What makes cities bigger and richer? Evidence from 1990-2000 in the US

González-Val, Rafael

Universidad de Zaragoza

9 June 2009

What makes cities bigger and richer? Evidence from 1990-2000 in the US

Rafael González-Val

Departamento de Análisis Económico

Universidad de Zaragoza

Abstract: This paper analyses the determinants of growth of American cities, understood as growth of the population or per capita income, from 1990 to 2000. This empirical analysis uses data from all cities with no size restriction (our sample contains data for 21,655 cities). The results show that while population growth in cities appears to be independent of initial size, the growth of city per capita income is negatively correlated to initial per capita income: the richest cities grew less in this period. To try to explain these differentiated behaviors, we examine the relationship between urban characteristics in 1990 and city growth (both in population and in per capita income) using a Multinomial Logit Model. The geographical situation of cities seems to play a key role in their growth.

Keywords: City growth, Multinomial logit.

JEL: R00, R11, R12.

Address: Rafael González-Val,

Dpto. de Análisis Económico, Universidad de Zaragoza

Facultad de CC. Económicas y Empresariales

Gran Vía, 2, 50005 Zaragoza (Spain)

E-mail: rafaelg@unizar.es

1. Introduction

Jacobs (1969) was the first to suggest that the city is the basic economic unit of each country when she stated "cities are also primary economic organs". Later, other writers would argue in the same way[1] (Duranton, 2000; Quigley, 1998; Fujita and Thisse, 2002). And indeed, some very special characteristics coincide in the city as an economic unit. First, among cities there is complete freedom of movement in labor and capital (they are completely open economies). Also, it is in cities where knowledge spillovers are most easily generated and transmitted, documented both at the theoretical level (Loury, 1979; Garicano and Rossi-Hansberg, 2006) and empirically (Glaeser et al., 1992; Henderson et al., 1995). Finally, the New Economic Geography adds that cities are a source of agglomeration economies (Duranton and Puga, 2004).

The starting point for this work is the idea that the city has a double nature, on one hand as a population centre and on the other as a motor of economic growth, and that the different external effects generated in cities can potentially have different effects on population growth and per capita income growth. In particular, this paper analyses the determinants of growth of American cities, understood as growth of the population or per capita income, from 1990 to 2000. This empirical analysis uses data from all cities with no size restriction (our sample contains data for 21,655 cities).

We will use a two-steps strategy. First, we analyse if the city population and city per capita income distributions have followed similar paths in the 1990s. The results show that while population growth in cities appears to be independent of initial size, the growth of per capita income is negatively correlated to initial per capita income: the richest cities grew less in this period. This explains why, while the empirical distribution of city population remains stable in the decade 1990-2000, the empirical distribution of city per capita income changes.

Second, to try to explain these differentiated behaviors, we examine the relationship between urban characteristics in 1990 and city growth (both in population and in per capita income) using a Multinomial Logit Model. Apart from initial levels of population and per capita income, we will focus on analysing the role played by employment, including variables reflecting the productive structure (percentage of employment by sector: agriculture, construction, manufacturing, services, etc.) and the unemployment rate. We will also use median travel time as a variable reflecting the costs of urban congestion, human capital variables, and geographical variables.

The American case has already been dealt with in earlier literature, using different econometric techniques and considering different periods and sample sizes. The two most direct precedents are Glaeser et al. (1995) and Glaeser and Shapiro (2003).

Glaeser et al. (1995) examine the urban growth patterns in the 200 most populous cities in the US between 1960 and 1990 in relation to various urban characteristics in 1960. They show income and population growths are

(1) positively related to initial schooling,

(2) negatively related to initial unemployment, and

[1] A good commentary on the relationship between cities and national economic growth can be found in Polèse (2005).

(3) negatively related to the initial share of employment in manufacturing.

This behavior would have continued during the decade 1990-2000, conclude Glaeser and Shapiro (2003) using a slightly larger sample size (they imposed a minimum population threshold of 25,000 cities, considering the 1,000 most populous cities). During this decade the three most relevant variables would be human capital, climate and individuals' transport systems (public or private). The growth of cities was determined by three main trends:

(1) cities with strong human capital bases grew faster than cities without skills,

(2) people moved to warmer, drier places, and

(3) cities built around the automobile replaced cities that rely on public transportation.

Other empirical studies exist analysing American population and per capita income growth, although the geographical unit analysed is not the city. At the county level, Beeson et al. (2001) studied the evolution of population from 1840 to 1990, while Young et al. (2008) analyse the evolution of income distribution from 1970 to 1998. Mitchener and McLean (2003) use data beginning in 1880 to study variations among states in labor productivity. Finally, Yamamoto (2008) examined the disparities in per capita income in the period 1955-2003 using different geographical levels (counties, economic areas, states and regions).

The main contribution of this paper compared to earlier studies is the use of the distribution of all cities, without size restrictions. The reason is that larger cities present very concrete characteristics, which also differentiate them from other cities in the distribution. By focusing only on the most populous cities, part of the story was not being told.

Table 1 presents the values of the averages and standard deviations of different variables for the entire distribution of cities in 1990, and for the 1000 and 200 largest cities. We can see how the most populous cities bear a greater congestion cost, measured by travel time, although its inhabitants enjoy higher levels of education. However, the most interesting differences are in productive structure. In the biggest cities, the services sector has more weight, while the employment percentage in the agriculture, forestry, fishing, mining, construction, and manufacturing sectors is below average when considering the whole sample. The most populous cities are also characterised by a higher unemployment rate and lower economic growth.

However, it could be said that our sample includes places which should not be considered urban, due to their small populations. Despite this, the results we obtain with our sample of 21,655 cities are similar to those of Glaeser et al. (1995). Thus, we find that the probability of a city being in the 25% of cities with most growth in income or population (i.e., the probability of the growth rate of per capita income or population being in the top quartile of the distribution) depends

(1) positively on the initial percentage of inhabitants with higher educational levels (some college or higher degree), although the sign and intensity of the effect change when considering a wider concept of education (high school graduate or higher degree);

(2) negatively on initial unemployment levels, and

(3) negatively on the initial percentage of employment in the manufacturing sector, although this sector seems to have lost weight, as other economic sectors have a greater influence on probability.

Geography also seems to have a strong influence on cities' per capita income or population growth rate.

The next section studies the evolution of per capita income and population growth in cities in the 1990s. The analysis continues in section 3, using a Multinomial Logit Model (MNLM) to examine the relationship between urban characteristics in 1990 and city growth, both in population and in per capita income. The paper ends with our conclusions.

2. City Population and City Per Capita Income: Twin paths or not?

Our first step is to analyse if city population and city per capita income distributions followed similar paths in the 1990s. Figure 1 shows scatter plots of city per capita income growth and city population growth (logarithmic scale) against initial levels in 1989 and 1990, respectively. We use data from the entire distribution of cities without any size restriction: 21,655 places.

We can observe that while in the case of city per capita income there is a clear negative relationship between the initial income level and the growth rate, for population growth it is difficult to deduce any relationship between initial size and growth. Thus, while the slope β of the line adjusted with OLS in the case of city per capita income growth is a clearly significant and negative coefficient (-0.1471), with population growth this coefficient is very close to zero (0.0026) and while it is significantly different to zero at 5%, it is not at 1%. This result, that initial population size does not influence its growth, is not new in urban economics. In fact, proportionate growth is a well-known empirical regularity known as Gibrat's law[2]. Recently Eeckhout (2004) studied the case of American cities during the period 1990-2000, also using data from the entire distribution, and concluded that Gibrat's law was fulfilled in that decade.

We would expect this different behavior to have different consequences in the evolution of distributions. Figure 2 shows the estimated empirical distributions using an adaptive kernel of city size, whether in per capita income or in population. It highlights an important change in the distribution of city per capita income. The negative relationship observed earlier between initial city per capita income and growth, which we can identify with convergent growth, has clearly produced a rightwards displacement of the distribution[3]. Meanwhile, there is hardly any change in the population distribution of the cities, as a consequence of their proportionate population growth.

Finally, Figure 3 relates city population growth and city per capita income growth. Have the cities which grew most in terms of population also grown the most in income, or vice versa? The graph shows a cloud of points with no apparent

[2] Gibrat (1931) observed that the size distribution (measured in sales or number of employees) of firms tends to be lognormal, and his explanation was that the growth process of firms could be multiplicative and independent of firm size. Starting from the 90s, this proposition has given rise to numerous empirical studies in the field of urban economics, testing its validity for the city size distribution.

[3] Everything seems to indicate that this behavior has been produced for decades. Figure 2 of Young et al. (2008), corresponding to the evolution of the Distribution of U.S. Counties' Log Per Capita Incomes from 1970 to 1998, presents a very similar effect to that observed in our estimated kernel of city per capita income distribution from 1989 to 1999.

relationship[4], leading us to conclude that during this period there was no relationship between economic growth and population growth in American cities.

However, the differentiated behavior observed in the growth rates of cities' per capita income and population seems to corroborate our initial idea: the different external effects generated in cities can produce different effects in population growth and per capita income growth. Therefore, the next section analyses the relationship between city characteristics in 1990 and city growth, both in population and in per capita income.

3. Empirical model and results

3.1. Data description

We use data for all cities in the Unites States (21,655), without imposing any minimum population cut-off point, as our proposal is to cover the entire distribution. The data came from the census[5] for 1990 and 2000. We identified cities as what the US Census Bureau calls places. This generic name, since the 2000 census, includes all incorporated and unincorporated places.

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 the New England states, New York, and Wisconsin), borough (except in Alaska and New York), or village and having legally prescribed limits, powers, and functions. On the other hand there are the unincorporated places (which were renamed Census Designated Places, CDPs, in 1980), which designate a statistical entity, defined for each decennial census according to Census Bureau guidelines, comprising a densely settled concentration of population that is not within an incorporated place, but is locally identified by a name. Evidently, the geographical boundaries of unincorporated places may change if settlements move, so that the same unincorporated place may have different boundaries in different census. They are the statistical counterpart of the incorporated places. The difference between them in most cases is merely political and/or administrative. Thus for example, due to a state law of Hawaii there are no incorporated places there; they are all unincorporated.

The explicative variables chosen are similar to those in other studies on city growth in the US and city size, and correspond to the initial 1990 values. The influence of these variables on city size has been empirically proven by other works studying the largest cities (see Glaeser and Shapiro, 2003). Table 1 presents the variables, which can grouped in four types: congestion cost variables, human capital variables, productive structure variables, and geographical variables. It is apparent that in general, standard deviations are somewhat lower in the biggest cities, which shows that the most populous cities are very similar in their economic structure, while by considering all population centres, we collect more heterogeneous behaviors.

Urban congestion cost variables are basically intended to reflect the effect of city size on urban growth. For this we use two variables: a dummy variable taking value 1 if the city population in 1990 is more than 25,000 inhabitants, enabling us specifically to control the most populous cities of the sample, and the variable Median travel time to work (in minutes), representing the commuting cost borne by workers. This is one of the most characteristic congestion costs of urban growth, explicitly considered in some

[4] In this case the adjusted line is not shown because the estimated slope β (-0.0153) is not significantly different to zero even at 5%.

[5] The US Census Bureau offers information on a large number of variables for different geographical levels, available on its website: www.census.gov.

theoretical models; that is, the idea that as a city's population increases, so do costs in terms of the time taken by individuals to travel from home to work.

As for human capital variables, there are many studies demonstrating the influence of human capital on city size, as cities with better educated inhabitants tend to grow more. We took two human capital variables: Percent population 18 years and over: High school graduate (includes equivalency) or higher degree, and Percent population 18 years and over: Some college or higher degree. The former represents a wider concept of human capital, while the latter centres on higher educational levels (some college, Associate degree, Bachelor's degree, and Graduate or professional degree).

The third group of variables, referring to productive structure, contains the unemployment rate and the distribution of employment by sectors. The distribution of labor among the various productive activities provides valuable information about other characteristics of the city. Thus, the employment level in the primary sector (agriculture; forestry; fishing and hunting; and mining) also represents a proxy of the natural physical resources available to the city (cultivable land, port, etc.). This is also a sector which, like construction, is characterised by constant or even decreasing returns to scale.

Employment in manufacturing informs us of the level of local economies of scale in production, as this is a sector which normally presents increasing returns to scale. The level of pecuniary externalities also depends on the size of the industrial sector. Marshall put forward that (i) the concentration of companies of a single sector in a single place creates a joint market of qualified workers, benefiting both workers and firms; (ii) an industrial centre enables a larger variety at a lower cost of concrete factors needed for the sector which are not traded, and (iii) an industrial centre generates knowledge spillovers. This approach forms part of the basis of economic geography models, along with circular causation: workers go to cities with strong industrial sectors, and firms prefer to locate nearer larger cities with bigger markets. Thus, industrial employment also represents a measurement of the size of the local market. Another proxy for the market size of the city is the employment in commerce, whether retail or wholesale.

Information is also included on employment in the most relevant activities in the services sector, which are more important in the most populous cities: Finance, insurance, and real estate, Educational, health, and other professional and related services, and employment in the Public administration.

Finally, we include several dummies which give us information about geographic localisation, and which take the value 1 depending on the region in which the city is located (Northeast Region, Midwest Region, or South Region; the West Region is used as a control category). Figure 4 is a map showing which states make up each of these regions, and how places of more than 10,000 inhabitants are distributed spatially. These dummies show the influence of a series of variables for which individual data are not available for all places, and which are directly related to the geographical situation (temperature, rainfall, access to the sea, presence of natural resources, etc.).

3.2. Empirical model

To try to explain the different evolution of growth in city per capita income and in city population in the 90s, we use a Multinomial Logit Model[6] (MNLM), relating cities' probability of being located in any of the distribution quartiles according to growth (both in per capita income and in population) to urban characteristics in 1990. We propose two separate models, one for the growth of city per capita income and another for city population growth, although as the explicative variables are the same, we can compare the results of both models.

The MNLM consists of transforming our dependent variable (the growth of city per capita income or of city population) into categories, which, to facilitate interpretation (and to ensure the groups are as homogeneous as possible in size), we make them coincide with the sample quartiles. This allows the results of the estimations to give us information about the probability (but not causality) of each variable affecting each category.

Thus, we rank the cities in descending order according to growth, and assign a value 1, 2, 3 or 4 according to which quartile the city's growth rate falls in, with 1 and 4 corresponding to 25% of cities with least and most growth, respectively. Figure 5 shows the box plots representing these quartiles graphically, and Table 2 shows the concrete values separating some quartiles from others. It will be seen that the distribution of income growth is much more concentrated than population growth, which at the tails shows values very far from the median[7]. To complete the information on the quartiles, Table 3 relates both distributions. The first conclusion to be extracted is that, as shown in Figure 3, there is no clear relationship between growth in city per capita income and in city population, as none of the groups is over 8%. It is worth pointing out, however, that the most numerous group, 7.61%, indicates that most of the cities with most income growth are those with least population growth.

With the MNLM we estimate a separate binary logit for each pair of categories of the dependent variable. Formally, the MNLM can be written as::

$$\ln \phi_{m|b} = \ln \frac{\Pr(K = m|\mathbf{x})}{\Pr(K = b|\mathbf{x})} = \mathbf{x}'\beta_{m|b} \qquad \text{for} \quad m = 1 \text{ a } J\,, \tag{1}$$

where b is the base category (in our case this will be category 1, the quartile containing the 25% of cities in the distribution with the lowest growth rate), $J = 4$ and $\mathbf{x}$ is the vector of the explicative variables, reflecting urban congestion costs, human capital, productive structure or geographical situation[8]. We propose studying how these

---

[6] The exogenous variables chosen are strongly correlated with each other and with growth rates, which could mean problems of endogeneity and simultaneity if we propose OLS regressions.

[7] Another advantage of this methodology is that by transforming growth rates into categories we eliminate the large variance they present (which could be the main problem when working with all population centres).

[8] The MNLM makes the assumption known as the independence of irrelevant alternatives (IIA). In this model: $\ln \frac{\Pr(K = m|\mathbf{x})}{\Pr(K = n|\mathbf{x})} = e^{\mathbf{x}'(\beta_{m|\mathrm{b}} - \beta_{n|\mathrm{b}})}$, where the odds between each pair of alternatives do not depend on other available alternatives. Thus, adding or deleting alternatives does not affect the odds between the remaining alternatives. The assumption of independence follows from the initial assumptions that the disturbances are independent and homoscedastic. We have considered one of the commonest tests developed for testing the validity of the assumption, the Small-Hsiao (1985) test, and we cannot reject the null hypothesis, that is, the odds are independent of other alternatives, indicating that the MNLM is appropriate. The model corresponding to city per capita income growth also passes the Hausman test (Hausman and McFadden, 1998), for the same null hypothesis.

explicative variables affect the odds of a city being located in one category (quartile) or another, focusing in particular on quartiles 1 and 4, representing the cities (25% of the distribution) which grew least and most, respectively. For example, if the percentage of individuals with higher level education (Percent population 18 years and over: Some college or higher degree) increases, does the probability of the city belonging to that 25% of cities with highest growth also increase?

To deal with these questions we use odds ratios (also known as factor change coefficients). Maintaining the other variables constant, the change in the odds of the outcome m against outcome n, when x_i increases by δ, equals:

$$\frac{\phi_{m|b}(\mathbf{x}, x_i + \delta)}{\phi_{n|b}(\mathbf{x}, x_i)} = e^{\beta_{i,m|n}\delta}. \qquad (2)$$

Thus, if $\delta = 1$ the odds ratio can be interpreted as follows: for each unitary change in x_i it is expected that the odds of m versus n change by a factor $e^{\beta_{i,m|n}}$, maintaining the other variables constant.

3.3. Results

This model includes many coefficients, making it difficult to interpret the effects for all pairs of categories. To simplify the analysis odds-ratio plots were developed, shown in Figures 6, 7 and 8 for different groups of variables. To analyse the marginal effect of each variable in the change in the probability of a city being in one quartile or another, Tables 4 and 5 are presented, relative to the models of growth of city per capita income and of city population respectively, showing the marginal effects for each category and the absolute average change in probability.

In an odds ratio plot, each independent variable is represented in a separate row, and the horizontal axis indicates the relative magnitude of the coefficients β associated with each outcome[9]. The numbers which appear (1, 2, 3 or 4) are the four possible outcomes, the categories (coinciding with the sample quartiles) which we previously constructed.

These graphs reveal a great deal of information (for more details, see Long and Freese, 2006). To begin, if a category is to the right of another, this indicates that increases in the independent variable make the outcome to the right more likely. Also, the distance between each pair of numbers indicates the magnitude of the effect. And when a line connects a pair of categories, this indicates a lack of statistical significance for this particular coefficient, suggesting that these two outcomes are tied together. The three graphs take outcome 1 as the base category. We are especially interested in categories (quartiles) 1 and 4, corresponding to the tails of the distribution, the 25% of cities with least and most growth, respectively.

Initial levels

Regarding the effect of initial levels of city per capita income and population, Table 4 shows that in the model corresponding to income growth the variable presenting the greatest absolute average change in probability (0.3498) is the initial city per capita income in 1989. Also, the signs of the coefficients clearly indicate that the cities with the highest initial per capita income have a greater probability of ending up in quartiles

[9] The values of the coefficients β are shown in Tables A1 and A2 in the Appendix.

1 and 2 (below median growth); i.e., the richest cities grew less in this period, relating directly to the negative relationship observed in Figure 1. In contrast, the effect of initial population on income growth is not so clear, as the most likely categories are 2 and 3, simply indicating that with a greater population in 1990 the most likely outcome is in the centre of the distribution. In the case of the model corresponding to population growth (Table 5) the effect of both variables is much less.

Congestion cost variables

In principle, the bigger the city, the greater the median travel time borne by workers. Figure 6 points to category (quartile) 4 in both models as most likely, which would indicate that indeed, where there is an increase in a unit of median travel time, the most likely outcome is that the city belongs to the 25% of cities with the highest growth, whether in per capita income or in population. In other words, increases in travel time correspond to the cities which grew most, in population or in income, although the effect is greater in the case of population growth.

The other variable is a dummy which takes value 1 if the population of the city in 1990 is more than 25,000 inhabitants, enabling us to control specifically the most populous cities of the sample. Figure 8 indicates that in the case of population growth none of the odds ratios is significant, relating directly to proportionate growth and the absence of a significant relationship between the initial population and growth (see Figure 1). On the contrary, the relationship with income growth appears to be negative: if a city had more than 25,000 inhabitants in 1990 it is most likely that it did not grow much in per capita income (the most likely outcome is quartile 1, the 25% of cities with the least income growth).

Human capital variables

The results show the opposite behavior for the two human capital variables we introduced, both in population growth and in per capita income growth. Thus, if we focus on category 4, representing the 25% of cities which grew most in population or income, Figure 6 shows that increases in the percentage of the population with the most education (some college or higher degree) have a positive impact on growth, as the most likely outcome is that the city will end up in quartile 4, while if we increase the percentage of the population with a wider concept of human capital (high school graduate or higher degree) outcome 4 becomes the least likely.

These results coincide with those of other studies analysing the influence of education in city growth. Glaeser and Shapiro (2003) also find workers have a different impact depending on their education level[10] (high school or college). Simon and Nardinelli (2002) analyse the period 1900-1990 for the USA and conclude that the cities with higher average levels of human capital grew faster over the 20th century, and Glaeser and Saiz (2003) analyse the period 1970-2000 and show that this is due to skilled cities being more economically productive (relative to less skilled cities).

Productive structure variables

In general, productive structure variables appear to have a very similar effect on the per capita income and population growth.

Figure 7 shows that per capita income and population growth depend negatively on the initial unemployment level. Thus, with an increase of 1% in the unemployment

[10] In their sample of cities the different effect is completely due to the impact of California.

rate, the most likely outcome in both models is 1, the 25% of cities with the lowest growth in per capita income or population.

For the distribution of employment by sectors, Table 4 shows that in the model corresponding to the growth of city per capita income, the sector presenting the greatest average absolute change in probability (0.0035) is the primary sector (agriculture, forestry, fishing, and mining). If we interpret this variable as a proxy for the natural physical resources available to the city (cultivable land, access to the sea, etc.) Figure 7 points to by far the most likely outcome being category (quartile) 1. In other words, higher employment in the primary sector means a higher probability that the growth rate of the city will be in the lowest quartile, the 25% of cities with the lowest income growth. This negative effect is because the primary sector usually presents constant or even decreasing returns to scale. The effect on population growth seems to be the same, with quartile 1 being the most likely outcome.

In contrast, employment in construction has a positive effect on growth, as Figure 7 shows 4 as the most likely category (quartile). The larger the percentage of employment in construction, the higher the probability that the city's growth rate belongs to the 25% of cities with the highest growth rate, both in per capita income and in population, while the average absolute change in probability is greater in the population growth model (Tables 4 and 5).

In the case of employment in manufacturing, the probability of per capita income or population growth being in the top quartile of the distribution (category 4) depends negatively on the initial percentage of employment in the manufacturing sector (Figure 7). This result coincides with that obtained by Glaeser et al. (1995) for the period 1960-1990, and its explanation is related to the depreciation of capital, suggesting that cities followed the fortunes of the industries that they were exposed to initially. However, in the 1990s the manufacturing sector seems to have lost importance, as the other sectors of activity had a greater influence on probability (Tables 4 and 5).

In services, it will be observed that only employment in finance, insurance, and real estate have a positive effect (the most likely outcome is category 4) on the growth rate of per capita income. Employment in professional services has a negative effect (the most likely outcome is category 1) and employment in wholesale and retail trade does not have a significant effect (the odds ratios are not significant). The influence of the services sector on the population growth rate seems to be much lower, as almost all the odds ratios are not significant.

The role of Geography

Until now the variables analysed seem to have a very similar effect on the growth both of per capita income and of population, as Figures 6 and 7 present a similar ordering of the categories in both models. Therefore, none of these variables is much help in explaining the divergence observed in the behavior of the distributions of per capita income and population in cities.

If we return to Tables 4 and 5, the variables presenting the greatest average absolute change in probability (after the initial levels) in both models are the dummies corresponding to geographical location, which would indicate that the location of cities in one region or another is one of the most influential factors in the growth rate of per capita income or the population of a city. Also, the odds ratio plot (Figure 8) shows a completely different order between the two models, which would indicate that the effect on the growth of per capita income and of population is different.

Remember that this dummy was used to record the influence of a series of variables for which no individual data was available for all the places, and which were directly related to the geographical situation: temperature, rainfall, access to the sea, the presence of natural resources, the availability of farming land, and even differences in economic and productive structures.

The influence of these variables has already been proven in other works. Glaeser and Shapiro (2003) find that in the 1990s people moved to warmer, dryer places. Black and Henderson (1998) conclude that the extent of city growth and mobility is related to natural advantage, or geography. Beeson et al. (2001) show that access to transportation networks, either natural (oceans) or produced (railroads) was an important source of growth over the period 1840-1990, and that weather is one of the factors promoting population growth. Access to the sea seems to influence not only the growth rate of cities, but their location itself. In Figure 4 we can see how many cities are located on the coast. And Mitchener and McLean (2003) find that some physical geography characteristics account for a high proportion of the differences in state productivity levels.

While the variable we introduced to control geography is a dummy at the regional level, the differences between cities at this geographical level are important. Table 6 shows averages for the different variables by regions, and we can observe large differences from the averages of the entire sample. Thus, the West Region is where cities grew most in population, while the cities of the Midwest Region grew most in per capita income. In contrast, the cities of the Northeast Region grew the least, both in per capita income and in population. Also, the cities present differences in their productive structures. The cities of the West Region present the highest unemployment rate, as well as a higher proportion of employment in the primary sector (agriculture, forestry, fishing, and mining), in construction and in the public administration. Employment in agriculture should indicate greater availability of land (as we can see in the map of Figure 4) in this region. The cities of the Midwest Region have a higher proportion of employment in manufacturing and wholesale and retail trade. And the cities of the Northeast Region have the inhabitants with the highest levels of human capital, and a higher proportion of employment in the services sector.

These different economic structures, and geographical characteristics, seem to be the key to explaining different behavior of per capita income and population growth in the cities in the 1990s.

4. Conclusions

This paper analyses the determinants of growth of American cities, understood as growth of the population or of per capita income, from 1990 to 2000. This empirical analysis uses data from all cities with no size restriction (our sample contains data for 21,655 cities). The results show that while population growth in cities appears to be independent of initial size (the empirical regularity known as Gibrat's law), the growth of city per capita income is negatively correlated to initial per capita income: the richest cities grew less in this period. This explains why, while the empirical distribution of city population remains stable in the decade 1990-2000, the empirical distribution of per capita income changes.

To try to explain these differentiated behaviors, we examine the relationship between urban characteristics in 1990 and city growth (both in population and in per capita income) using a Multinomial Logit Model. Apart from initial levels of population and per capita income, we used variables for congestion costs, human capital, productive structure and geographical variables. The results we obtained with our sample of all cities are similar to those of other studies which focused only on the most populous cities. Thus, we find that the probability of a city being in the 25% of cities with most growth in income or population (i.e., the probability of the growth rate of per capita income or population being in the top quartile of the distribution) depends

(1) positively on the initial percentage of inhabitants with higher educational levels (some college or higher degree), although the sign and intensity of the effect change when considering a wider concept of education (high school graduate or higher degree);

(2) negatively on initial unemployment levels, and

(3) negatively on the initial percentage of employment in the manufacturing sector, although this sector seems to have lost weight, as other economic sectors have a greater influence on probability.

Also, the location of cities on one region or another is one of the most influential factors on the growth rate of a city's per capita income or population.

References

[1] 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.

[2] Black, D., and V. Henderson, (1998). Urban evolution in the USA. Brown University Working Paper No. 98-21.

[3] Duranton, G., (2000). Urbanization, urban structure, and growth, in: J.-M. Huriot and J. F. Thisse (Eds) Economics of Cities: Theoretical Perspectives, 290–317. Cambridge: Cambridge University Press.

[4] Duranton, G., and D. Puga, (2004). Micro-foundations of urban agglomeration economies. Handbook of urban and regional economics, Vol. 4, J. V. Henderson and J. F. Thisse, eds. Amsterdam: Elsevier Science, North-Holland, Chapter 48, 2064-2117.

[5] Eeckhout, J., (2004). Gibrat's Law for (All) Cities. American Economic Review, American Economic Association, vol. 94(5), pages 1429-1451.

[6] Fujita, M. and J.-F. Thisse, (2002). Economics of Agglomeration. Cambridge: Cambridge University Press.

[7] Garicano, L., and E. Rossi-Hansberg, (2006). Organization and inequality in a knowledge economy. The Quarterly Journal of Economics, MIT Press, vol. 121(4): 1383-1435.

[8] Gibrat, R., (1931). Les inégalités économiques, París: Librairie du Recueil Sirey.

[9] Glaeser, E. L., H. D. Kallal, J. A. Scheinkman, and A. Shleifer, (1992). Growth in Cities. The Journal of Political Economy, Vol. 100(6), Centennial Issue, 1126-1152.

[10] Glaeser, E. L., and A. Saiz, (2003). The Rise of the Skilled City. Harvard Institute of Economic Research, Discussion Paper number 2025.

[11] Glaeser, E. L., J. A. Scheinkman, and A. Shleifer, (1995). Economic growth in a cross-section of cities. Journal of Monetary Economics, 36: 117-143.

[12] Glaeser, E. L. and J. Shapiro, (2003). Urban Growth in the 1990s: Is city living back? Journal of Regional Science, Vol. 43(1): 139-165.

[13] Hausman, J., and D. McFadden, (1984). Specification tests for the multinomial logit model. Econometrica 52, 1219-1240.

[14] Henderson, V., A. Kuncoro and M. Turner (1995). Industrial development in cities. Journal of Political Economy 103 (5), 1067-1090.

[15] Jacobs, J., (1969). Economy of Cities. Vintage, New York.

[16] Long, J. S., and J. Freese, (2006). Regression Models for Categorical Dependent Variables Using Stata. Stata Press Publication, 2nd ed.

[17] Loury, G. C., (1979). Market Structure and Innovation. The Quarterly Journal of Economics, MIT Press, vol. 93(3): 395-410.

[18] Mitchener, K. J., and I. W. McLean, (2003). The Productivity of US States since 1880. Journal of Economic Growth, 8: 73-114.

[19] Polèse, M., (2005). Cities and national economic growth: A reappraisal. Urban Studies, 42(8): 1429-1451.

[20] Quigley, J. M., (1998). Urban diversity and economic growth. Journal of Economic Perspectives, 12(2): 127–138.

[21] Simon, C. J., and C. Nardinelli, (2002). Human capital and the rise of American cities, 1900-1990. Regional Science and Urban Economics, 32: 59-96.

[22] Small, K. A., and C. Hsiao, (1985). Multinomial logit specification tests. International Economic Review 26, 619-627.

[23] Yamamoto, D., (2008). Scales of regional income disparities in the USA, 1955–2003. Journal of Economic Geography, Vol. 8(1): 79-103.

[24] Young, A. T., M. J. Higgins, and D. Levy, (2008). Sigma Convergence versus Beta Convergence: Evidence from U.S. County-Level Data. Journal of Money, Credit and Banking, Vol. 40(5): 1083-1093.

Tables

Table 1.- Means and standard deviations, city variables in 1990

	Mean			Stand. dev.		
Variable	All sample	Top 1000	Top 200	All sample	Top 1000	Top 200
Population Growth (ln scale), 1990-2000	0.09	0.10	0.10	0.25	0.16	0.13
Per Capita Income Growth (ln scale), 1989-1999	0.43	0.37	0.36	0.17	0.09	0.08
Congestion cost variables						
Median Travel Time to Work (in minutes)	20.41	21.27	21.16	6.04	6.35	4.13
Human capital variables						
Percent population 18 years and over: Some college or higher degree	27.30	38.23	37.14	13.00	11.80	9.82
Percent population 18 years and over: High school graduate (includes equivalency) or higher degree	53.04	59.03	56.79	11.20	9.71	8.85
Productive structure variables						
Unemployment rate	6.82	6.23	7.11	5.00	2.73	2.66
Percent employed civilian population 16 years and over:						
Agriculture, forestry, fishing, and mining	5.15	1.63	1.66	6.59	1.87	1.84
Construction	6.85	5.60	5.53	4.58	2.08	1.65
Manufacturing (durable and nondurable goods)	19.59	16.75	15.75	11.60	7.56	6.71
Wholesale and Retail trade	21.15	22.50	21.66	15.32	4.49	2.56
Finance, insurance, and real estate	5.14	7.41	7.48	7.90	3.03	2.12
Educational, health, and other professional and related services	22.79	24.64	24.73	22.59	7.99	5.74
Public administration	4.80	5.03	5.38	4.95	4.20	3.43

Source: 1990 and 2000 Census, www.census.gov

Table 2.- City Per Capita Income Growth and Population Growth (ln scale): Sample quartiles

Percentile	Population Growth	Per Capita Income Growth
25%	-0.0383	0.3378
50%	0.0471	0.4263
75%	0.1672	0.5206

Source: 1990 and 2000 Census, www.census.gov

Table 3.- Cities by sample quartiles

		Per capita income growth			
		Quartiles			
		1	2	3	4
Population Growth quartiles	1	5.23%	6.05%	6.11%	7.61%
	2	6.23%	7.25%	6.42%	5.10%
	3	6.73%	6.64%	6.28%	5.35%
	4	6.82%	5.05%	6.18%	6.95%

Source: 1990 and 2000 Census, www.census.gov

Table 4.- City Per Capita Income Growth: Marginal effects for each category and the average absolute change in the probability

	Categories (quartiles)				
Initial levels	1	2	3	4	Total average
City Population (ln scale) in 1990	-0.0250***	0.0493***	0.0286***	-0.0530***	0.0390***
Per Capita Income (ln scale) in 1989	0.6171***	0.0826***	-0.1935***	-0.5062***	0.3498***
Congestion cost variables					
Median Travel Time to Work (in minutes)	-0.0018***	-0.0008	-0.0004	0.0030***	0.0015***
Big city dummy variable (population in 1900>25,000)	0.2221***	-0.0028***	-0.0869***	-0.1324***	0.1110***
Human capital variables					
Percent population 18 years and over: Some college or higher degree	-0.0096***	-0.0038***	0.0031***	0.0103***	0.0067***
Percent population 18 years and over: High school graduate (includes equivalency) or higher degree	-0.0007	0.0043***	0.0002	-0.0037***	0.0022***
Productive structure variables					
Unemployment rate	0.0054***	0.0005***	-0.0028***	-0.0031***	0.0030***
Percent employed civilian population 16 years and over:					
Agriculture, forestry, fishing, and mining	0.0047***	0.0024***	-0.0024***	-0.0046***	0.0035***
Construction	-0.0022**	-0.0006	0.0011**	0.0017***	0.0014**
Manufacturing (durable and nondurable goods)	-0.0001	0.0014**	0.0006	-0.0018***	0.0010***
Wholesale and Retail trade	0.0001	-0.0008	0.0004	0.0004	0.0004
Finance, insurance, and real estate	-0.0047***	-0.0010**	0.0016***	0.0041***	0.0029***
Educational, health, and other professional and related services	0.0025***	0.0012**	-0.0005***	-0.0031***	0.0018***
Public administration	0.0042***	0.0022*	-0.0017***	-0.0047***	0.0032***
Geographical dummy variables					
Northeast Region	-0.0440***	0.0930***	0.0160***	-0.0649	0.0545***
Midwest Region	-0.1736***	0.0321***	0.0992***	0.0423***	0.0868***
South Region	-0.0592***	0.0293***	0.0295***	0.0005***	0.0296***

***Significant at the 1% level, **Significant at the 5% level, *Significant at the 10% level

Table 5.- City Population Growth: Marginal effects for each category and the average absolute change in the probability

	Categories (quartiles)				
Initial levels	1	2	3	4	Total average
City Population (ln scale) in 1990	-0.0285***	0.0371***	0.0188***	-0.0275	0.0280***
Per Capita Income (ln scale) in 1989	-0.1209***	0.0684***	0.0474***	0.0051***	0.0605***
Congestion cost variables					
Median Travel Time to Work (in minutes)	-0.0068***	-0.0022***	0.0020***	0.0070***	0.0045***
Big city dummy variable (population in 1900>25,000)	-0.0058	-0.0483	0.0252	0.0289	0.0271***
Human capital variables					
Percent population 18 years and over: Some college or higher degree	-0.0030***	-0.0036	0.0002***	0.0064***	0.0033***
Percent population 18 years and over: High school graduate (includes equivalency) or higher degree	0.0029***	0.0019	-0.0013***	-0.0035***	0.0024***
Productive structure variables					
Unemployment rate	0.0035***	0.0017*	-0.0034***	-0.0019***	0.0026***
Percent employed civilian population 16 years and over:					
Agriculture, forestry, fishing, and mining	0.0044***	0.0020**	-0.0014***	-0.0050***	0.0032***
Construction	-0.0038***	-0.0028	0.0031***	0.0034***	0.0033***
Manufacturing (durable and nondurable goods)	0.0003	0.0017	0.0005	-0.0025***	0.0012***
Wholesale and Retail trade	-0.0001	0.0002	-0.0002	0.0000	0.0001
Finance, insurance, and real estate	0.0000	0.0014	-0.0005	-0.0009	0.0007
Educational, health, and other professional and related services	0.0021***	0.0035	0.001	-0.0066***	0.0033***
Public administration	0.0027***	-0.0003**	-0.0012***	-0.0012***	0.0014**
Geographical dummy variables					
Northeast Region	0.3255***	0.0751***	-0.1408***	-0.2598***	0.2003***
Midwest Region	0.1547***	0.0889***	-0.0366***	-0.2070***	0.1218***
South Region	0.1823***	0.0141***	-0.0648***	-0.1316***	0.0982***

***Significant at the 1% level, **Significant at the 5% level, *Significant at the 10% level

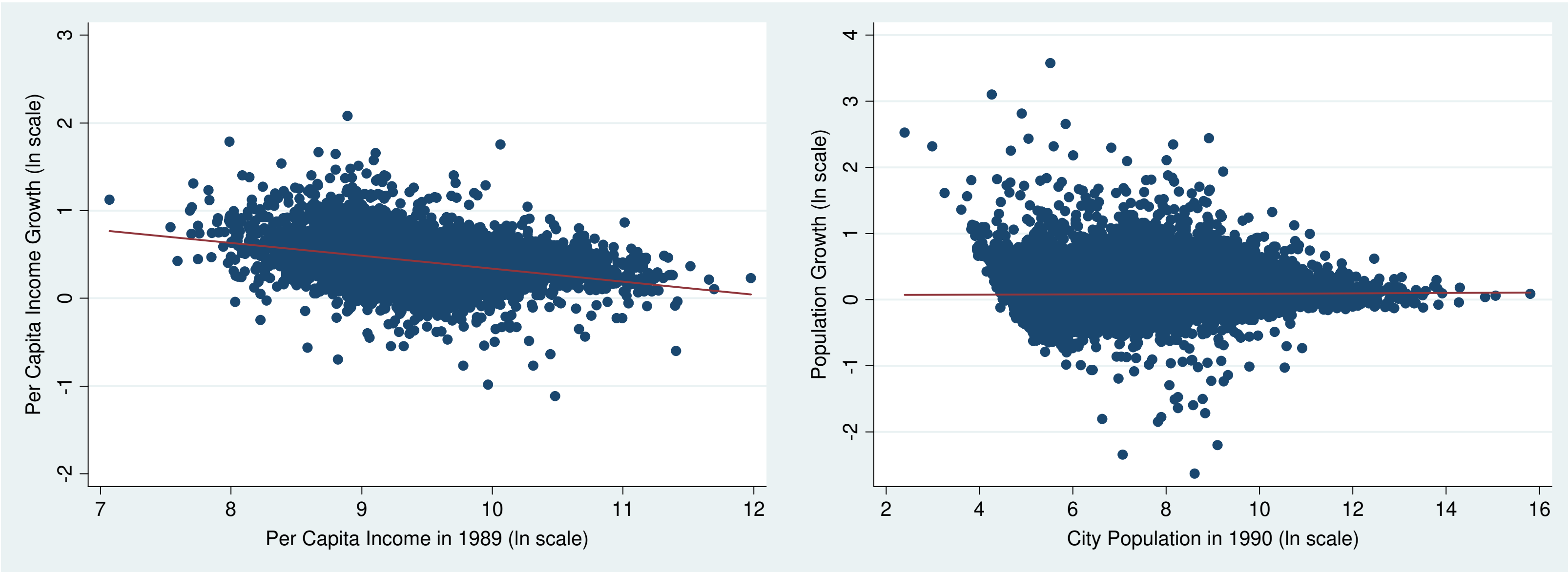
Table 6.- City variables in 1990: Means by Region

	Mean				
Variable	All sample	Northeast Region	Midwest Region	South Region	West Region
Population Growth (ln scale), 1990-2000	0.09	0.01	0.06	0.11	0.20
Per Capita Income Growth (ln scale), 1989-1999	0.43	0.38	0.46	0.44	0.41
Congestion cost variable					
Median Travel Time to Work (in minutes)	20.41	21.81	19.91	21.10	19.61
Human capital variables					
Percent population 18 years and over: Some college or higher degree	27.30	32.24	25.32	25.51	31.25
Percent population 18 years and over: High school graduate (includes equivalency) or higher degree	53.04	60.15	53.89	48.91	53.04
Productive structure variables					
Unemployment rate	6.82	5.86	6.33	7.17	8.25
Percent employed civilian population 16 years and over:					
Agriculture, forestry, fishing, and mining	5.15	1.99	4.87	5.52	8.28
Construction	6.85	6.43	6.37	7.25	7.55
Manufacturing (durable and nondurable goods)	19.59	19.28	21.81	20.38	12.58
Wholesale and Retail trade	21.15	21.49	21.89	20.42	20.65
Finance, insurance, and real estate	5.14	6.61	5.00	4.80	4.71
Educational, health, and other professional and related services	22.79	25.42	22.93	21.41	22.88
Public administration	4.80	4.50	3.81	5.26	6.51
Sample size	21655	3276	7922	7278	3179

Source: 1990 and 2000 Census, www.census.gov

Figures

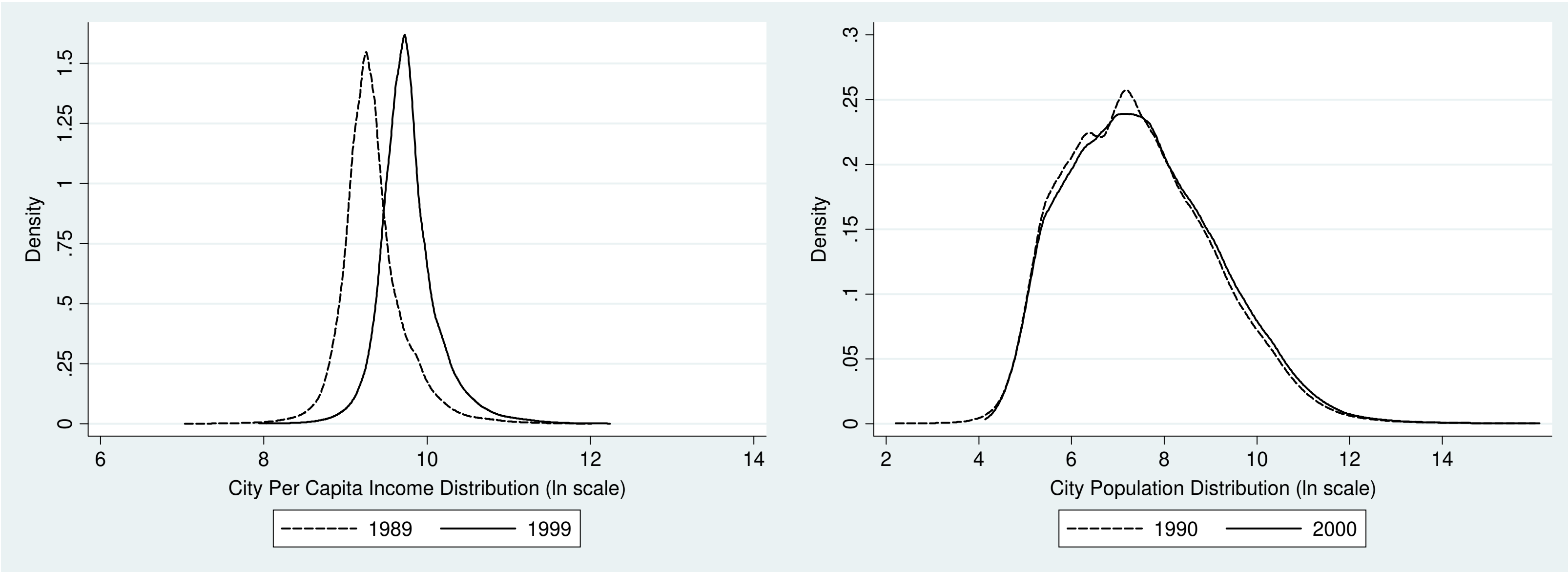
Figure 1.- Scatter Plots of City Growth (ln scale) against initial level



Note: Line fitted as $(\ln y_{it} - \ln y_{it-1}) = \alpha + \beta \ln y_{it-1}$.

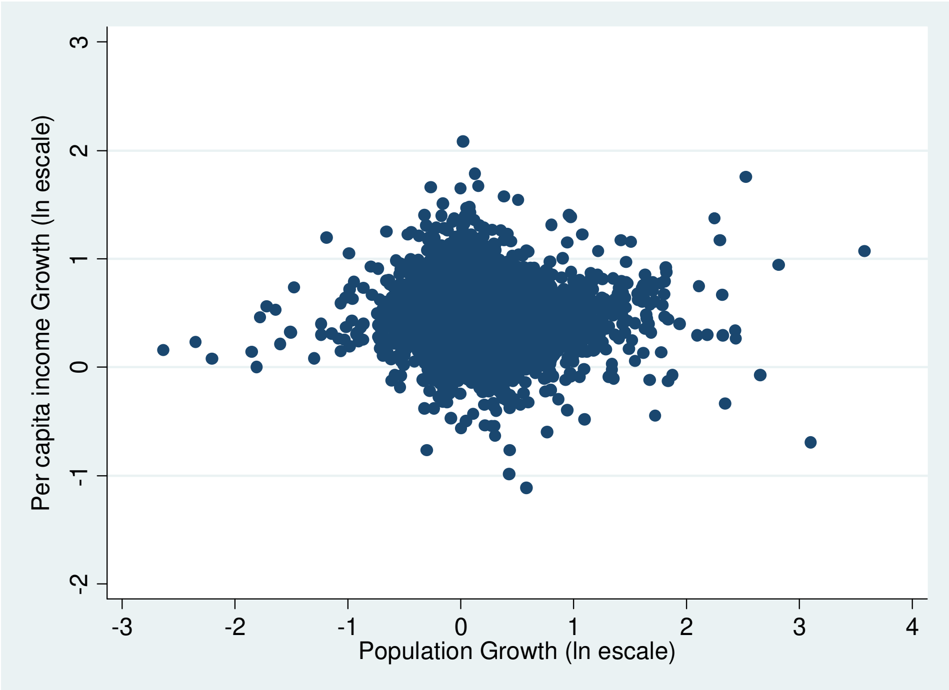
Source: 1990 and 2000 Census, www.census.gov

Figure 2.- Kernel density estimation (ln scale) of City Per Capita Income and City Population Distributions



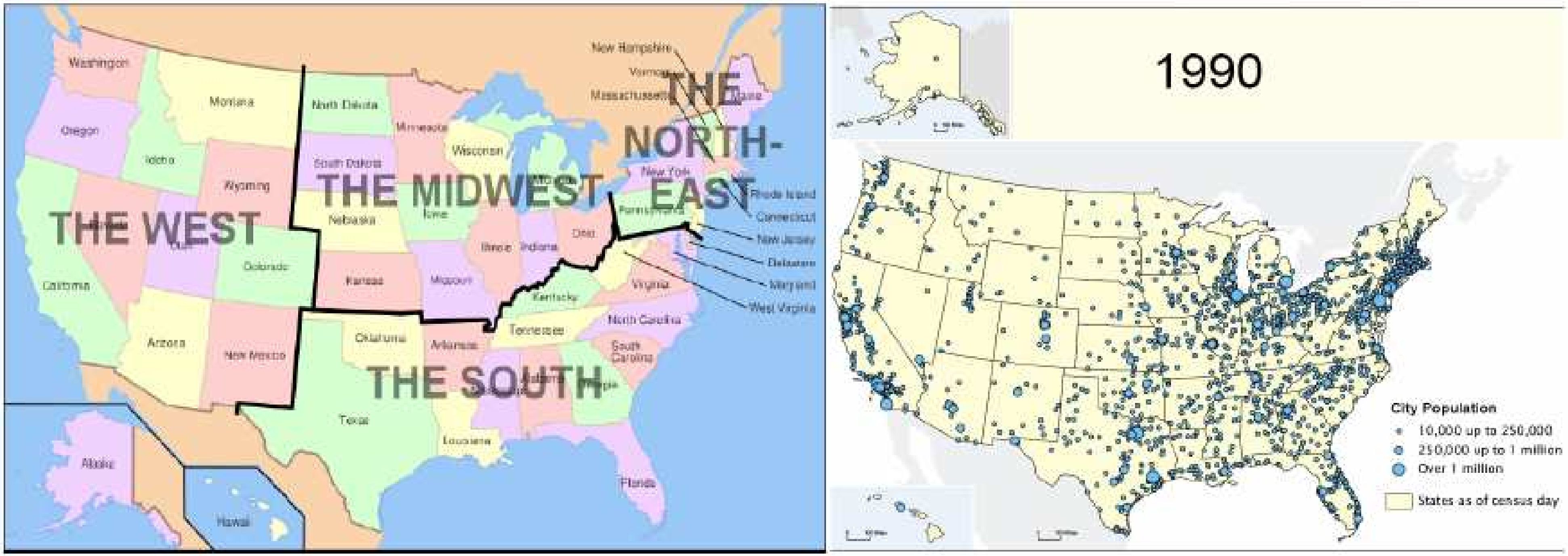
Source: 1990 and 2000 Census, www.census.gov

Figure 3.- Scatter Plot of City Per Capita Income Growth (ln scale) against City Population Growth (ln scale)



Source: 1990 and 2000 Census, www.census.gov

Figure 4.- Cities by Region

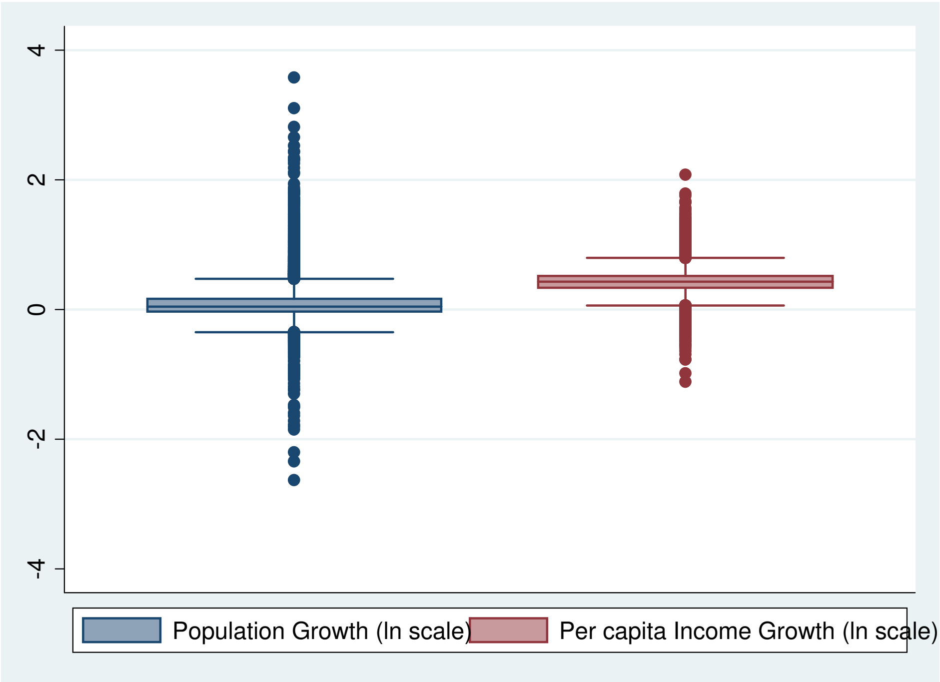


Sources:

Wikimedia Commons: http://commons.wikimedia.org/wiki/File:Map_of_USA_showing_regions.png

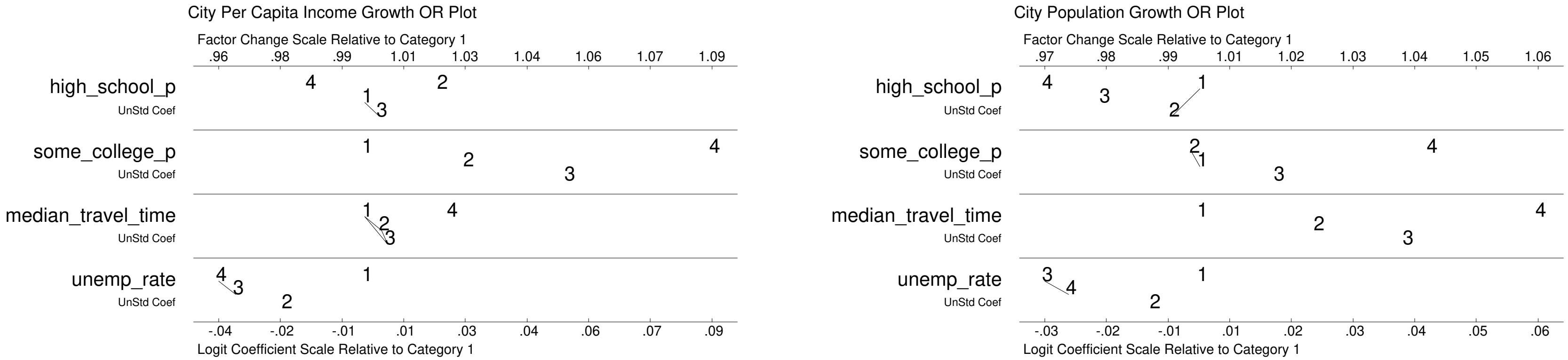
US Census Bureau: http://www.census.gov/dmd/www/map_1990.pdf

Figure 5.- Box Plots of City Per Capita Income Growth (ln scale) and City Population Growth (ln scale)



Source: 1990 and 2000 Census, www.census.gov

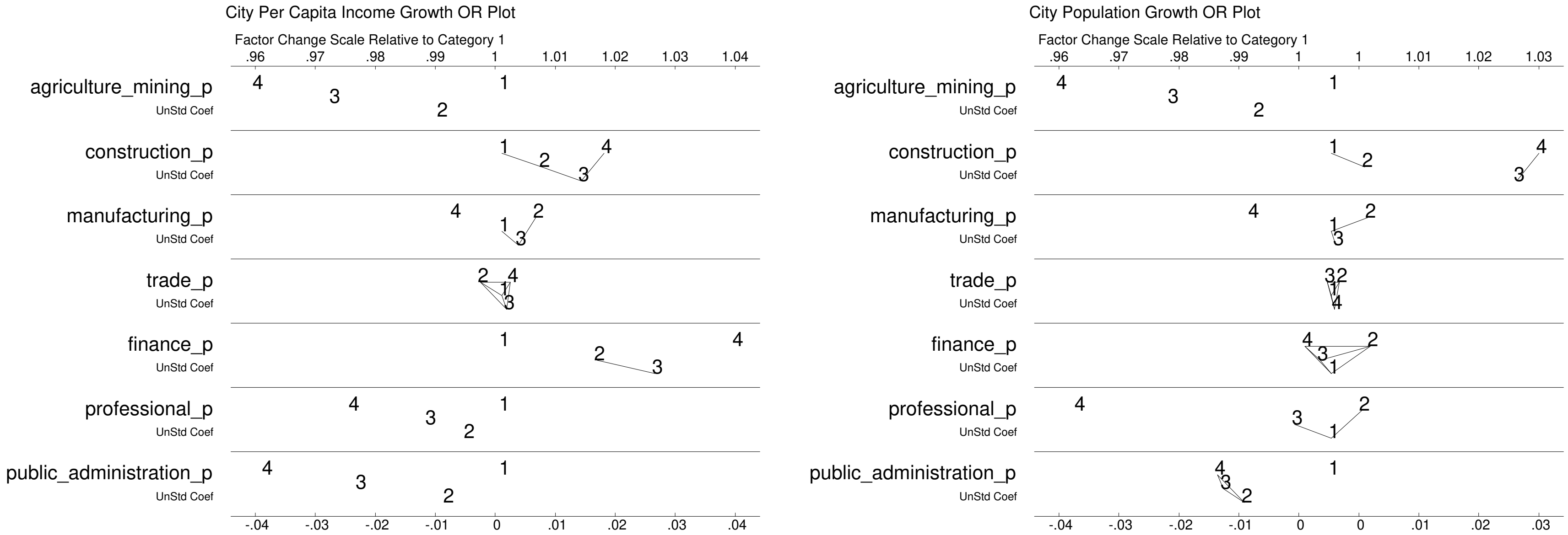
Figure 6.- Odds ratio plots of human capital variables, median travel time and unemployment rate



Note:

High_school_p: Percent population 18 years and over: High school graduate (includes equivalency) or higher degree. Year 1990.
Some_college_p: Percent population 18 years and over: Some college or higher degree. Year 1990.
Median_travel_time: Workers 16 years and over who did not work at home: Median travel time to work (in minutes). Year 1990.
Unemp_rate: Unemployment rate. Universe: Unemployed persons 16 years and over. Year 1990.

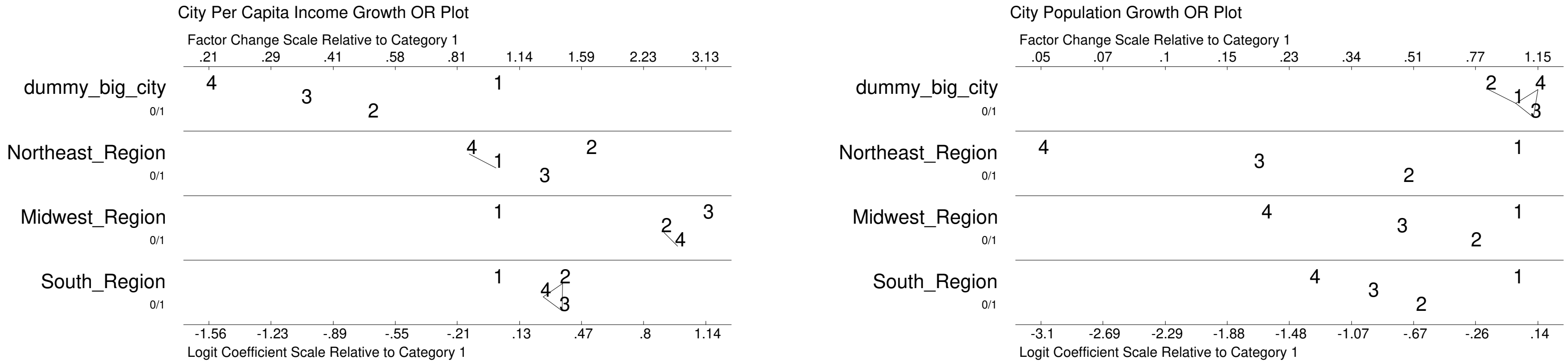
Figure 7.- Odds ratio plots of productive structure variables



Note:

Percent employed civilian population 16 years and over:
- Agriculture_mining_p: Agriculture, forestry, fishing, and mining.
- Construction_p: Construction.
- Manufacturing:_p Manufacturing (durable and nondurable goods).
- Trade_p: Wholesale and Retail trade.
- Finance_p: Finance, insurance, and real estate.
- Professional_p: Educational, health, and other professional and related services.
- Public_administration_p: Public administration.

Figure 8.- Odds ratio plots of geographical dummy variables



Note:

Dummy_big_city: Big city dummy variable (population in 1900>25,000).
Northeast_Region: The Northeast Region includes the following states: Maine, New Hampshire, Vermont, Massachusetts, Rhode Island, Connecticut, New York, New Jersey, and Pennsylvania.
Midwest_Region: The Midwest Region includes the following states: Ohio, Indiana, Illinois, Michigan, Wisconsin, Minnesota, Iowa, Missouri, North Dakota, South Dakota, Nebraska, and Kansas.
South_Region: The South Region includes the following states: Delaware, Maryland, District of Columbia, Virginia, West Virginia, North Carolina, South Carolina, Georgia, Florida, Kentucky, Tennessee, Alabama, Mississippi, Arkansas, Louisiana, Oklahoma, and Texas.

Appendix: Estimated Multinomial Logit coefficients

Table A1.- City Per Capita Income Growth: Multinomial Logit coefficients relative to Category (quartile) 1

	Categories (quartiles)		
Initial levels	2	3	4
City Population (ln scale) in 1990	0.2873***	0.2070***	-0.1440***
Per Capita Income (ln scale) in 1989	-2.2360***	-3.2478***	-4.9018***
Congestion cost variables			
Median Travel Time to Work (in minutes)	0.0042	0.0058	0.0214***
Big city dummy variable (population in 1900>25,000)	-0.6839***	-1.0468***	-1.5649***
Human capital variables			
Percent population 18 years and over: Some college or higher degree	0.0255***	0.0509***	0.0873***
Percent population 18 years and over: High school graduate (includes equivalency) or higher degree	0.0189***	0.0036	-0.0142***
Productive structure variables			
Unemployment rate	-0.0202***	-0.0325***	-0.0368***
Percent employed civilian population 16 years and over:			
Agriculture, forestry, fishing, and mining	-0.0103***	-0.0281***	-0.0409***
Construction	0.0067	0.0132**	0.0170***
Manufacturing (durable and nondurable goods)	0.0056**	0.0028	-0.0081***
Wholesale and Retail trade	-0.0036	0.0009	0.0014
Finance, insurance, and real estate	0.0158**	0.0254***	0.0387***
Educational, health, and other professional and related services	-0.0058**	-0.0122***	-0.0249***
Public administration	-0.0093*	-0.0238***	-0.0393***
Geographical dummy variables			
Northeast Region	0.5063***	0.2518***	-0.1453
Midwest Region	0.9134***	1.1424***	0.9886***
South Region	0.3624***	0.3601***	0.2567***

1 is the base outcome. ***Significant at the 1% level, **Significant at the 5% level, *Significant at the 10% level

Table A2.- City Population Growth: Multinomial Logit coefficients relative to Category (quartile) 1

	Categories (quartiles)		
Initial levels	2	3	4
City Population (ln scale) in 1990	0.2629***	0.1890***	0.0007
Per Capita Income (ln scale) in 1989	0.7728***	0.6831***	0.5308***
Congestion cost variables			
Median Travel Time to Work (in minutes)	0.0204***	0.0360***	0.0592***
Big city dummy variable (population in 1900>25,000)	-0.1788	0.1138	0.1434
Human capital variables			
Percent population 18 years and over: Some college or higher degree	-0.0014	0.0134***	0.0402***
Percent population 18 years and over: High school graduate (includes equivalency) or higher degree	-0.0049	-0.0171***	-0.0272***
Productive structure variables			
Unemployment rate	-0.0083*	-0.0272***	-0.0230***
Percent employed civilian population 16 years and over:			
Agriculture, forestry, fishing, and mining	-0.0111**	-0.0237***	-0.0402***
Construction	0.0050	0.0274***	0.0306***
Manufacturing (durable and nondurable goods)	0.0054	0.0007	-0.0119***
Wholesale and Retail trade	0.0012	-0.0006	0.0005
Finance, insurance, and real estate	0.0057	-0.0017	-0.0038
Educational, health, and other professional and related services	0.0045	-0.0054	-0.0375***
Public administration	-0.0128**	-0.0159***	-0.0168***
Geographical dummy variables			
Northeast Region	-0.7153***	-1.6892***	-3.0963***
Midwest Region	-0.2752***	-0.7572***	-1.6395***
South Region	-0.6339***	-0.9430***	-1.3269***

1 is the base outcome. ***Significant at the 1% level, **Significant at the 5% level, *Significant at the 10% level