics*. 2, 233-254.
- Henderson, J.V., Wang, H.G., 2007. Urbanization and city growth: the role of institutions. *Regional Science and Urban Economics*. 37 (3), 283-313.
- Ioannides, Y.M., Overman, H.G., 2003. Zipf's Law for Cities: An Empirical Examination. *Regional Science and Urban Economics*. 33, 127-137.

Kim, S., Margo, R.A., 2004. Historical Perspectives on US Economic Geography, in Henderson, J.V., Thisse, J.F. (Eds.), *Handbook of Urban and Regional Economics*, vol 4. Amsterdam: Elsevier Science, North-Holland, Chapter 66, pp.2982-3019.

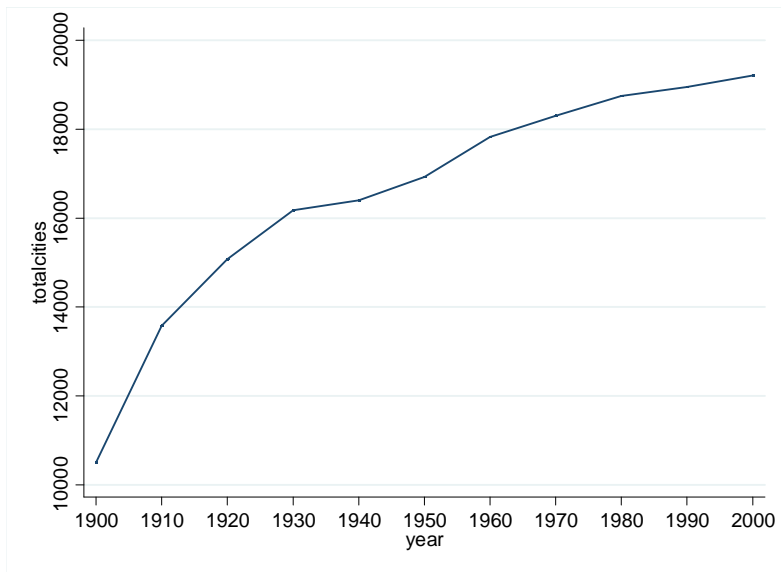
Michaels, G., Rauch, F., Redding, S.J., 2012. Urbanization and structural transformation. *Quarterly Journal of Economics*. 127(2), 535-586.

Mitchener, K., McLean, J., 2003. The productivity of US states since 1880. *Journal of Economic Growth*. 8 (1), 73-114.

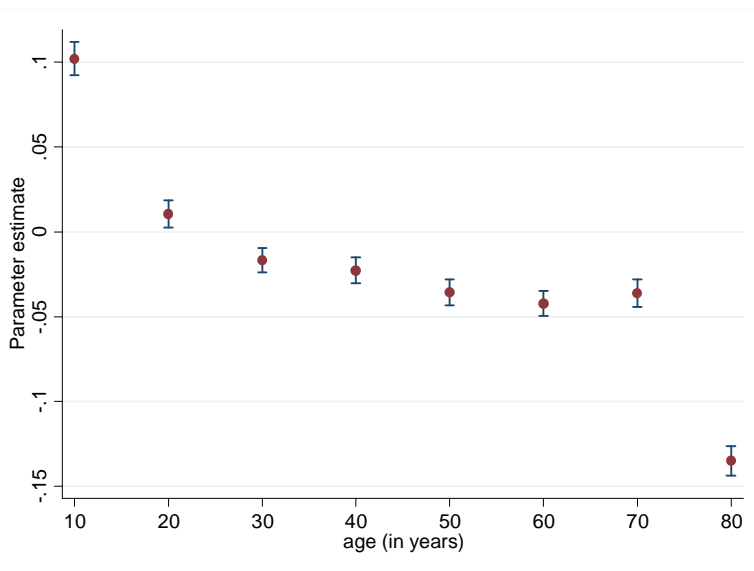
Rappaport, J., 2007. Moving to nice weather. *Regional Science and Urban Economics*. 37(3), 375-398.

Rappaport, J., Sachs, J., 2001. The US as a coastal nation. *Journal of Economic Growth*. 8 (1), 5-46.

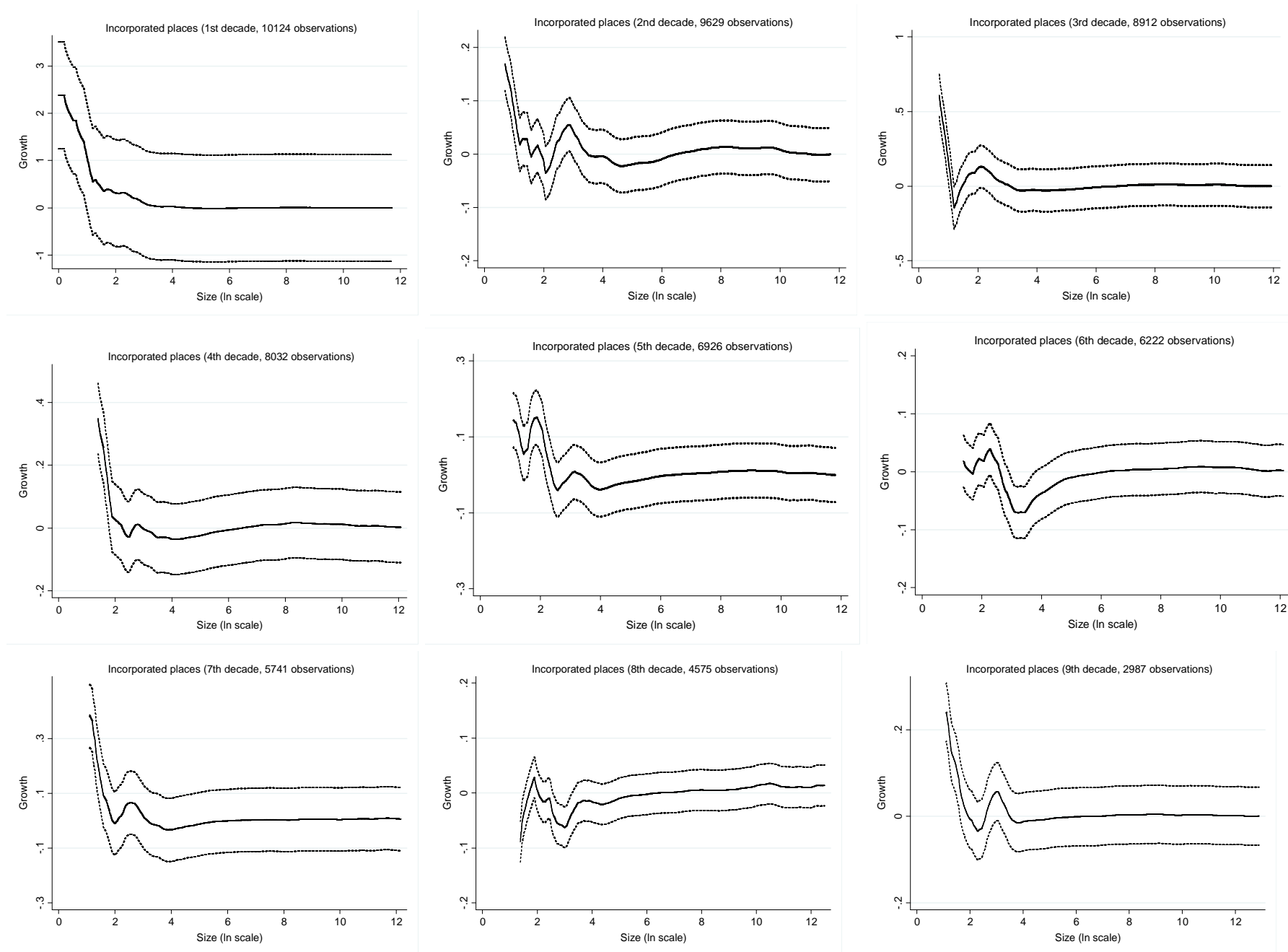
**Figure 1.** Evolution of the number of cities in the U.S. over the 20<sup>th</sup> century



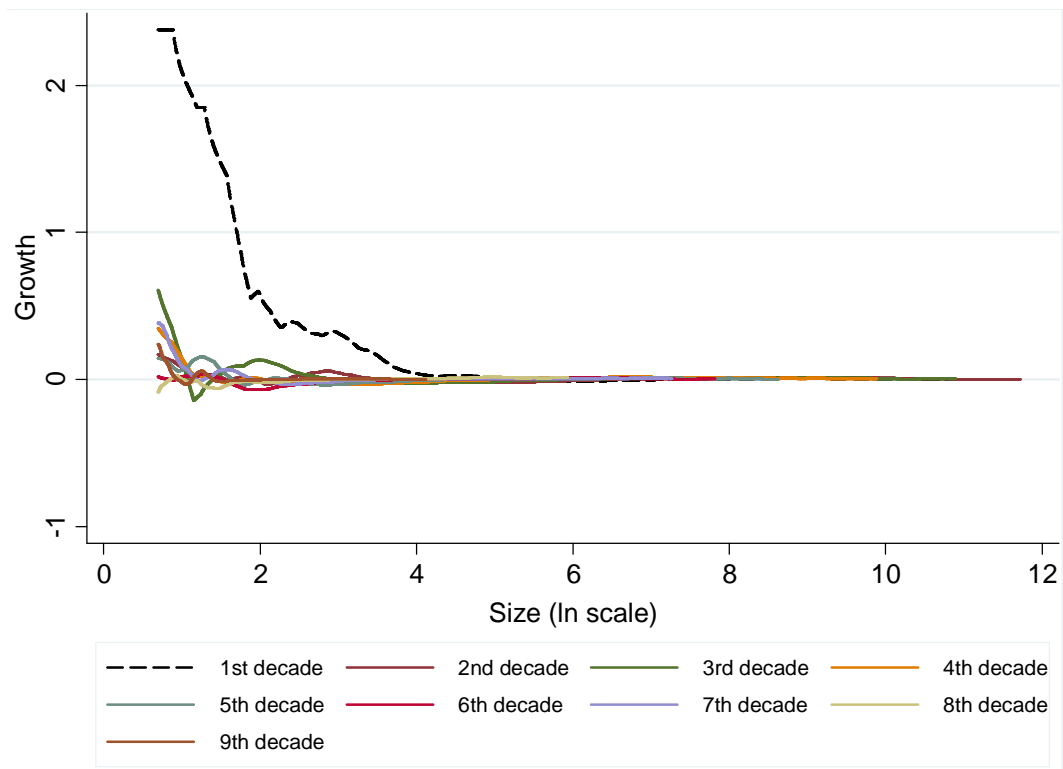
**Figure 2.** Estimated parameters for regression (5)



**Figure 3. Growth and size by age**



**Figure 4.** Population growth and size



**Table 1.** Descriptive statistics of cities

Year	Cities	Mean Size	Standard Deviation	Minimum	Maximum
1900	10,496	3,468.27	42,617.51	7	3,437,202
1910	13,577	3,610.36	50,348.78	7	4,766,883
1920	15,073	4,087.61	57,540.69	3	5,620,048
1930	16,183	4,771.31	68,462.35	1	6,930,446
1940	16,400	4,977.44	72,001.37	1	7,454,995
1950	16,923	5,662.07	76,487.59	2	7,891,957
1960	17,825	6,455.86	75,195.01	1	7,781,984
1970	18,302	7,149.50	75,690.26	4	7,895,563
1980	18,752	7,431.72	69,475.36	2	7,071,639
1990	18,953	7,998.27	72,178.75	2	7,322,564
2000	19,211	8,939.77	78,175.03	1	8,008,278

Note: Alaska, Hawaii and Puerto Rico are excluded

**Table 2.** Descriptive statistics of MSAs

Year	MSAs	Mean Size	Standard Deviation	Minimum	Maximum
1900	104	280,915	586,361	52,577	5,048,750
1910	130	307,261	719,325	50,731	7,049,047
1920	139	362,905	847,072	51,284	8,490,694
1930	145	445,147	1,063,769	50,872	10,900,000
1940	148	473,984	1,125,419	51,782	11,700,000
1950	150	570,480	1,127,541	56,141	12,900,000
1960	264	477,991	1,095,872	51,616	13,000,000
1970	268	561,378	1,318,920	53,766	16,100,000
1980	279	617,269	1,455,040	57,118	18,900,000
1990	348	588,405	1,457,107	51,359	19,500,000
2000	350	658,734	1,510,498	52,457	18,300,000

Notes: (1) The minimum threshold is 50,000 inhabitants  
(2) Alaska, Hawaii and Puerto Rico are excluded

**Table 3.** Evolution of cities over the 20th century

year	d <sub>1</sub>	d <sub>2</sub>	d <sub>3</sub>	d <sub>4</sub>	d <sub>5</sub>	d <sub>6</sub>	d <sub>7</sub>	d <sub>8</sub>	d <sub>9</sub>
1910	3,291	0	0	0	0	0	0	0	0
1920	1,747	3,229	0	0	0	0	0	0	0
1930	1,267	1,711	3,171	0	0	0	0	0	0
1940	505	1,245	1,684	3,132	0	0	0	0	0
1950	646	489	1,210	1,657	3,088	0	0	0	0
1960	1,046	627	470	1,164	1,614	3,025	0	0	0
1970	756	1,025	619	459	1,155	1,597	3,010	0	0
1980	553	750	1,008	612	457	1,143	1,588	2,987	0
1990	313	553	750	1,008	612	457	1,143	1,588	2,987
Total	10,124	9,629	8,912	8,032	6,926	6,222	5,741	4,575	2,987

Source: Self elaboration with US Census Bureau data

**Table 4.** Evolution of MSAs over the 20th century

year	d <sub>1</sub>	d <sub>2</sub>	d <sub>3</sub>	d <sub>4</sub>	d <sub>5</sub>	d <sub>6</sub>	d <sub>7</sub>	d <sub>8</sub>	d <sub>9</sub>
1910	26	0	0	0	0	0	0	0	0
1920	9	26	0	0	0	0	0	0	0
1930	6	9	26	0	0	0	0	0	0
1940	3	6	9	26	0	0	0	0	0
1950	2	3	6	9	26	0	0	0	0
1960	114	2	3	6	9	26	0	0	0
1970	4	114	2	3	6	9	26	0	0
1980	11	4	114	2	3	6	9	26	0
1990	69	11	4	114	2	3	6	9	26
Total	244	175	164	160	46	44	41	35	26

Source: Self elaboration with US Census Bureau data

**Table 5.** Estimation of the dynamic effects of cities

Dependent variable: population growth at the city level

Decades of existence	(1)	(2)	(3)	(4)	(5)	(6)
d <sub>1</sub>	0.142*** (0.005)	0.169*** (0.005)	0.154*** (0.005)	-0.079 (0.079)	0.111*** (0.004)	0.106*** (0.005)
d <sub>2</sub>	0.048*** (0.004)	0.07*** (0.004)	0.070*** (0.004)	-0.129 (0.079)	0.017*** (0.004)	0.023*** (0.004)
d <sub>3</sub>	0.017*** (0.003)	0.036*** (0.003)	0.036*** (0.003)	-0.144* (0.079)	-0.010*** (0.003)	-0.004 (0.003)
d <sub>4</sub>	0.003 (0.003)	0.019*** (0.003)	0.004 (0.003)	-0.159** (0.079)	-0.016*** (0.004)	-0.0267*** (0.004)
d <sub>5</sub>	-0.016*** (0.003)	-0.0007 (0.003)	-0.023*** (0.004)	-0.173** (0.079)	-0.029*** (0.004)	-0.044*** (0.004)
d <sub>6</sub>	-0.025*** (0.003)	-0.009*** (0.003)	-0.016*** (0.003)	-0.160** (0.079)	-0.036*** (0.003)	-0.0281*** (0.004)
d <sub>7</sub>	-0.028*** (0.003)	-0.013*** (0.003)	-0.015*** (0.003)	-0.155* (0.079)	-0.031*** (0.004)	-0.009*** (0.004)
d <sub>8</sub>	-0.096*** (0.003)	-0.082*** (0.003)	-0.033*** (0.004)	-0.170** (0.079)	-0.131*** (0.004)	-0.024*** (0.004)
d <sub>9</sub>	-0.02*** (0.004)	-0.005 (0.004)	-0.020*** (0.005)	-0.162** (0.079)	---	---
City size <sub>t-1</sub>		0.025*** (0.0005)	0.028*** (0.0006)	-0.219*** (0.003)	0.006*** (0.0009)	0.007*** (0.0009)
City fixed effects	No	No	No	Yes	No	No
Time effects	No	No	Yes	Yes	No	Yes
County effects	No	No	No	No	Yes	Yes
State effects	No	No	No	No	Yes	Yes
Region effects	No	No	No	No	No	Yes
Navigable waters	No	No	No	No	No	Yes
Sun & Rust Belts	No	No	No	No	No	Yes
Observations	160,292	160,292	160,292	160,292	130,836	130,836
R-squared	0.019	0.034	0.042	0.194	0.156	0.174

Notes: (1) Robust standard errors in parentheses (\*\*\*) p&lt;0.01, \*\* p&lt;0.05, \* p&lt;0.1)

(2) d<sub>9</sub> is not included in regressions (4) and (5) because collinearity problems with the county effects(3) The number of observations varies between regressions (5) and (6) and the rest due to the exclusion of d<sub>9</sub>



**Table 6.** Estimation of the dynamic effects of MSAs

Dependent variable: population growth at the MSA level

Decades of existence	(1)	(2)	(3)	(4)	(5)	(6)
d <sub>1</sub>	0.053*** (0.012)	0.035** (0.014)	0.117*** (0.018)	0.007 (0.035)	-0.028*** (0.008)	0.035* (0.021)
d <sub>2</sub>	0.067*** (0.015)	0.052*** (0.016)	0.121*** (0.020)	0.025 (0.033)	-0.042*** (0.015)	0.042** (0.021)
d <sub>3</sub>	0.010 (0.022)	-0.001 (0.023)	0.082*** (0.025)	-0.014 (0.031)	-0.027* (0.015)	-0.011 (0.028)
d <sub>4</sub>	0.028** (0.013)	0.019 (0.013)	0.091*** (0.014)	0.003 (0.026)	-0.086*** (0.024)	-0.001 (0.017)
d <sub>5</sub>	0.180*** (0.033)	0.173*** (0.034)	0.151*** (0.030)	0.064* (0.033)	-0.064*** (0.014)	0.067** (0.029)
d <sub>6</sub>	0.003 (0.018)	0.001 (0.019)	0.069*** (0.018)	-0.006 (0.023)	0.084** (0.033)	-0.011 (0.019)
d <sub>7</sub>	0.026 (0.021)	0.025 (0.022)	0.098*** (0.022)	0.035 (0.025)	-0.081*** (0.017)	0.019 (0.022)
d <sub>8</sub>	-0.047 (0.042)	-0.045 (0.043)	0.032 (0.042)	-0.021 (0.027)	-0.052*** (0.020)	-0.050 (0.039)
d <sub>9</sub>	-0.039** (0.017)	-0.035* (0.019)	0.053*** (0.018)	---	-0.122*** (0.017)	-0.022 (0.018)
MSA size <sub>t-1</sub>		-0.013* (0.003)	0.009 (0.009)	-0.138*** (0.020)	-0.028*** (0.004)	-0.002 (0.004)
MSA fixed effects	No	No	No	Yes	No	No
Time effects	No	No	Yes	Yes	No	Yes
State effects	No	No	No	No	Yes	Yes
Region effects	No	No	No	No	No	Yes
Navigable waters	No	No	No	No	No	Yes
Sun & Rust Belts	No	No	No	No	No	Yes
Observations	1,975	1,975	1,975	1,611	1,975	1,975
R-squared	0.036	0.040	0.153	0.244	0.201	0.301

Notes: (1) Robust standard errors in parentheses (\*\*\*) p&lt;0.01, \*\* p&lt;0.05, \* p&lt;0.1)

(2) d<sub>9</sub> is not included in regression (3) because collinearity problems with the city-fixed effects(3) The number of observations varies between regression (4) and the rest due to the exclusion of d<sub>9</sub>