Regional Science and Urban Economics 31 (2001) 701–731
www.elsevier.com/locate/econbase

Spatial interactions among U.S. cities: 1900–1990

Linda Harris Dobkins[a,*], Yannis M. Ioannides[b]

[a]*Department of Economics, Emory and Henry College, Emory, VA 24327, USA*
[b]*Department of Economics, Tufts University, Medford, MA 02155, USA*

Received 8 June 2000; received in revised form 6 November 2000; accepted 19 December 2000

Abstract

We test implications of economic geography by exploring spatial interactions among U.S. cities. We use a data set consisting of 1900–1990 metro area populations, and spatial measures including distance from the nearest larger city in a higher-tier, adjacency, and location within U.S. regions. We also date cities from their time of settlement. We find that among cities which enter the system, larger cities are more likely to locate near other cities. Moreover, older cities are more likely to have neighbors. Distance from the nearest higher-tier city is not always a significant determinant of size and growth. We find no evidence of persistent non-linear effects on urban growth of either size or distance, although distance is important for city size for some years.

Keywords: Spatial interactions; Urban growth; New economic geography; Neighbors

JEL classification: R00; C14

1. Introduction

Why do cities locate where they do? What does location, growth and the age of a city tell us about its economic relationship to other cities in a system? To answer such questions, spatial economics has recently returned to the mid-twentieth

*Corresponding author.
E-mail addresses: lhdobkin@ehc.edu (L. Harris Dobkins), yannis.ioannides@tufts.edu (Y.M. Ioannides).

0166-0462/01/$ – see front matter
PII: S0166-0462(01)00067-9

century works of geographers such as Christaller (1933) and Pred (1966), and furthered their basic ideas with the tools of modern economics, including models of imperfect competition under increasing returns.

This paper constructs an empirical platform within which one may examine broad predictions of the economic geography of urban systems. We do so by focusing on the United States' system of cities from 1900 to 1990. Dobkins and Ioannides (2000a) explores a data set for U.S. cities spanning the century, looking at patterns of city growth and the distribution of city sizes as new cities enter the distribution, which is important for the United States. We augment those data by means of additional geographical and other information to examine *spatial* aspects of the U.S. system. We consider the presence of neighboring cities, regional influence, distance between cities, and the time since first settlement, 'age,' of cities in the system.

There are three distinct but intertwined explanations for the location and sizes of cities within an urban system: Christaller's central place theory, Pred's agglomerative forces, and initial advantage. These explanations invoke a variety of theoretical devices, from the mercantilist theory of geographers to recent research in 'new economic geography', which rests on monopolistic competition models with increasing returns. Central-place theory describes the relationship of cities of different sizes, but says nothing about their location. We combine these three strands in order to illuminate the relationship between central place theory, location and age. We make novel use of a distance variable related to functional urban tiers and of a variable denoting the presence of neighbors as measures of spatial proximity, and of an age variable as a proxy for initial advantage, to examine these three explanations.

The U.S. system is characterized by the entry of new cities, which we define as new settlements which appear anew or grow above a minimum size of 50 000 to qualify for admission to our data. Some new cities cluster near existing ones, thus becoming neighbors of older cities. We look at the growth of existing cities and the location of new cities as a function of proximity to other cities and relate these results to Herbert Simon's theory of random urban growth (Simon, 1955; Krugman, 1996a). These, and other findings, fill a needed gap in informing and illuminating the wealth of theories, new and old, on city growth, location and formation.

Section 2 outlines the theoretical points that guide the questions we ask. Section 3 describes our data set. Section 4 details the empirical questions we consider. Section 5 describes our stylized findings, regression results, and answers to the questions posed, and Section 6 concludes.

2. Literature review

Pred (1966) identifies three factors that may explain the size and location of cities, relative to their hinterlands in a given spatial system. Those factors are: first, agglomerative forces, including scale economies; second, central place considerations a la Christaller (1933); and third, initial advantage. Current approaches to spatial economics involve various combinations of those three factors.

The pioneering work of J. Vernon Henderson centers on the play of agglomerative forces. In a series of works, (Henderson, 1974, 1987, 1988), this approach explores urban space (but ignores national space) and is premised on Marshallian localized external effects; external, that is, to the industries and individuals in a city but internal to the city economy. City sizes vary by the types of cities, and new cities enter the economy in proportion to the growth of population in the system (Henderson and Ioannides 1981).[1] The question of where new cities locate when they 'appear' cannot, however, be addressed by the Henderson model.

Simon (1955) proposes a statistical law by which cities may locate in space alone or next to another city, as an urban system evolves. Krugman (1996a) simplifies Simon's theory this way: the likelihood that a new city locates next to an existing city is an increasing function of the size of the existing city. As in Fujita and Mori (1996) and Marshall (1989) (below), this links initial advantage to the age of settlement, as the system evolves.

Fujita and Mori (1996, 1997) and Fujita et al. (1999a) bring some of those insights to bear on the question of emergence of new cities. In the evolutionary model of Fujita and Mori (1997), the economy starts with a single city, in which a variety of manufacturing goods are produced, and a hinterland which produces agricultural goods. In keeping with Pred's delineation, we might say that a central place exists because of agglomeration benefits. However, Christaller's central place system does not evolve. Fujita and Mori's system allows for an expanding population, which means that the hinterland's population starts to look attractive to entrepreneurs who see market opportunities, and a new city is born. The economy's population reaches a critical value at which the monocentric equilibrium becomes unstable; then a catastrophic bifurcation occurs and a duocentric system emerges. Fujita and Mori conclude that new cities are created as the population increases, and that the pattern of these new cities approaches a regular, central-place type system.

And how does Pred's initial advantage fit into this story? Krugman (1991) and Arthur (1994) describe evolutionary patterns in which history matters. Fujita and Mori (1996) emphasize the importance of port cities, thus putting a comparative advantage argument in an evolutionary setting. In their analysis, cities may have been important because of water access, or any other initial advantage in the past. However, those cities almost always remain important because of the 'lock-in effect' generated by the kind of circular causation discussed above.

[1]Part 2 of that paper identifies difficulties associated with the introduction of a second city, but fails to model it as a bifurcation of the dynamic system, which is the key contribution of Fujita et al. (1999a). The assumption of product differentiation in Ioannides (1994) confers an element typically associated with endogenous growth models. As national population grows, so does the number of cities, and through that, product variety.

Another look at the relationship between central place theory and the evolution of the system is laid out by John Marshall (1989). Noting that Christaller's central place system is a static theory, indeed better suited to the geography of Europe, Marshall recommends Vance's (1970) mercantile theory to explain the urban systems of North America. In the mercantile model, the urban system evolves from 'points of attachment' on the coast inland, along trade routes, to the 'depots' needed by the wholesalers. According to Marshall, this historical view of the development of the North American urban system must be reconciled with central place and manufacturing considerations in order to explain the current system. It is a view that in part conflicts with the Fujita–Mori story: population growth in the hinterland comes first in their story and then a new city emerges.

Marshall's drawings of the succession of depot-centers westward fits well with Fujita and Mori's linear development of successive population centers. Furthermore, Marshall asserts that the mercantile model can help provide the missing dynamic element for central place theory in that the mercantile process of settlement ' 'fixes' the spatial pattern of towns destined to become the leading central places of the fully developed system'. (Marshall, 1989, p. 284) Marshall ascribes to Christaller the notion that the highest-order central places would be the oldest, starting small perhaps but always being the leading population center in a region.

Finally, we consider the potential role of threshold effects, particularly as developed by Fujita et al. (1999a) and by Fujita et al. (1999b), associated with the *spatial interaction* of cities and their hinterlands. They distinguish the case of a mature urban system and its hinterland, from that of a growing system, but explore in full detail examples set in simple linear topology. A most noteworthy feature of these works is their emphasis on national space (although occasionally by neglecting urban space), by means of an interplay of scale economies and transportation costs. Scale economies matter because manufacturing firms want to locate near demand, but demand implies the presence of the very people who work in the manufacturing plants. Thus manufacturing concentrates in 'cores', while the 'periphery' is the agricultural hinterland. Concentration of the population into discrete cores requires the presence of increasing returns in manufacturing production as well as two other considerations which may change over time: a fairly high percentage of income spent on non-agricultural goods, and low enough transportation costs to get some goods to and from the periphery, or hinterland. Suddenly, the story takes on a spatial aspect: where do the core areas form? And, at what critical combinations of decreasing transportation costs and rising percentages of manufactured-goods spending do the centers emerge? Fujita et al. (1999b) offer precise answers to these questions.

This result, of course, is reminiscent of the central place theory of Christaller (1933), a system in which the largest central places provide the greatest number of products and services. Nearby towns are much smaller in a central place system, because they provide only the most basic goods and services. These smaller neighbors in the Christaller system fall into what Krugman calls a city's 'agglomeration shadow.' While we certainly lack the data to test very specific implications of the theory in Fujita, Krugman and Venables, we do note that: one, non-linearities may show up in the dynamics of the system; and two, it is implicit in this analysis that once a city has entered, the presence of neighbors influences its subsequent growth. [2] These theories, then, spawn a number of questions. Not all of the questions are new, and only some of the answers may be addressed by our data and the techniques we employ. With the noteworthy exception of Hanson (1998), whose work arguably provides the only direct structural test of predictions of new economic geography, there essentially exists little research on spatial interactions. We acknowledge broadly related work by Black and Henderson (1999), which does not however address spatial interactions as defined in the present paper.[3] Ioannides and Overman (2000b) address some of the same questions by means of the sort of non-parametric empirical techniques emphasized by Quah (1996). We proceed next to describe the data.

3. Data

3.1. City definitions and date of settlement

There are, of course, a variety of ways to define cities. In this study we primarily use contemporaneous Census Bureau definitions of metropolitan areas, with adaptations for data availability. From 1900 to 1950, we have metropolitan areas defined by the 1950 census. That is, for years previous to 1950, we use Bogue's (1953) reconstructions of what populations would have been in each metropolitan area in each year if the cities had been defined spatially as they were in 1950. For each decennial year from 1950 to 1980, we use the metropolitan area definitions that were in effect for those years. Between 1980 and 1990, the Census Bureau redefined metropolitan areas in such a way that the largest U.S. cities would seem to have taken a huge jump in size, and several major cities would have been lost. While this might be appropriate for some uses of the data, we want to be able to track cities as neighbors. Therefore, we reconstructed the metro areas for 1990, based on the 1980 definitions, much as Bogue did earlier. We believe that this gives us the most consistent definitions of U.S. cities (metropolitan areas) that we are likely to find.

An alternative method would be to have a standard geographical definition of a given city and use that definition for all years, (as Bogue essentially did in 1953

[2] We note that this discussion pertains to spatial proximity. Yet, it is obvious that cities that are physically distant from one another do interact through markets for their similar products. We abstract in this paper from such considerations, as we wish to emphasize the impact of physical proximity.

[3] The first version of the present paper was completed prior to Black and Henderson (1999).

for the previous half of the century). That data set has not been constructed.[4] We believe it would actually be less useful because a set geographic boundary ignores changing technology. The use of contemporaneous definitions allows us to take advantage of the Census Bureau's knowledge of commuting patterns, which reflect technology among many other economic interactions.

The method also raises a question as to which cities, as defined or reconstructed, should be included. In the years from 1950 to 1980, we use the Census Bureau's listing of metropolitan areas. Although the wording of the definitions of metropolitan areas has changed slightly over the years, the number 50 000 is the minimum requirement for a core area within the metropolitan area. Therefore, we used 50 000 as the cutoff for including metropolitan areas as defined by Bogue prior to 1950. Consequently we have a changing number of cities over time, from 112 in 1900 to 334 in 1990. While it is often difficult to deal with an increasing number of cities econometrically, we think that this is a key aspect of the U.S. system of cities.

For a more basic analysis on the number of cities and national population, we use a second and simpler data set. We simply use the number of cities in the U.S. from 1790 to 1990. Two points must be made here. First of all, these are 'urban territories' as defined by the government. This definition has changed slightly over time, but basically points to an incorporated area, with some exceptions being made for densely populated areas that were not legally incorporated. This definition of a city yields more entities than the metropolitan area designation in use in this century, so that we reach a maximum of 555 'urban territories' with population over 50 000 by 1990. The second point is that this data set is affected by an ambiguous cutoff size in the early years. In 1790, no city had reached the size of 50 000, but New York City was much larger than other urban places. By 1800, Baltimore and Philadelphia were approaching New York City's 1790 size, and Boston did so in 1810. Essentially, we use a sliding scale for the inclusion of cities until 1860, at which time the 50,000 cutoff seems appropriate given the size distribution. A listing of the number of cities in each census year is shown in Table 1. While there might be some argument about which cities to include at which date, there are not enough cities involved to make a significant difference in the early years.

Initial advantage is a challenging concept for operationalization. We use the date of settlement for each city, reasoning that it reflects geography and randomness. One would suppose that the east to west settlement of the country would determine settlement dates. Yet, we find early settlement dates in the west and late ones along the east coast. The notion of settlement does not reflect incorporation as a city, but it is a historical reference to the earliest indication of the use of location. This is

[4] In a conversation with one of the authors a Census Bureau official noted that such a project had been considered, but never undertaken.

Table 1
US population growth and number of cities[a]

Year	Number of cities	Cities included	Percentage change in U.S. population (%)
1790	1	New York City	
1800	3	plus Philadelphia and Baltimore	35
1810	4	plus Boston	36
1820	5	plus New Orleans	33
1830	6	plus Cincinnati	34
1840	8	plus Pittsburgh and Louisville	33
1850	11	plus Newark, Washington D.C. and St. Louis	36
1860	16	all cities over 50 000	35
1870	25		27
1880	35		26
1890	58		25
1900	78		21
1910	109		21
1920	144		15
1930	191		16
1940	199		7
1950	232		15
1960	333		19
1970	396		13
1980	463		11
1990	555		10

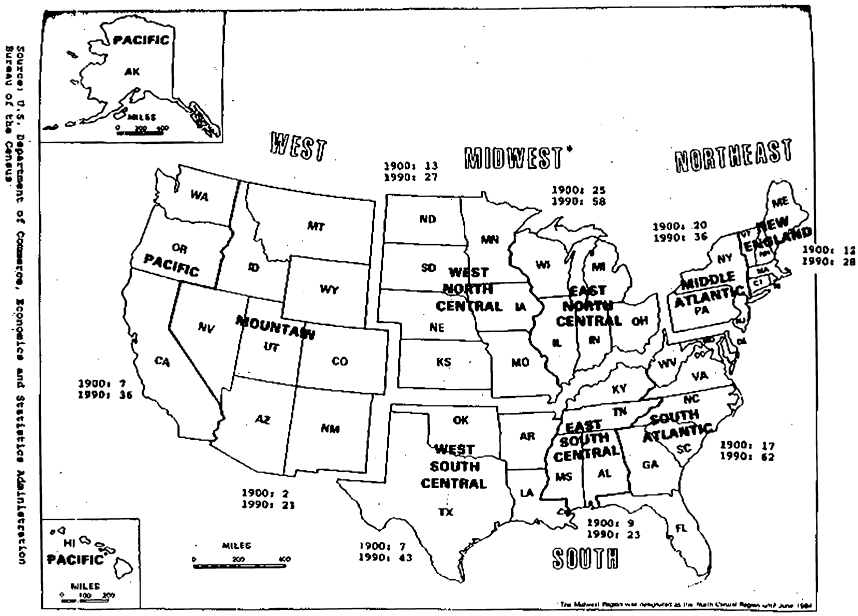
[a] As noted in the body of the paper, we count cities early in the country's history based on relative size. New York City was very close to 50 000 population in 1790, and other cities are added as they come close to New York City's size. By 1860, we include all cities with population above 50,000. The sources are *Historical Statistics of the United States from Colonial Times to the Present, Vol. 1 and 2, and Statistical Abstract of the United States, 1993, 1996.*

our own newly defined variable, and was compiled by sifting through historical records, and comparing the earliest identifications of use.

In a number of cases, the dates are references to military forts. We use those dates because often the site of the fort determined the site of the city that grew up nearby. The earliest date is that of Jacksonville, Florida, which includes the St. Augustine area and dates to 1564; the latest is Richland, Washington, originally the site of a nuclear facility settled in 1944. It is an interesting statistic in and of itself to see how age of settlement correlates with city size. If older age (a better site) makes a city larger, which indicates importance in the system, then we would expect the 'date' variable to have a negative effect, of course, this variable also has implications for Marshall's prediction, as noted above.

3.2. Spatial measures

We consider the importance of space in several ways. First, we control for the



The number of cities in our sample from each region are given for 1900 and 1990.

Fig. 1. U.S. regions and their cities, 1900–1990.

particular regions and Census divisions within the U.S. where cities are located and new urban centers appear. Second, we account for distance, where we use a particular measure that combines physical and economic distance. Third, we account for whether or not cities are adjacent to one another.

As noted above, spatial expansion over geographical regions is an important feature of the U.S. experience. The Census Bureau divides the country into nine regions (which we recombine into five regions, when necessary). See Fig. 1 for the spatial distribution of the cities in our data. The east-west movement that is at the heart of mercantile theory would predict a steady increase of cities in the Midwest, Mountain and Pacific Coast areas. A loose interpretation of Fujita and Mori (1997) would suggest that the 'frontier cities' might initially be larger than cities just to the east; those cities would then, in turn, bear the competition of a newer frontier city, which would be larger.[5] We look at regions to see if such a pattern exists, although a cursory knowledge of recent trends would lead us to believe that the

[5] This is a loose interpretation of Fujita and Mori's model because they see expansion in both directions along a line, not just East to West.

movement pattern is from the Northeast and Middle Atlantic to first the West, and then the Southeast during the century.

In order to examine central place considerations, we have created a variable referred to below as *distance*. It is a measurement of driving distances from each city in the sample to the nearest larger city in a higher tier. (Means for each census year are given in Table 2, column 7. We note that the distances, as defined here, decrease slightly over the century, particularly from the earliest years, as one might expect given the increase in the number of cities). We construct the tiers by grouping the cities in a given decade by function. These classifications should change over time; therefore, we used historical sources that rank cities by function. These 'central place' considerations identify cities for the top tiers that offer diverse economic functions, and are not based solely on population. Some cities change tiers over the years, as in the case of Detroit, New Orleans and Miami, to give a few examples. Detroit is a tier-one city in the early years of the century, but falls to a tier-four city as it becomes increasingly identified as a manufacturing center. New Orleans, never a particularly large urban center, is a tier-one city in the early years and falls only to tier-two as its importance as a port wanes. Miami, which does not even enter the distribution until 1930, rises swiftly through the tiers thereafter.

The cities are divided into four tiers. For the early and middle years, our categories are based on Pickard (1959). For later years, we use a classification from Knox (1994) [itself based on Noyelle and Stanback, 1984] that identifies 'nodal' centers (e.g. Houston) for our top tier; regional nodal centers (e.g. Dallas)

Table 2
Descriptive statistics: decennial data 1900–1990[a]

1	2	3	4	5	6	7
Year	U.S. Pop. (000)	U.S. Pop.: urban (000)	Mean size	Number	GNP billion $	Distance miles
1900	75 995	29 215	260 851	112	71.2	221
1910	91 972	39 944	287 367	139	107.5	215
1920	105 711	50 444	338 549	149	135.9	178
1930	122 775	64 586	411 378	157	184.8	178
1940	131 669	70 149	438 434	160	229.2	178
1950	150 697	85 572	528 223	162	354.9	178
1960	179 323	112 593	536 158	210	497.0	176
1970	203 302	139 419	573 742	243	747.6	175
1980	226 542	169 429	526 177	322	963.0	169
1990	248 710	192 512	576 383	334	1277.8	169

[a] All figures are taken from *Historical Statistics of the United States from Colonial Times to 1970*, Volumes 1 and 2, and *Statistical Abstract of the United States, 1993*. Column 6: GNP adjusted by the implicit price deflator, constructed from sources above; 1958 = 100.

for the second tier; and subregional nodal centers (e.g. Memphis) for the third tier. All other cities fall into the fourth tier.[6]

The distance variable, as noted above, is the driving distance (as published by Rand McNally) to the nearest city in a higher tier. For cities in the top tier, the relevant distance variable is to the nearest city in that tier. (For Honolulu and Anchorage, we simply use arbitrarily large numbers, because driving distance is irrelevant). Obviously, distance varies with time because cities shift into different tiers over the century. Thus, the distance variable says something about the shifting spatial configuration of the U.S. urban system.

Central place theory would predict that larger cities are farther away from each other, so that the distance variable ought to correlate positively with population, if central place theory is meaningful at all. Because the top tier cities are obviously large, it would seem that the variable is 'stacked' in this sense, but the location of cities is still a powerful influence, reflecting the usual criticism of central place theory that reality is very different from the featureless plane. Furthermore, the agglomeration effects as evidenced by the presence of neighbors comes into play, as we will see below.

Another measure of proximity that we employ is whether or not two cities are adjacent. We consider cities to be adjacent if the Census Bureau has ever grouped them together in various extended, but pertinent, definitions. While the Census Bureau definitions clearly do not answer all questions about proximity, (nor were they meant to), they do provide a ready-made proxy. We are again relying on the Census Bureau's knowledge of commuting patterns in the cities it considers 'consolidated.' For example, the Census Bureau's consolidated metropolitan area for Los Angeles includes San Bernardino/Riverside, Anaheim and Oxnard. We consider these as three separate cities in our sample. (Other cities that may be as close as Los Angeles and San Bernardino may not be considered neighbors by the Census Bureau, presumably because of Census Bureau information, such as commuting patterns. Therefore that information is reflected in our data set). When these cities enter the data set on their own, they are denoted as neighbors to Los Angeles and to each other. The average number of neighbors (of cities with neighbors) fluctuates around 1.00, until the 1960's, after which time it starts varying between a low of 1.383 and a high of 2.111.

Neighbors 'happen' in several ways. In some cases, cities simply grew up near each geographically, as in the case of Dallas and Fort Worth. In other cases, neighboring cities may have been a part of a city's hinterland and simply grow with the core city until they reach a population threshold and enter the distribution. An example of this is Rock Hill, SC, which enters in 1980 as a neighbor to Charlotte, NC. In other cases, neighbors enter and in so doing separate from an existing city. The most dramatic case is Nassau and Suffolk counties in New York

[6] Details of the cities included in the four tiers in each decade are available in Dobkins and Ioannides (2000b).

state, which enter in 1980 at more than two million population, lowering the population of New York City, of which they were previously a part, by that amount. See Appendix A for additional details on neighbors.

4. Empirical questions

Next we proceed by looking first at descriptive aspects of the data and then turning to econometric analyses. A first question is whether the predictions of Fujita and Mori regarding increasing numbers of cities as a function of increasing population are borne out. Several other models have very different premises. Eaton and Eckstein (1997) develop a model in which cities may grow in parallel – that is the same group of cities simply become larger while their distribution remains invariant as the country's population grows. They propose this as a model for at least two countries: Japan from 1925 to 1985 and France from 1876 to 1990. The U.S. has expanded, of course, not only in terms of population but of the number of cities – a fact in agreement with Fujita and Mori, op. cit. – and its geographical area as well. Furthermore, Henderson and Ioannides (1981) predicts a proportional growth of the number of cities with respect to population. We might expect that the prediction would not be borne out due to the aspatial nature of their model. We look at this issue with a simple data set relating the number of cities and population growth from 1790 to 1990 for the U.S.

We might also ask if the creation of new cities has followed an East to West pattern, as suggested by John Marshall, with Christaller, Pred, and Fujita and Mori (1996) in broad agreement. This would support a significant role for initial advantage, proxied in our data set by date of settlement. Also related is whether or not the highest order/function cities are the oldest. These questions are related to central place theory's fundamental grid-pattern, which is difficult to test. But we can ask if the lower function cities are smaller, as Christaller would have predicted. We can also examine the relationships of distance and function.

Along with Christaller, Simon and Krugman, we might ask where new cities locate specifically. Do they locate near other cities or do they locate in relative isolation. Are entering neighbors smaller or larger than their existing neighbors? Are they smaller or larger than other cities (and other entering cities) on average? Are older cities more likely to attract neighbors? Simon's model, as popularized by Krugman (1996b), considers that a new city, a lump, may either, with probability ϖ, locate on its own, or, with probability $1 - \varpi$, attach itself to a 'clump', an existing agglomeration. The probability that a lump will join an existing agglomeration is assumed to be proportional to the clump's size (measured in lumps). A strict interpretation of this assumption is that new cities are of the same size and will locate according to the size of existing cities. That is, the probability that a city will locate in an isolated site is a decreasing function of the population of city i at time t, P_{it}. We would expect that spatial evolution according to this theory exhibits strong history dependence. That is, sites that were settled first are more likely to acquire neighbors; and, once they have been settled, they would grow faster. Our data set, which includes the presence of 'neighbors', and also includes times when settlements were founded from which cities developed, allows us to examine this theory. We also consider the role of threshold effects. Our regression results probe the data for non-linearities in the distance variable as well as in the population variable.

4.1. Econometric analysis

Let us define as $\mathcal{I}_t$, the set of cities extant in time t, and for each city i, an integer-valued variable $\theta_{i,t}$, $\theta_{i,t} = 0,1,\ldots$, to indicate the number of its neighbors as of time t. The number of new neighbors city i acquires between time periods $t-1$ and t is thus expressed by $\Delta\theta_{i,t} \equiv \theta_{i,t} - \theta_{i,t-1}$. Let Θ_t denote the vector comprised of all θ_{it}'s: $\Theta_t \in R_+^{I_t}$. The definition applies to both existing cities as well as entering cities.

A theory of the spatial evolution of the city size distribution may be expressed succinctly in the form of a reduced-form dynamic system involving the vectors of city sizes, P_t, of the number of each city's neighbors, Θ_t, and of settlement sites, G_t. Let $\Phi_{\Theta_t},\Phi_{P_t},\Phi_{G_t}$ denote a vector of interdependent random shocks. The evolution of (Θ_t,P_t,G_t) may be expressed through a system of equations as follows:

$$\Theta_t = Y_\Theta(\Theta_{t-1},P_{t-1},G_{t-1};\text{s};\Phi_{\Theta_t}). \tag{1}$$

$$P_t = Y_P(\Theta_{t-1},P_{t-1},G_{t-1};\text{s};\Phi_{P_t}), \tag{2}$$

$$G_t = \Gamma(\Theta_{t-1},P_{t-1},G_{t-1};\text{s};\Phi_{G_t}). \tag{3}$$

We eschew a separate description for $\mathcal{I}_t$, as its evolution is implicit in the above system, provided that we agree to the account for new entries by adding components to the three vectors. This system may be used for studying spatial-size evolution by focusing, alternatively, at new entries (or exits, if they occur) or at existing cities. We may accomplish this by conditioning appropriately, e.g. newly entering cities between $t-1$ and t are described by newly added components of the above vectors, corresponding to entries $j \in \mathcal{I}_t \backslash \mathcal{I}_{t-1}$.

5. Results

We present our results by starting with qualitative aspects of the spatial pattern of urban settlements in the United States. We follow up with considering the notion of central place in the light of the data. Next we examine econometrically the location of new cities and the impact on urban growth of a city's proximity to other cities. We conclude our econometric analysis by examining in depth statistical aspects of observed growth patterns.

5.1. Cities and population patterns

This section examines the predictions of the Fujita/Mori model on the number of cities in an economy and their pattern of distribution, as well Marshall's prediction that older cities are larger cities.

To answer the Fujita/Mori question regarding the number of cities relative to total population, we use our simpler data set, described above and detailed in Table 1. As predicted by the Fujita and Mori model, the Pearson correlation coefficient between population and number of cities is estimated at 0.980 (at the 0.01 level of significance) over the full two centuries. A cross correlation analysis indicates that U.S. population is a leading indicator, and very significantly so, for the number of cities.[7] The reverse is never true. The proportion of new cities entering the system is not proportional to the change in population. This result ignores the complexities of city types, a key element of the Henderson model, as well as spatial considerations.

As to the East to West pattern, we refer readers to Fig. 1, a map showing the increase in the number of cities in each region from 1900 to 1990. In terms of percentage increase, the number of cities increases in an arc, slicing south of the Midwest districts, through the South, and enlarging to include all of the Mountain and Pacific districts of the West (which extend from the northern to the southern boundaries).

However, we also use regression results for a more detailed analysis. We specify Eq. (2) in first-difference form:

$$\Delta \ell n P_{i_t} = \beta + \beta_p \ell n P_{i_{t-1}} + BX_{it} + \psi_{it}, \tag{4}$$

In Dobkins and Ioannides (2000a) we examine patterns in U.S. urban growth by working with Eq. (4) and regressing the growth in population for cities in each census year against its own lagged population and eight Census regions, with the West North Central region (made up of North and South Dakota, Minnesota, Iowa, Nebraska, Kansas and Missouri) omitted.[8] Using t-statistics to judge significant growth or decline, we have noted expected but interesting patterns. Generally, the estimated R^2's range from 0.08 to 0.53. The lagged value of log population is significant only for 1980 and 1990. This suggests that growth in the cross-sectional sense is primarily driven by regional patterns. It is the time-series setting that provides support for the notion of persistent dynamics, with the estimated

[7]The two series are differenced because of stationarity.

[8]This analysis uses the fuller 1900 to 1990 data set. We refrain from reporting the actual results here so as to avoid unnecessary duplication.

autoregression coefficient being very close to 1. However, Dobkins and Ioannides (2000a) rejects the hypothesis that an AR(1) specification of log population has a unit root.

In the first decade of the century, we see westward movement with positive growth in the Mountain and Pacific Coast regions as well as the West South Central region, which includes Texas, Oklahoma, Arkansas and Louisiana. The years from 1910 to 1920 record positive movement into the West South Central and Pacific Coast regions. In the 1930s those two regions again show positive growth along with the East South Central region, which includes Kentucky, Tennessee, Mississippi and Alabama.

It is during the 1940s that we see the first signs of significant decline in cities' population growth rates. The Middle Atlantic region, which includes only New York, Pennsylvania and New Jersey, declined. There was positive growth in all the regions listed above along with the South Atlantic region, the area that is bordered on the north by West Virginia, Maryland, Delaware, and Washington DC, and extends south through Virginia, North and South Carolina, and Georgia to Florida.

The 1950s saw decline in New England, with growth in the West South Central, Mountain and Pacific Coast regions. In the 1960s, there was positive growth in the South Atlantic, Mountain and Pacific Coast regions, and perhaps surprisingly, the East North Central region, which includes Wisconsin, Michigan, Illinois, Indiana and Ohio. The 1970s saw negative growth in the Middle Atlantic region, with growth everywhere to the South and West. In the 1980s the trend reversed for the East North Central region, which saw a decline while the Mountain and Pacific Coast regions saw the only positive growth.

Throughout the twentieth century, growth is positive in the Pacific Coast region, which includes Hawaii and Alaska as well as Washington, Oregon and California. New England, the Middle Atlantic states and the East North Central region each go through periods of decline. If we were simply looking at negative and positive coefficients, we would see a decline in New England during the first part of the century, with that trend reversing in the 1960s. A breakdown by state would reveal more specific reasons for these trends at which we can only guess.[9]

Marshall predicts that older cities will be larger. If we consider each census year of our data set, we see that there is always the (appropriate) negative correlation between the size of cities in a given year and their age. Interestingly, the size of this correlation declines over time, beginning at -0.309 in 1900 and moving to -0.167 in 1990. These coefficients, although relatively small, are always statistically significant at the 0.01 level. We interpret this to mean that the initial advantage wanes over time, as one might expect in a system also sensitive to agglomeration effects, general population growth, and changing technologies.

Therefore, we conclude that the predictions of Fujita/Mori and Marshall are

[9]The association of Tennessee with Great Depression programs during the 1930s comes to mind here.

confirmed, with only slight qualifications. We find that the growing population in the United States leads to a larger number of cities. Growth over time does follow a pattern that is generally East to West – or more specifically, East to South and West. Finally, older cities do tend to be larger, although the benefits of initial advantage wane over time.

5.2. Central place considerations

Our data offer us the opportunity to examine the central place considerations of Christaller. Central place considerations yield several interesting results.

We first ask if higher order cities, based on the tiers established for the distance variable, are also the oldest. As it turns out, the correlation between functional order and date of settlement is always positive, but rarely is the Pearson correlation coefficient significant even at the 0.05 level, considered on a decade by decade basis. By the final two decades for which we have data, 1980 and 1990, the correlation is less than 0.1. (The correlation of 0.126 is significant at the 0.01 level for the century as a whole, but that is because of the large sample size, 1988 observations for all cities over all years).

We might ask about the correlation between central place and size, a key ingredient of the Christaller theory. In this case, the correlation is large (an absolute value of 0.66 over the entire period) and significant, but hardly 'complete'; that is, the list of the largest places would not duplicate the list of tier one cities.[10]

Thirdly, we look at the relationship between tier/function and distance apart. Central place theory would suggest that the higher order cities should be further apart than lower order places. Of course, Christaller's theory stipulated a 'featureless plain', and it is of interest to see how the relaxation of that assumption affects the relationship.

Tier-one (highest order) cities are a mean 392 miles apart, but with a standard deviation of 360 miles. (There are 77 occurrences of tier-one cities over the century). The minimum distance between tier-one cities is 101 miles, the maximum distance 2072 miles. Tier-two cities (with 131 occurrences over the entire period) are a mean 408 miles from the nearest city in tier one, with a minimum of 36 miles and a maximum of 1281. Tier-three cities (137 occurrences) are a mean 243 miles from a city in either of the higher tiers, although the standard deviation falls to 167 miles. Tier-four cities, the great bulk of the designation at 1643 occurrences, do tend to be relatively close to cities in a higher tier (any of the

[10] The reader is reminded, in regard to both of these results, of the example of Detroit. The city is quite large, with more than 4 million people in 1990, and relatively old, with a settlement date of 1701, but belongs in the fourth tier by 1990 because of its industrial, as opposed to higher function, role.

three higher tiers), with a mean of 146 miles and a standard deviation of 129 miles.[11]

It would seem that central place theory predictions do not hold on a 'plane with features'. There is no significant correlation between functional order and date of settlement, and the largest places are not necessarily tier-one cities. While the tier-four cities are close to higher tier cities, as would be expected, the tier-one cities are not necessarily furthest apart.

5.3. Cities with neighbors

Next, we turn to the questions involving when and where new cities locate in relation to existing cities.

5.3.1. Patterns among neighbors

Descriptive statistics for our data, given in Tables 2 and 3, but especially in Table 3, reveal important features of the force of agglomeration in U.S. economic geography. Roughly one fourth of all cities have neighbors over all years. However, of the 222 cities that enter after 1900, nearly 16% locate so as to have other cities as neighbors. No such entries occur in 1930, 1940, 1950 or 1990.

We see evidence in Table 3 of some of the enduring facts of U.S. economic geography, and an interesting spatial interpretation. The population 'boom' of the 1950s resulted in 48 new cities entering the system, almost a third of them as neighbors to either existing cities and/or to each other.[12] The so-called 'rural renaissance' of the 1970s, however, resulted in 79 new cities entering the system, with less than 10% of those being neighbors.

Table 3, columns 7 and 8, suggest that cities with neighbors and the neighbors themselves tend to be larger than isolated cities. Whereas column 7 shows that average city sizes generally grow over time, the opposite is true for city sizes relative to total urbanized population, reported in column 8. Column 7, Table 3, indicates that the average size of a city with no neighbors in 1900 was 192 000. The average size of a city with neighbors was 487 000 and the average size of the neighbors was 571 000. (These numbers differ because some cities have more than one neighbor and because not all neighbors to a central city are neighbors to each other). This pattern continues through the century.

Taken alone, these numbers might seem to contradict the central place notion that the largest concentrations must have small neighbors. Fujita and Mori (1997) also suggest that as new cities develop their neighbors will be smaller (because the

[11] The inclusion of Anchorage and Honolulu skews these numbers somewhat to be larger than they would be if we considered only the continental United States, but does not affect our qualitative results.

[12] The so-called population boom of the 1950s is, of course, relative and modern. The 19% increase in population would have rated as the smallest increase in the period from 1790 up through the first decade of this century.

Table 3
Sizes of cities and their neighbors, absolute and normalized growth rates

Year/ nei's	Number of cities	Number >1 nei	New cities	Growth rate, abs	Growth rate, nor.	S.D. nor.	Size (000's) with/of nei's	Size nor.
	1	2	3	4	5	6	7	8
1900/no nei's	86						192	0.007
1900/nei's	26	2:2					487/571	0.017
1900/all	112						261/133	0.009
1910/no nei's	109		25	0.233	−0.08	0.159	202	0.005
1910/nei's	30	2:2	2	0.340	0.027	0.248	597/687	0.015
1910/all	139		27	0.260	−0.052	0.190	287/148	0.007
1920/no nei's	113		7	0.216	−0.02	0.148	215	0.004
1920/nei's	36	2:2	3	0.250	0.166	0.180	726/818	0.014
1920/all	149		10	0.224	−0.009	0.156	339/198	0.007
1930/no nei's	117		6	0.219	−0.03	0.151	260	0.004
1930/nei's	40	2:2	2	0.243	−0.004	0.184	855/954	0.013
1930/all	157		8	0.223	−0.022	0.160	412/243	0.006
1940/no nei's	120		3	0.106	0.024	0.104	279	0.004
1940/nei's	40	2:2	0	0.085	0.002	0.074	916/1017	0.013
1940/all	160		3	0.101	0.018	0.097	438/254	0.006
1950/no nei's	122		2	0.233	0.034	0.151	342	0.004
1950/nei's	40	2:2	0	0.224	0.025	0.157	1096/1211	0.013
1950/all	162		2	0.231	0.032	0.152	526/299	0.006
1960/no nei's	150		33	0.227	−0.047	0.219	365	0.003
1960/nei's	60	4:2; 7:3; 1:6	15	0.176	−0.098	0.349	964/1425	0.009
1960/all	210		48	0.213	−0.062	0.262	535/407	0.005
1970/no nei's	173		24	0.178	−0.036	0.156	374	0.003
1970/nei's	70	11:2; 9:3; 5:4; 1:7	7	0.009	−0.204	0.715	1067/1560	0.008
1970/all	243		31	0.129	−0.085	0.412	575/449	0.004
1980/no nei's	244		72	0.185	−0.010	0.148	356	0.002
1980/nei's	78	16:2; 7:3; 8:4; 1:11	7	0.112	−0.083	0.218	1060/1706	0.006
1980/all	322		79	0.164	0.031	0.174	527/413	0.003
1990/no nei's	256		12	0.084	−0.04	0.162	391	0.002
1990/nei's	78	16:2; 7:3; 8:4; 1:11	0	0.118	−0.01	0.151	1184/1860	0.006
1990/all	334		12	0.093	−0.035	0.160	577/434	0.003
Total no	1490			0.174	−0.023	0.160	318	0.003
Total w/	498			0.150	−0.049	0.335		0.01
Total	1988		222	0.168	−0.029	0.220		0.005

previous 'frontier' cities now have competition on both sides). The fact is, when we consider the data carefully, we see a pattern that is not inconsistent with central place theory. Entering neighbors are large (up to 2 million, as noted above, at 'birth'), but they are usually relatively small compared to their already existing neighbor.

In our set of 78 cities that have neighbors over the years from 1900 to 1990, 56 are involved in either entering as a new neighbor or being the existing neighbor to a new entrant. The other 22 are cities that co-exist as a neighbor in 1900, and do not overlap the previous set. (For example, Bridgeport, CT is a neighbor to New York City in 1900 and is tallied among the 22. New York City is counted among the 56 with its ten other neighbors that enter over the century). Among the 56, all entering neighbors are smaller than their existing neighbors except for Greensboro, NC, which enters as a neighbor to Winston-Salem. These cities are an exception to the rule throughout the years, as the Greensboro–Winston-Salem–High Point area grows together quite quickly. Among the 56, excepting Greensboro and Winston-Salem, the average percentage of the size of the entering city to the size of the existing city is 18%. This includes such large concentrations as Nassau and Suffolk (NY) counties, noted above.

Interestingly, of the neighbors that coexist in 1900, the smaller neighbors are, on average, 32% of the size of their larger neighbors. This may highlight a feature of the data set, in that cities are designated as neighbors if they are ever grouped together by the Census Bureau. These groupings were published relatively late in the century. Perhaps, with less efficient transportation, these cities were actually further apart in a real sense in 1900. To check this, we note the average percentage of the same group of neighbors in 1990. This averages turns out to be 28%; it would be 21% if we were to leave out Scranton, PA and Wilkes–Barre, PA. This is another noteworthy group of cities (which the Census Bureau simply calls 'Northeast Pennsylvania' in 1980); and which reverses dominant size, with Scranton the smaller city in 1900 and the larger in 1990. Although these numbers deal with a small set of cities, the analysis does seem to bear out some of the theoretical predictions. Cities tend to be smaller than the core city in an 'agglomeration shadow', although the entire agglomeration is larger than isolated cities. Furthermore, cities with some initial advantage (in 1900, for instance), may 'lock-in' and remain relatively large even as a neighbor grows more rapidly.

We perform a number of probit regressions based on the sample of cities that enter during each decade, with the dependent variable being whether or not new cities enter and locate as neighbors of either existing cities or become isolated cities. We found (Table 4, column 6) that a new city's own size is the single most important determinant of whether it will locate so as to have neighbors. Being in the New England or Middle Atlantic states (the Northeast designation), was also statistically significant. This probit regression included census dummies along with regional dummies. Inclusion of a third degree term for log size (Table 4, column 8) shows that the non-linear structure falls short of statistical significance.

For the event that an existing city has either no neighbors or at least one neighbor as of time t, we assume that

$$\text{Prob}\{\theta_{it} > 0\} = \text{Prob}\{b_0 + b_1 \ \ell n\, P_{it-1} + B_0 X_{it} + \phi_{\theta it} > 0\}, \tag{5}$$

where B_0 denotes a vector of parameters, X_{it} denotes a vector of regressors, which will be discussed in more detail below, and $\psi_{\theta it}$ is a random variable. For convenience we will adopt a homogeneous probit specification for $\phi_{\theta it}$ and thus assume that it is IID across all observations and has a normal distribution.

A probit estimation for the event that a city has neighbors along the lines of Eq. (5) is reported in Table 4, column 1. It shows that a city's own size increases the likelihood that it has neighbors and very significantly so. Furthermore, being an older city increases the likelihood of having neighbors. Regional dummies are also important: being in the Southeast makes having a neighbor much less likely while being in the Pacific Coast group makes it more likely. Inclusion of a third-degree polynomial structure for size (see Table 4, column 7) is not statistically significant relative to just size.

We conclude that the data provide support for the essential intuition behind Simon's model of random urban growth, in that the larger an existing city the more likely it is that it will have neighbors. Our finding that the larger an entering city the more likely it is that it will locate so as to have neighbors confirms that the same intuition also applies to new entrants, which in Simon's theory are of the same size. Our findings do not support threshold effects, which might be implied by the newer theories noted above.

5.3.2. Spatial interactions and urban growth

Next we specify Eq. (2) as a selection model by working jointly with Eq. (5) and a system of equations like Eq. (4). This system explains the evolution of the size of city i conditional, respectively, on whether or not it has at least one neighbor, that is,

$$\Delta \ell n\, P_{i_t} = \beta_1 + \beta_{p1} \ \ell n\, P_{i_{t-1}} + \beta_n \ \ell n\, \overline{P}_{\nu(i)_t} + B_1 X_{i\nu(i)t} + \psi^1_{it}, \tag{6}$$

$$\Delta \ell n\, P_{i_t} = \beta_0 + \beta_{p0} \ \ell n\, P_{i_{t-1}} + B_0 X_{it} + \psi^0_{it}, \tag{7}$$

where the random variables $(\psi^1_{it}, \psi^0_{it})$ are assumed to be correlated with $\phi_{\theta it}$, the stochastic shock in Eq. (5). Estimation of Eqs. (6) and (7), under the assumption of endogenous selection, relate to the impact of spatial interaction on urban growth.

We underscore the economic significance of the switching regressions model here: the law explaining the evolution of city size is different once a city acquires neighbors. This is designed to express the divergence between the *sustain* and *break points*, that are critical features of nonlinear dynamic growth in the urban system [Fujita et al., Ch. 3].

Table 4
Urban growth and spatial interactions[a]

Regression	1	2	3	4	5	6	7	8
sample	Probit all	$GRi,t-1,t$ w/nei	$GRi,t-1,t$ w/o nei	$GRi,t-1,t$ w/nei	GR1,00,10 w/nei	Probit new	Probit all	Probit new
Constant	0.381	0.402	0.035	0.494	−0.043	−14.75	−11.13	614.7
	(0.297)	(3.40)	(0.382)	(3.88)	(−0.133)		(−0.212)	
ℓnP_{t-1}	0.292	−0.028	0.0005	−0.034	0.035	0.645	4.72	−145.9
	(7.64)	(−2.95)	(0.064)	(−3.37)	(1.13)	(2.72)	(0.384)	(−1.56)
$(\ell nP_{t-1})^2$							−0.478	11.31
							(−0.502)	(1.50)
$(\ell nP_{t-1})^3$							0.016	−0.289
							(0.643)	(−1.43)
$GR\nu(i),t-1,t$		0.171		0.118	0.073			
		(2.33)		(1.27)	(0.234)			
Date	−0.003						−0.003	
	(−4.40)						(−4.42)	
Dist		−0.0003	−0.00003	0.00009	−0.002			
		(−0.158)	(−0.160)	(0.425)	(−2.47)			
$Dist^2$		2.54×10^{-7}	7.41×10^{-7}	6.79×10^{-8}	6.03×10^{-6}			
		(0.876)	(1.23)	(0.208)	(3.13)			
$Dist^3$		-1.31×10^{-10}	-5.97×10^{-10}	-6.63×10^{-11}	-2.38×10^{-9}			
		(−1.23)	(−1.64)	(−0.551)	(−3.21)			

North East	0.008	−0.0009	−0.042	−0.007	0.0008	1.19	0.064	1.20
	(0.076)	(−0.039)	(−3.09)	(−0.292)	(0.010)	(3.15)	(0.572)	(3.06)
South East	−0.409	0.075	0.079	0.080	−0.205	−0.153	−0.365	0.028
	(−3.92)	(2.64)	(5.78)	(2.78)	(−3.13)	(−0.416)	(−3.48)	(0.075)
South West	−0.185	0.174	0.108	0.212	0.377	−0.697	−0.174	−0.727
/Mountain	(−1.65)	(4.27)	(5.64)	(4.18)	(1.53)	(−1.63)	(−1.56)	(−1.63)
Pacific	0.437	0.240	0.198	0.249	0.084	0.165	0.477	0.273
	(3.44)	(5.13)	(7.39)	(5.21)	(0.911)	(0.395)	(3.71)	(0.623)
Error		−1.59	−1.67					
		(−17.2)	(−40.78)	(−0.121)				
Corr. Coef/t		0.059	−0.182					
			(−2.82)					
Observations	1654	418	1220	420	26	217	1654	217
LLF	−859.20	−764.07	−547.31			−68.68	−848.45	−65.26
χ^2 p	185.51							
R^2	0.097			0.355	0.900	0.318	0.109	0.352
F				11.97	25.57			

[a] Columns 1, 2 and 3 report a selection model for the growth rate of population; column 1 is a probit regression, and columns 2 and 3 are switching regressions for the 10-year growth rate of city population. Wave dummies are also included in these regressions. Column 4 reports the uncorrected regression for the 10-year growth rate of city population. Column 5 reports a regression for the GR*i*,00,10. Column 6 reports a probit regression for the event that a newly entering city locates so as to have neighbors, where the row for ℓnP_{t-1} corresponds to ℓnP_{t}. Columns 7 and * correspond to columns 1 and 6, respectively. *t*-Statistics in parentheses.

Columns 2 and 3, Table 4, report a selection model for the growth rate of population, which uses the probit estimates according to Eq. (5), reported in column 1 for regime switching. Distance, as defined in this study in a central place configuration, is not significant for either cities with or without neighbors. The results in Column 2 do show that, when distance in this central place sense is accounted for, the impact of a city's own size on its growth differs between cities that have neighbors and those that do not. For a city with neighbors (column 2), own size has a negative impact on growth, but the growth rate of neighboring cities is a positive stimulus to own growth. For cities without neighbors, own size has little impact on growth.

Column 4 reports the uncorrected regression for the 10-year growth rate of city population. This regression does not account for the panel structure of the data, and it is for this reason that we have also carried out a number of additional regressions by conditioning on each pair of successive periods. One such regression for the $GR_{i,00,10}$ is reported in column 5 of Table 4.

The regression in column 4 (Table 4) pertains to the pooled data for all cities with neighbors. The average growth rate among a city's neighbors is still a positive determinant of a city's own growth rate, but is not statistically significant in this uncorrected regression. When we look at specific pairs of successive periods, the positive coefficient is statistically significant in three of the nine periods: 1910 to 1920, 1930 to 1940, and 1940 to 1950. More typical is the period shown in column 5, for the decade from 1900 to 1910. Regional dummies are significant in the pooled data and accord with the intuition obtained from observing the urban expansion away from the Northeast and towards the South and West. Distance from the nearest center is statistically significant in these regressions only in 1900 to 1910.

Another set of regressions explaining city populations using the date variable as well as the distance variable are reported in part in Table 5. We regressed the date of settlement (after subtracting the earliest year, 1564, when Jacksonville/St. Augustine, FL was founded), distance (divided by ten), distance squared (divided by 100) and distance cubed (divided by 1000), as well as regional dummies, against the natural log of population of each city in each time period. We then repeated the exercise for the set of cities with neighbors and the set of cities without neighbors in each period. We found that, for cities with neighbors, the date variable was highly significant in the earlier years and declines somewhat in statistical significance after 1950. (The variable was statistically significant in all years except 1970). Interestingly enough, in the years up to and including 1950, the date variable always carried a higher '*t*-value' for the cities with neighbors, although the variable was most often statistically significant for both. In 1980 and 1990, the cities without neighbors were more influenced by the date of settlement than were cities with neighbors, at about the level of significance we see for the entire time period. Perhaps this is because the cities with neighbors are often large,

Table 5
Urban populations and spatial interactions[a]

Year	1910	1910	1950	1950	1980	1980
Sample	w/o/nei	w/nei	w/o/nei	w/nei	w/o/nei	w/nei
Constant	12.662	13.208	12.610	13.410	12.517	13.175
	(28.14)	(29.49)	(29.14)	(26.29)	(39.66)	(37.85)
Date	−0.0033	−0.0099	−0.0030	−0.0089	−0.0030	−0.0036
	(−2.86)	(−5.25)	(−3.05)	(−4.46)	(−3.51)	(−2.40)
Distance	−0.0150	0.1460	0.0310	0.2010	0.0380	0.0440
	(−0.441)	(4.42)	(0.926)	(3.11)	(1.99)	(1.25)
Dist^2	0.0010	−0.0020	−0.0002	−0.0046	−0.0004	−0.00006
	(0.971)	(−3.54)	(−0.182)	(−2.09)	(−0.792)	(0.058)
Dist^3	−0.0000	0.0000	0.0000	0.00003	0.0000	−0.0000
	(−0.939)	(3.24)	(0.126)	(1.66)	(0.553)	(−0.378)
North/	−0.030	–	0.176	–	0.077	–
Northeast	(−0.128)	(–)	(0.825)	(–)	(0.423)	(–)
Southeast	−0.431	−1.04	−0.0500	−1.865	0.158	−0.244
	(−1.74)	(−1.34)	(−0.230)	(−4.25)	(0.865)	(−0.838)
Southwest/	−1.044	−1.871	−0.5040	0.1920	−0.0550	0.3280
Mountain	(−2.90)	(−2.65)	(−2.13)	(0.350)	(−0.296)	(0.714)
Pacific	−0.7940	−0.4640	0.3890	0.7170	0.2910	0.8710
	(−2.09)	(−0.952)	(1.19)	(1.60)	(1.32)	(2.43)
Observations	109	30	122	40	244	78
R^2	0.216	0.721	0.257	0.638	0.160	0.356
F	3.44	8.11	4.89	8.05	5.59	5.54

[a] Columns reflect cities with and without neighbors for 1910, 1950, and 1980. The regressions are the adjusted date variable, distance, distance squared and distance cubed (all adjusted), and the regional dummy variables against the natural log of population. *t*-Statistics in parentheses.

'core' cities, whose advantage in settlement carried them through the first half of the twentieth century.

As for the role of distance, there is a pattern similar to that of the date variable. Early in the century, up through 1950, the distance from a city in a higher tier is a significant explanatory variable for population in cities with neighbors. That changes after 1950, when for cities without neighbors, distance becomes insignificant, until this changes again in 1980 and 1990. Because distance to the nearest higher tier city helps determine whether a city is a neighbor, this result seems reasonable for the early years. Something in the dynamics of 'neighborliness' must change after the rural renaissance of the 1970s, but the trend is not clear without another decade-worth of data.

We find that the cities that have neighbors are large cities. Cities that enter as neighbors may be large, but their existing neighbors are even larger. The entire agglomeration is larger than isolated cities. Our findings offer some support for the Simon predictions that larger cities are more likely to draw neighbors. However we do not find evidence of threshold effects. We find interesting interactions between neighbors: the city with neighbors responds positively to its neighbor's growth, but negatively to its own size. For cities without neighbors, own size has little impact on growth.

5.3.3. The structure of urban growth

Finally, we take a preliminary non-parametric look at statistical aspects of the observed growth rates. Absolute growth rates, defined in terms of absolute populations, and relative growth rates, defined in terms of urban populations relative to total U.S. urbanized population, are reported in Table 3, columns 4 and 5, respectively. Column 6, Table 3, reports the standard deviation of the relative growth rate.

As one may conclude from the evidence, normalized growth rates give a very different picture of urban growth from that of absolute ones. Analysis of variance for growth rates shows that wave dummies,[13] regional dummies and their interactions explain 28% of total variance.

We examine the statistical variation of the absolute growth rate across its own deciles as well as the deciles of lagged population, separately for each year as well as for the entire panel. The variance of the growth rate across its own deciles for each year suggest (albeit very roughly) a U-shaped pattern: the mean growth rate declines as we move to the upper deciles of lagged population though not uniformly and with several deviations. The variance of the growth rate across its own deciles for each year also suggests a U-shaped pattern, though less pronounced than that for the mean. Unfortunately, this analysis is rather inconclusive and therefore we do not report it here in further detail. It is available from the authors on request.

Gabaix (1999) explains Zipf's Law in terms of Gibrat's Law for city sizes: if city growth rates are independent and identically distributed random variables, then Zipf's Law holds in the upper tail of the size distribution. Gabaix attributes a critical role to the elasticity of the variance of the growth rate with respect to normalized city size by linking it to the existence and magnitude of the exponent of Zipf's Law. If the mean growth rate is independent of size, then a sufficient condition for the Zipf' Law exponent to be less than 1 is that the variance of the growth rate decline with size.

We test this particular feature of Gabaix's theory, performing in effect a structural test of Zipf's Law, by computing the variance of the growth rate within

[13] We believe that wave dummies are necessary to account for census-specific conditions that may be at work. We also believe that conventional serial correlation assumptions are not appropriate in view of 10-year intervals between our observations.

each decile of population size in each year, and by regressing it against the mean (or the median) of its respective decile and wave dummies. While this regression is significant with $R^2_{\text{adj.}} = 0.287$ and an F statistic which is significant with probability 0.0001, it yields a statistically insignificant positive coefficient for the decile median (alternatively, mean), and three significant coefficients for wave dummies. Therefore, our data, at least, provides no support for this key assumption in Gabaix (1999).[14]

6. Conclusions

We end by summarizing our key findings, some of which are essentially descriptive of the United States system of cities, and others that relate to geography and spatial interactions. It is an important fact that in the United States, in contrast to France and Japan (see Eaton and Eckstein, 1997), population growth has spawned new cities. The model of Fujita and Mori (1997) accords with this finding. Furthermore, in the United States, older cities are larger, as suggested by the mercantile model of Marshall and his rewriting of Christaller's central place theory. For the first part of the century, at least, older cities seemed to benefit, population-wise, from their age, suggesting that initial advantage conferred a benefit that began to wane only in the latter part of the twentieth century.

Spatial considerations are important in urban growth. Urban expansion away from the Northeast and toward the South and West shows up repeatedly in a number of different configurations. The likelihood that an entering city will locate so as to have neighbors is increasing with its own size and its age. This seems to us to support the Simon/Krugman notion of 'lumps' locating near 'clumps'. Distance is not always an important determinant of size and growth and we see no evidence of non-linear effects in the distance variable. This is a very simplistic way of looking at the threshold effects implied by the new economic geography, and we note that the result might be different if distance is interpreted without regard to the functions of cities. Our switching regressions tell us that, for cities with neighbors, growth rates are closely interdependent. For cities without neighbors, own size has little impact on growth.

We recognize that our data set is not perfect: even in a century of phenomenal increase in the use of economic numbers, definitions and procedures change, requiring that researchers make any number of judgment calls on data gathering. Furthermore, there are many more tests we might exercise to test predictions.

We end by reiterating that our data set was designed to enhance our understanding spatial interactions within the U.S. system of cities. We see that the present paper complements ongoing research by many others, which utilize disaggregated

[14] Ioannides and Overman (2000a) using non-parametric techniques, including numerical integration, do provide support for Gabaix's prediction.

data (either micro data on firms or industry data (Black and Henderson, 1999) and aims at understanding the main forces determining patterns in U.S. regional specialization and localization over the last century and more (Kim, 1995).

Acknowledgements

We are grateful to the referees and to Konrad Stahl for numerous constructive suggestions and comments. We also thank Ed Glaeser and Stuart Rosenthal for valuable comments, Xavier Gabaix for useful suggestions and Tim Harris, Pamela Jia and Tracey Seslen for exceptional research assistance. Ioannides thanks the John D. and Catherine T. MacArthur Foundation and the National Science Foundation for generous research support. We acknowledge comments received during presentations of earlier versions at the 1998 Econometric Society Meetings, Chicago, and at Brown University. This is a major revision of the May 26, 1999 version of the paper.

Appendix A. Details on the data and definition of neighbors

Descriptive statistics on the entire data set are given in Table A.1 below.

We provide here additional details on the definitions we employ in our construction of the data on new neighbors. We also elaborate on the size, actual geographical location and composition in terms of counties of entering neighbors. In contrast, Table 3 looks only at the average size of neighbors in any given decade.

How should we deal with a city that enters the data set in a given decade and

Table A.1
Descriptive statistics for all cities, 1900–1990[a]

Variable	Mean	Std. Dev.	Skewness	Kurtosis	Min	Max
Population (000)	479.5	1001.5	6.6	58.8	50.7	9372.0
Log(Population)	12.4028	0.9895	1.0	4.1	10.8343	16.374
Growth rate (%)	10.62	41.98	−1.1	5.8	−0.999	187.52
New England	0.0879	0.2833	2.9	9.5	0.00	1.00
Mid Atlantic	0.1276	0.3338	2.2	6.0	0.00	1.00
South Atlantic	0.1673	0.3734	1.8	4.2	0.00	1.00
East North Central	0.2030	0.4023	1.5	3.2	0.00	1.00
East South Central	0.0663	0.2489	3.5	13.1	0.00	1.00
West North Central	0.0910	0.2876	2.8	9.1	0.00	1.00
West South Central	0.1221	0.3275	2.3	6.3	0.00	1.00
Mountain	0.0462	0.2100	4.3	19.7	0.00	1.00
Pacific	0.0884	0.2840	2.9	9.4	0.00	1.00

[a] 1 998 Observations.

becomes a neighbor to an existing city? For example, in 1980, Rock Hill, NC, and Salisbury, NC enter as neighbors to Charlotte, NC. Table 3 includes Charlotte in the count of cities with new neighbors, (because it acquires neighbors that cause the number of cities with neighbors to increase from 70 to 78). Charlotte is an existing city which has never before had neighbors. It is appropriate to count Charlotte in the count of cities that are neighbors of another city (as we had done earlier). Yet, for purposes of looking at sizes of entering new cities, it is appropriate to count Rock Hill and Salisbury, but not Charlotte. Clearly, Rock Hill and Salisbury are influenced by the size and existence of their larger neighbor. To continue with 1980 examples, Santa Cruz, CA enters and is a neighbor to San Francisco, CA. San Francisco does not count as an entering neighbor, as Charlotte does, because it previously had neighbors.

The cities that enter as neighbors to an existing city are very few and can be listed individually. Such a list follows. The table that follows provides averages for these categories. It shows that before 1950 entering neighbors were generally smaller that the average size of an entering city (all entering cities, both neighbors and those without neighbors). After 1950, the average size of an entering neighbor is much larger than the average size of all entering cities. Recall that we rely on data from Bogue (1953) for census years prior to 1950 and including 1950.

In 1910, Table 3 shows two new neighbors. Racine, WI joins as a neighbor to Milwaukee, WI and is much smaller. Riverside, CA joins Los Angeles, CA and is a tenth its size. In 1920, three neighbors join existing cities (so that Table 3 lists six new neighbors in all).

Kenosha enters with a population of 51 000, as a neighbor to Chicago, IL. Galveston, TX with a population of 53 000, is a neighbor to Houston, TX. Winston-Salem, NC enters as a neighbor to Greensboro, NC which had entered in the previous census. Their populations are similar, although Greensboro, NC is slightly larger (79 000 to 77 000.) In 1930, Durham, NC enters as a neighbor to Raleigh, NC. Durham's population is 67 000, compared to 95 000 for Raleigh. Ogden, UT entered as a neighbor to Salt Lake City, UT and is much smaller, just over 50 000. In 1940 and 1950, no new cities entered as neighbors.

In 1960, we see the outcome of a large number of new cities' entering, (48 new cities), and of different methods of designating metropolitan areas by the Bureau of the Census. Table 3 notes 20 new neighbors, which includes 12 cities by our definition for present purposes. A look at each of these cities is illuminating. Lawrence, MA and Lowell, MA both became designated as separate cities, with populations of 188 000 and 158 000, respectively. Bogue (1953) notes that Lawrence and Lowell had previously been counted with Boston. Lawrence, however, is actually the designation for Lawrence, MA, and Haverhill, MA, and some additional towns in New Hampshire. The actual land area of Lawrence is only a small part of the total area, although Lawrence has the single largest population. This follows the Census Bureau's definition (which we adopt) of designating a city to 'stand on its own' when the central city (Lawrence in this case) exceeds 50 000 and other nearby areas seem economically linked. That matter is particularly important in New England, where metropolitan areas may involve parts of counties.

Another New England example is New Britain, CT which enters with a population of 129 000, just about a fifth of the size of its neighbor, Hartford, CT. Bogue lists New Britain as having been included with Hartford in 1950 and before, although by later definitions, New Britain would have been large enough to be a separate city earlier. The same is true of Waterbury, CT, which enters by definition in 1960 as a neighbor to New Haven, CT (at a third of New Haven's size).

Steubenville, OH and Wheeling, WV were included together in the 1950 definition and were separated by the Census Bureau in 1960, so that Steubenville seems to enter as a neighbor to Wheeling, with Steubenville having a slightly larger population.

Some cities may actually have been too small to count as a metropolitan area before 1960, but grew in the general 1950s growth spurt to be large enough to fit the Census definitions. Ann Arbor, MI with 172 000, enters as a neighbor to Detroit, MI. Ft. Lauderdale, FL entered as a neighbor to Miami, with a third its population. Newport News, VA joins Norfolk, VA with a population of 225 000, compared to 579 000 for Norfolk).

Perhaps the situation involving New York City requires particular elaboration. Bogue did not include (using Census Bureau definitions) Newark, NJ (or Jersey City, NJ) as separate cities in 1950, although both are clearly large enough to have been included even in 1900. When Newark enters as a separate city in 1960, it has 1.7 million people. Paterson, NJ joins that year with 1.2 million. On the other hand, Norwalk, CT joins with 97 000, and Connecticut areas were not included in the 1950 definition. (Norwalk and Stamford, CT, which enters later, were included in Bridgeport, CT which is counted as a separate city but a neighbor to New York City since 1900). So, while New York City itself loses population in our data set from 1950 to 1960, it would have actually gained population if we had been able to separate out Newark and Jersey City and possibly other New Jersey cities earlier. It is important to note a difference in definition, where city boundaries for the New Jersey municipalities are used at an earlier date, and Census-defined metropolitan areas later.

A similar situation may apply to Gary, IN. Gary becomes a city by the new definition in 1960 and enters with nearly a half million population. Parts of East Chicago, IN are included in the Gary definition, and probably were all included with Chicago earlier, although it is not clear that all of the accompanying rural areas were involved. These kinds of discrepancies go a long way toward explaining the average population figures laid out below.

In 1970, Anaheim, CA (and the rest of Orange County) enter with a population of 1.4 million as a neighbor to both Los Angeles and Riverside. Oxnard, CA (and the rest of Ventura county) joins as neighbor to Los Angeles and Riverside with 377 000. In California, each of these four cities are also separate counties. It appears that in 1960 and all years before, the metropolitan area of Los Angeles was only Los Angeles County (including the cities of Los Angeles and Long Beach, CA).

Several cities join with much smaller populations. Bristol, CT which was included in the Hartford metropolitan area in 1950, reaches a population of 66 000 and becomes a separate city and neighbor to Hartford. Nashua, NH joins Boston, Lowell and Lawrence as a neighbor, with 78 000 population. Danbury, CT joins New York City as a neighbor with 78 000 population. Vineland, NJ joins Philadelphia, PA and Wilmington, DE with a population of 121 000; and Petersburg, VA joins Richmond, VA with a population of 129 000. Santa Rosa, CA and Vallejo, CA join San Francisco and San Jose with 205 000 and 250 000 populations, respectively.

In 1980, Rock Hill and Salisbury join Charlotte. Santa Cruz joins San Francisco and the other Bay area neighbors with a population of 188 000. Again, the cities that join the data set as neighbors of New York City are to be listed individually. In New York state, the Nassau and Suffolk counties become a new metro area (known as Nassau in our data set) in 1980, with a population of 2.6 million. These two counties were part of New York City in 1970. (Looking at the fine print, one sees that New York City also gained a small county in New York, and a much larger one in New Jersey. In spite of these changes, New York City's population falls). Three other areas become metropolitan areas, neighboring New York City: Orange County, NJ, with population 260 000; Monmouth, NJ, with population 500 000; and New Brunswick, NJ, with population 596 000. No new neighbors are added in 1990.

These numbers highlight the dilemma of this or any data set. Changing compositions are necessary in order to account properly for the notion of metropolitan areas as construed by the U.S. Bureau of the Census. Particularly over a 10-year span, it is reasonable to accept that changing definitions are

Table A.2

Census	Cities entered since previous census	Average size entrants	Number of cities that are neighbors	Average size of entering neighbors
1910	27	61 247	1	57 000
1920	10	70 172	3	60 000
1930	8	68 717	2	60 000
1940	3	63 749	0	n.a.
1950	2	57 535	0	n.a.
1960	48	216 297	12	405 000
1970	33	183 284	9	299 000
1980	79	184 367	7[a]	613 000[a]
1990	12	135 417	0	n.a.

[a] The numbers for 1980 are increased because of an outlier, Nassau County, NY. Excluding Nassau County, the number of entering would be six and the average size would be 282 000.

necessary. And finally, as we see from the discussion above, the alternative to using these definitions is worse: considering the alternative of using cities like Newark as separate from New York City at all times, we would end up using city proper data. That surely is not a better alternative.

Given these concerns, it is of special interest to refer to Table A.2 below, which gives the number and average size of entering cities and number and average size of entering neighbors (in the sense defined above) in each decade after 1900.

References

Arthur, W.B., 1994. Increasing Returns and Path Dependence in the Economy. The University of Michigan Press, Ann Arbor, Michigan.

Black, D., Henderson, J.V., 1999. Urban Evolution in the USA. Brown University and London School of Economics, mimeo.

Bogue, D., 1953. Population Growth in Standard Metropolitan Areas 1900 – 1950. Scripps Foundation in Research in Population Problems, Oxford, Ohio.

Christaller, W., 1933. In translation (1966), Central Places in Southern Germany, translated by Baskin, C.W. Prentice–Hall, Englewood Cliffs, New Jersey.

Dobkins, L.H., Ioannides, Y.M., 2000a. Dynamic evolution of the U.S. city size distribution. In: Huriot, J.-M., Thisse, J.-F. (Eds.), The Economics of Cities. Cambridge University Press, Cambridge, pp. 217–260.

Dobkins, L.H., Ioannides, Y.M., 2000. A Data Set for U.S. Cities, 1900 – 1990, mimeo.

Eaton, J., Eckstein, Z., 1997. Cities and growth: theory and evidence from France and Japan. Regional Science and Urban Economics 27, 443–474.

Fujita, M., Krugman, P., Mori, T., 1999a. On the evolution of hierarchical urban systems. European Economic Review 43, 209–251.

Fujita, M., Krugman, P., Venables, A.J., 1999b. The Spatial Economy. MIT Press, Cambridge, Massachusetts.

Fujita, M., Mori, T., 1996. The role of ports in the making of major cities: self-agglomeration and hub-effect. Journal of Development Economics 49, 93–120.

Fujita, M., Mori, T., 1997. Structural stability and evolution of urban systems. Regional Science and Urban Economics 27, 399–442.

Gabaix, X., 1999. Zipf's Law for cities: an explanation. Quarterly Journal of Economics, CXIV, August, 739–767.

Hanson, G., 1998. Market Potential, Increasing Returns and Geographic Concentration, University of Michigan, September (mimeo).

Henderson, J.V., 1974. The types and size of cities. American Economic Review 64, 640–656.

Henderson, J.V., 1987. Systems of Cities and Inter-City Trade. In: Pierre Hansen et al. (Eds.), Systems of Cities and Facility Location. Harwood Academic Publishers, Chur, Switzerland, pp. 71–119.

Henderson, J.V., 1988. Urban Development: Theory, Fact and Illusion. Oxford University Press, Oxford.

Henderson, J.V., Ioannides, Y.M., 1981. Aspects of growth in a system of cities. Journal of Urban Economics 10, 117–139.

Ioannides, Y.M., 1994. Product differentiation and economic growth in a system of cities. Regional Science and Urban Economics 24, 461–484.

Ioannides, Y.M., Overman, H.G., 2000a. Zipf's Law for Cities: An Empirical Examination, May, working paper, Tufts University; http://ase.tufts.edu/econ/papers/index.html.

Ioannides, Y.M., Overman, H.G., 2000b. Spatial Evolution of the US Urban System, August; working paper, Tufts University, `http://ase.tufts.edu/econ/papers/index.html`.
Kim, S., 1995. Expansion of markets and the geographic distribution of economic activities: the trends in U.S. manufacturing structure, 1860–1987. Quarterly Journal of Economics 95, 881–908.
Knox, P., 1994. Urbanization: An Introduction to Urban Geography. Prentice-Hall, Englewood Cliffs, New Jersey.
Krugman, P., 1991. History versus expectations. Quarterly Journal of Economics 106 (2), 651–668.
Krugman, P., 1996a. The Self-Organizing Economy. Blackwell Publishers, Oxford, UK and Cambridge, Massachusetts.
Krugman, P., 1996b. Confronting the mystery of urban hierarchy. Journal of the Japanese and the International Economies 10, 399–418.
Marshall, J.U., 1989. The Structure of Urban Systems. University of Toronto Press, Toronto.
Noyelle, T.J., Stanback, Jr. T.M., 1984. The Economic Transformation of American Cities. Rowman and Allanheld, Totowa, New Jersey.
Pickard, J.P., 1959. Metropolitanization of the United States. The Urban Land Institute, Washington, D.C.
Pred, A.R., 1966. The Spatial Dynamics of U.S. Urban-Industrial Growth, 1800–1914: Interpretive and Theoretical Essays. MIT Press, Cambridge, Massachusetts.
Quah, D.T., 1996. Regional convergence clusters across Europe. European Economic Review 40, 951–958.
Simon, H., 1955. On a class of skew distribution functions. Biometrika 44, 425–440.
Vance, Jr. J.E., 1970. The Merchants' World: The Geography of Wholesaling. Prentice-Hall, Englewood Cliffs, New Jersey.