Journal of Economic Geography **3** (2003) pp. 343–372

Urban evolution in the USA

Duncan Black and Vernon Henderson***

Abstract

On a sustained basis, cities are of non-uniform relative sizes. This paper addresses three basic issues which arise from this simple observation by examining the size distribution of US cities over the period 1900–1990. First, we explore the reasons why there is a wide distribution of city sizes. Second, we characterize the evolution of the size distribution of cities, documenting growth in sizes and numbers of cities. We ask whether the relative size distribution of cities has remained stable over time, or if it has displayed, instead, a tendency to collapse, flatten, or otherwise change its shape. We also examine evidence on whether the size distribution obeys Zipf's Law. Third, we examine the degree and determinants of mobility of individual cities within this distribution, asking to what extent cities are moving up and down in the distribution and how this movement is influenced by cities' geographic characteristics. We use a newly constructed data with consistent metropolitan area definitions over this century, discussing the issues and linking our results to the relevant literature.

Keywords: city size distribution, urban concentration, urban growth, urban specialization, Markov processes
JEL classification: O0, R0
Date submitted: 7 May 2002 **Date accepted:** 12 March 2003

1. Introduction

The reasons for the existence of a non-degenerate distribution of city sizes follow from standard ideas in the urban literature. Firms concentrate geographically in order to exploit benefits of agglomeration which are traded off against urban diseconomies (e.g. commuting costs) in determining city sizes. Cities may be highly specialized in production due to external scale economies arising from intra-industry spillovers (Henderson, 1974), or more diversified with firms benefiting from either Jacobs-type externalities or upstream—downstream linkages (Fujita et al., 1999). In either case, theory predicts that cities will have varying industrial compositions and sizes, as the tradeoff between the costs and benefits of agglomeration is resolved at different population levels for different city types, reflecting the varying benefits of agglomeration across industries and locations. In this paper, we use clustering algorithms to group cities into types according to their production patterns, show that these patterns differ significantly across these types, and demonstrate that there is indeed a relationship between city types and sizes.

* Department of Economics, University of California–Irvine, USA.
email <DBlack@brynmawr.edu>
** Corresponding author at: Economics Department, Brown University, Providence, RI 02912, USA.
email <Vernon.Henderson@Brown.edu>

How should we expect city size distributions to evolve over time? In a related paper Black and Henderson (1999b), we model urban evolution, building on work by Eaton and Eckstein (1997) to develop an endogenous growth model in a system of different types of cities. With national population growth and human capital accumulation, city sizes and perhaps numbers grow. In this model of a closed economy without new product development, growth is 'parallel' across cities, which maintains a constant relative size distribution. However, product cycle considerations, changes in export/import behavior, and corresponding changes in national output composition suggest a potential basis for increasing concentration in recent decades, something we find in the data. Under an urban product-cycle hypothesis, product development occurs in major metro areas with their diverse industrial-occupational environments. Once products and production processes are standardized, production is decentralized to smaller, perhaps more specialized cities with their lower wage and land costs (Hekman, 1982; Fujita and Ishii, 1994; Duranton and Puga, 2000). In the US, this pattern has been altered in recent decades as is evident from the relative growth in imports and decline in manufacturing. Standardized production is now decentralized not so much to smaller US cities, but more to other countries. The overall composition of US production has shifted toward the service sector—services in finance, research, management and consulting, law and education, engineering and architecture, and business (advertising, credit, computer, personnel, etc.). As we will see later in the paper, a number of these activities tend to be disproportionately located in large metro areas (Kolko, 1999), enabling increased urban concentration.

In examining the evolution of the US city size distribution, we begin by revisiting the traditional rank-size rule and Pareto distribution approaches (Hamilton and Mills, 1994) as updated by (Gabaix, 1999a,b). We then proceed to build on work which is based on less restrictive assumptions about underlying stochastic processes and which uses a more non-parametric approach (Eaton and Eckstein, 1997; Dobkins and Ioannides, 2000; Ioannides and Overman, 2001). By modelling the transition process of cities directly, we can examine the evolution of and trends in the city size distribution and address the third main topic in this paper—the degree and determinants of size mobility of cities within this evolving distribution.

The literatures on economic geography (Krugman, 1993; Fujita et al., 1999) and on scale economies (Arthur, 1990) reach specific conclusions about the time patterns of the size mobility of cities. In particular, Arthur develops a firm-location matching model of the evolution of agglomeration, which has a direct adaptation to cities. The model characterizes the tension in location choices between the scale benefits of historically evolved concentrations and improving matches of new heterogeneous economic agents to heterogeneous site attributes, as an economy grows and location patterns emerge. Arthur predicts that as an economy grows, local relative employment fluctuations will dampen, and locational patterns as measured by locational shares of national employment or population will become fixed.

In opposition to this is empirical work of Davis et al. (1996) who postulate that locations experience on-going allocative shocks, which effectively maintain turbulence in the system and lead to continuous shifts in locational patterns. National aggregate shocks may have differing effects across industries. If cities are fairly specialized in production, national sector shocks will affect cities according to their current specialization. Even without sector-specific shocks or urban specialization, if shocks are randomly distributed across firms spatially, some locations will be hurt (or helped) more than others by the accident of the random patterns of firm hits within and across cities. Temporary shocks would be expected to lead to ongoing mobility within the size distribution of cities. More permanent shocks to national output composition of the type discussed above might lead to changes in the overall shape of the distribution.

We test directly whether fluctuations in relative city populations have dampened over time, by examining whether transition speeds have slowed. The first half of the century is characterized by rapid urbanization as well as high national population growth (not to mention a major depression, two world wars, and massive interregional migration). In recent decades, urbanization rates have stagnated, national population growth has slowed, and interregional migration rates may have dampened. Has the degree of city-size mobility also slowed or are sources and effects of turbulence ongoing?

In addition to exploring the degree of size mobility over time, it is of interest to examine the effects of observed heterogeneity of city site characteristics on mobility. The literature on economic geography and scale economies stresses the role of physical and economic geography in influencing city size, stressing three aspects of geography: (a) first nature attributes of landscape (coast, mountains), weather, and raw resource deposits, (b) the role of neighbors, and (c) internal historically developed aspects of cities whether physical (i.e., transport infrastructure) or human (culture, laws, and traditions). For neighbors, a key variable in theoretical work Fujita et al. (1999) and empirical work (Davis and Weinstein, 1997) concerns market potential. More and bigger neighbors nearer to a city enhance its growth by providing markets for its products. Additionally, neighbors are a source of information spillovers, which is the heart of Eaton and Eckstein's work. However, neighbors are also competitors in regional markets and congest the landscape. In location models, competitor cities/firms may gain in separating themselves—by differentiating their locations, they differentiate themselves in regional markets. Hierarchy models picture location patterns where the biggest cities are furthest apart, while the smallest locate next to the biggest (Beckmann, 1968). Empirically we will argue that the neighbor relationship is complex and an increase in market potential may not always help city growth.

In Section 2 we provide an overview of the evolution of the US system from 1900–90, looking at changes in the relative size distribution and mobility of cities through the size distribution. In Section 3, we examine heterogeneity of urban sites and the effect on urban growth of geography. In Section 4, we look at urban production specialization and its relationship to relative city sizes.

To aid the reader in following the paper, we preview our key findings. While urban and population growth involve on-going increases in city numbers, they are accompanied by much higher rates of population growth of existing cities, due to technological change. Over the century there has been a stable wide relative size distribution of cities which is supported by different types of cities with different industrial compositions. The size distribution has exhibited some increasing relative concentration, which as noted above, we associate with the recent relative growth of services in the US economy. In terms of mobility of cities through the size distribution, there has been a stable transition process over the century, in opposition to Arthur's prediction. However, one notable aspect of the transition process is that the largest cities exhibit minimal downward mobility. In addition to the general downward immobility of large cities, the greater market potential, or regional market sizes in the Northeast and Midwest helps explain why existing large metro areas in those regions have stayed large, even though development of new large metro areas is in warm, drier coastal areas, particularly in the South and West.

2. Evolution of the US urban system 1900–1990

This section examines growth in the US urban system, changes in the relative size distribution of cities, and mobility of cities through the distribution, over a 90-year period. In that time, the US population rose by 225% from 75 to 244 million. In 1900 the modal and median size US county had 17,000–18,000 population. In 1990, the mode was identical and the median had risen by under 50%. The 225% increase in total US population was accommodated by a bulge in the right tail of the county population distribution, representing urbanization and overall geographic concentration. The percent urbanized rose steadily from 40.5% in 1900 to 70.0% in 1960 and then edged up to 75.1% in 1990. Over the 90-year period the urban population thus rose by 500%.

Within this geography, we focus on a special entity—what we conceive of as major cities, or essentially permanent metropolitan areas. Of the geographic units with any urban population in any decade (metro areas and non-metro urban counties), we will examine only about 20% of those, accounting for 84–89% of the total urbanized population. Conceptually these are established urban areas, each with high population density and significant overall relative size and with economic activity traditionally agglomerated around a major central city. Cities are geographic entities with sufficient size and density to possess costly congestion effects, which distinguishes them from other areas. This metro area portion of the US represents the world of urban growth that we model elsewhere and we have complete data on it.

2.1. The data

To explore these issues, we want to construct a data set with metropolitan areas defined consistently over the century. This is rather different than the approach in Dobkins and Ioannides (2000) and Ioannides and Overman (2000, 2001). Those authors define metro areas based on the definitions in existence in each particular decade, which vary enormously from decade to decade as the views on what is a metro area have changed. The result is considerable inconsistency across decades. To try to construct consistently defined metro areas, we begin by utilizing the decennial Population Census to obtain figures for total and urban populations by county. The metropolitan area definitions that we use are based on county boundaries. A metro area can consist of either a single county or a collection of contiguous counties. Approximately half are composed of a single county. Four basic problems arise in defining these entities consistently over time.

The first problem is that county definitions have changed over time, apart from minor boundary redrawings. Historical counties have split into two or more modern counties and historical counties have been combined into single modern counties, with most changes occurring prior to 1940. We recombine all splits, either modern or historical, to achieve a set of 'common denominator counties', with a goal of defining consistently over time a set of metro areas. Doing so reduces the 3043 modern counties to 2708 in our data set. In terms of metro areas this reconsolidation reduces the sample of 317 metro areas in the continental US in 1990 covering 742 modern counties to 282 metro areas covering 695 consolidated counties.[1] MSAs are dropped when reconsolidation leaves us unable to distinguish modern MSAs; this process removes all of Florida from the data. In

1 We drop Hawaii and Alaska, and one oddity, a 1990 'MSA' with well under 50,000 urban residents, from an original set of 320 metro areas. In terms of MSA definitions consolidated metro areas (CMSAs) are split into

an earlier version of this paper, Black and Henderson (1998), we show that, for the period 1940–1990, our results for the full sample of 317 metro areas present in 1990 are very close to those for the working sample of 282 metro areas, used in this paper.

The second problem concerns how to define a sample of metro areas in decades other than 1990. Contemporaneous definitions do not work. The Census Bureau did not start using metro area concepts until 1950 and metro area definitions have changed significantly since then.[2] We do not have the detailed incorporated and unincorporated place data to construct consistent definitions over time. Instead we follow Donald Bogue's strategy of applying current land area definitions to prior decades. We follow the component counties of the metro areas in existence in 1990 back in time to 1900. We assume, in essence, that once formed metro areas are permanent. We use urban population numbers for these counties, combining them to form the set of metro areas in each decade.

The third problem we face is how to determine when an area qualifies to enter the sample of cities in each decade. In 1900, Phoenix has zero urban population, so it is not a city. When does it become one? Even at a single point in time the choice of cutoff is subjective. This problem is compounded when looking across the entire century. A standard approach is to use an absolute cutoff point by assuming that an urban area qualifies only as a metro area in any decade if it has 50,000 urban residents, a threshold which matches 1990 definitions. However, it appears that technological progress affecting the relative costs (e.g. improvements in commuting technology) and benefits (arising from human capital accumulation, as in Black and Henderson, 1999a) of urban agglomeration has led to an increase in the size of a typical city. A city which would have been considered of medium size in 1900 would be thought of as little more than a town by 1990 standards. This implies that we should use a threshold that changes over time.

The key is to choose a minimum population cutoff point in each decade, above which an area is included in our sample in that period. We want to use a cutoff that reflects this growth in typical city sizes. For our basic sample, we select a relative cutoff point which is equal to the 1990 ratio of the minimum to the mean MSA urban population, or $52262/527508 = 0.099$. We apply this criterion to every decade, selecting all cities above this cutoff as the set of MSAs in that decade.[3] Using this cut-off point in defining metro areas,

primary metro areas (PMSAs). Thus the New York consolidated metro area of almost 20 million is split into 15 PMSAs. Consolidation is a recent Census practice. In part it derives from declining urban population densities and urban sprawl, such that traditional spatially separated metro areas start to overlap at the fringes. For most of the century most current PMSAs in a CMSA were spatially separated. Since we follow metro areas over time, it would be implausible to define consolidated metro areas for most decades. Even today we would argue most PMSAs retain their own industrial bases and labor markets. Also we are going to argue that relative urban concentration has increased over time, and we want to ensure arguments are not based on definitional consolidations.

2 Since 1950, definitions have been changed from a primary focus on political city population measures to density measure approaches in order to accommodate significantly populated urbanized area. This has resulted from the changing character of urban areas, especially in the West.

3 In selecting each decade's cut-off point for the base 282 metro areas in 1990, we rank these cities by size from highest, 1, to lowest, 282, in each decade. We then choose the first s cities as our city set for that decade, such than

$$\min\left\{ s; N_{s+1} \Big/ \left(\sum_{i=1}^{s+1} N_i / s+1 \right) \le 0.099 \right\}.$$

Given the LHS of the equality is a non-monotonic function of s, we check in each decade that there is a unique cross-over point. That is violated once only, in 1920, where, if we chose the max s instead of the min s, we would have added two cities to our sample for that decade.

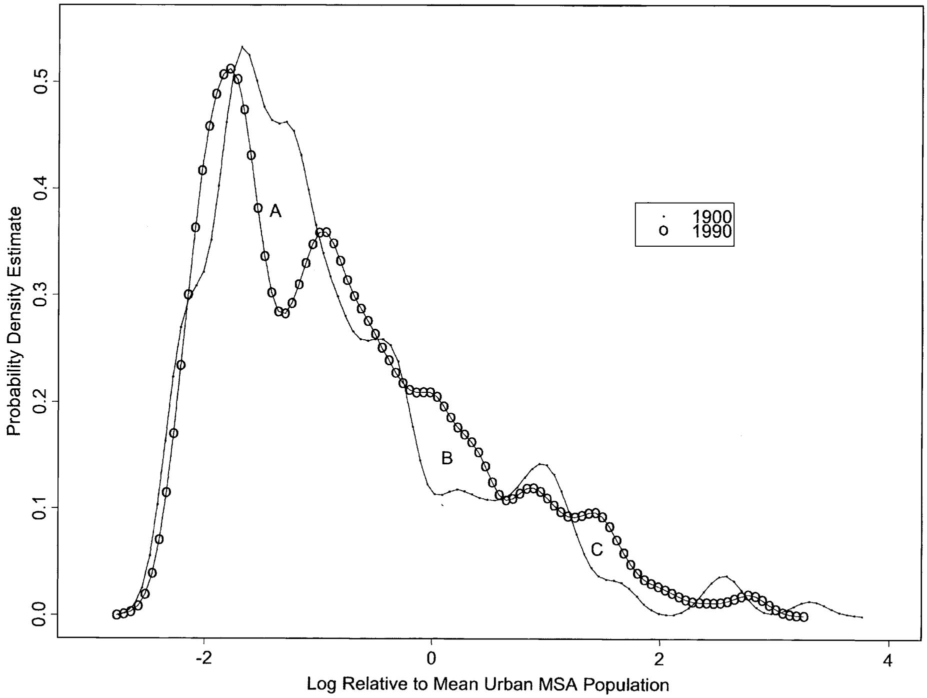

Figure 1. Density functions for MSA size distributions

from 282 metro areas in 1990 with a sample mean size of 527,508 we obtain a set of 194 metro areas in 1900 with a sample mean size of 129,596. These metro areas account for 89 and 94% of the total US urban populations in 1990 and 1900 respectively. As illustrated in Black and Henderson (1999a), results are very similar if we use a harsher cut-off point (minimum 1990 size of 86,850 with relative cut-off point of 0.136 and 148 metro areas in 1900).

In addition, to examine the evolution of the distribution over time we want to normalize for this growth in typical metro area sizes. The distribution is shifting to the right as mean city sizes increase. We therefore look at relative size distributions, with sizes in any decade normalized by dividing by the average metro area size.[4] The impact of the normalization and cut-off point selection in defining metro areas and their size distribution is illustrated in Figure 1. This displays an estimate of the density function of the normalized city size distributions for 1990 and 1900.[5] The relative distributions are quite similar. The main difference seems to be the 1990 loss of density relative to 1900 in the area marked *A* in lower size ranges, with this loss added in the upper size ranges in 1990 in the areas marked *B* and *C*. This pictured shift to larger relative sizes will be analysed

4 Note we do not want to normalize by total national population to obtain city population shares as the unit of observation. Since the number of cities increases over time, typical shares must decline steadily, giving a false impression of deconcentration.

5 Figure 1 was calculated and drawn in S-Plus, using a Gaussian kernel estimates evaluated at 100 points with a window width of 0.6.

below. Regardless there is remarkable stability in the distribution over 90 years of rapid technological advance, urbanization, and population growth.

The final problem in defining metro area sizes concerns whether to use urban or total population of MSAs as the size measure. Using total population is problematic because of changes in rural–urban composition over time. In 1990 while about 86% of all population of metro area counties is urban, in 1900 the figure is about 65% reflecting changes in rural–urban compositions of counties. Unfortunately, the definition of urban itself has changed over time, with the main definitional shift occurring in 1950 with the treatment of unincorporated places. We can't reconstruct definitions, so we have no choice but to use the contemporaneous definition in each decade. We believe the modifications are minor, based on Bogue's (1953) analysis comparing 1950 with 1940 definitions and based on the notion that prior to 1940–50 heavy suburbanization and urban settlements beyond urban places were relatively unimportant. Thus, though we fix the land areas of MSAs over time, we focus only on the urban population within those counties.

2.2. Modeling changes in size distributions and mobility of cities

To characterize the spatial evolution of the urban system, we examine the relative size distribution of local urban populations across existing metro areas in 1900 and analyse how that distribution changes by decade to 1990. We start by applying Zipf's Law to the 20th century US city size distribution. We then turn to a more general approach to examining the evolution of the size distribution of cities which accounts for entry of new cities and a general transition process of existing cities through the size distribution. Examining evolution in this fashion allows us (1) to test for stationarity vs acceleration or dampening in mobility rates over time, (2) to quantify mobility of cities in terms of mean first passage times, and (3) to characterize the evolution of the relative size distribution to determine if it is collapsing (to a common city size), going bimodal, or reconcentrating at upper (mega-cities) or lower ('manageable' communities) ends. With our findings we can then address the key questions posed at the beginning of the paper.

We start in Table 1 with some basic facts about our data. Then we turn to an analysis of the evolution of the size distribution and of city mobility. Column (1) shows that except for the Depression, the growth rate in average sizes of existing cities in the base year far exceeds the growth rate in numbers of new cities. Growth in sizes and numbers of cities has been fairly steady over the decades, but with growth rates in sizes slowing in recent decades. Black and Henderson (1999b) show that average city sizes grow with technological change, or local knowledge accumulation, accentuating the scale externalities of cities. In that paper, MSA and time fixed effects regressions show MSA sizes being strongly related to local educational attainment. Column (3) indicates few cities ever experience the type of large population declines, envisioned in Krugman's bifurcation model, (Krugman, 1996) of city formation whereby formation involves 'oversize' cities catastrophically splitting into two or more parts.[6] The largest decline is 24% for

6 We report declines in total, rather than urban population. The shifting definitions of urban versus rural combined with issues in defining cities in Virginia produce about a half dozen anomalies in the urban population data that don't affect total county populations in early years. For the record we reran basic results in Table 5 below excluding five anomalies, with absolutely no noticeable impact on results. The five dropped MSAs have the following characteristics: their urban and total populations move in opposite directions in one Census period and the gap exceeds 40 percentage points.

Table 1. Numbers and sizes of cities

	Inter-decade growth in average sizes (urban populations) of metro areas (%)	Inter-decade growth in numbers of metro areas (%)	Numbers of existing metro areas where total pop. declines exceed 10%
1900–10	36.3	1.5	1
1910–20	27.3	4.6	2
1920–30	27.1	4.4	1
1930–40	5.4	8.4	0
1940–50	29.1	6.0	1
1950–60	29.2	4.8	0
1960–70	18.9	3.1	0
1970–80	9.4	5.6	0
1980–90	11.4	0	5

	Average size	Median size	No. of metro areas	Minimum size	Maximum size
1900 level	129,596	39,568	194	13,103	3,576,826
1990 level	527,508	198,591	282	52,262	8,786,271

Buchanan, Missouri (the river town of St Joseph) in 1900–10. The decade with the largest number of declines is 1980–90, as Midwestern cities adjusted to the precipitous decline in manufacturing nationally.

2.2.1. Zipf's Law, or the rank-size rule

We begin our exploration of the evolution of the US city size distribution by revisiting Zipf's Law, or the rank-size rule. A variety of authors have focused on and attempted to explain this regularity, perceived to be consistent across countries and time periods, in the relationship between the rank and sizes of cities (see in particular Ioannides and Overman, 2000; as well as Rosen and Resnic, 1980; Krugman, 1996; Eaton and Eckstein, 1997; Gabaix, 1999a; Dobkins and Ioannides, 2000). Begin by ranking cities existing in each decade by size from 1 (largest) to m (smallest) to obtain rank $R(n)$ for city size n. According to Zipf's Law, the distribution of cities should follow a Pareto distribution, with $R(n) = An^{-a}$.[7] Empirically, the approach is to estimate the following equation for each decade:

$$\ln R(n_{it}) = \ln A_t - a_t \ln n_{it} + \epsilon_{it} \quad i = 1, \ldots, m_t \tag{1}$$

for city i in time t. The rank size-rule is obtained for $a_t = 1$, so $R(n) \cdot n$ is the same value for every city. The usual hypothesis is that a_t has declined over time in the US, starting near one with the rank-size rule holding exactly in 1900, and now well below one (e.g. Hamilton and Mills, 1994, pp.69–77). The notion is that the rank-size line (apart from shifting upwards) has rotated counter-clockwise over time, and that represents

7 Specifically, Zipf's Law states that $P(n > N) = AN^{-a}$, where $P(n > N)$ is the probability that a city has a size greater than N.

increasing urban concentration—relatively higher ranks (lower in the pecking order) for traditional larger cities.

Table 2, Part 1, column (a) presents our estimates of a_t for our MSA sample in each decade. The Pareto parameter in all years is around 0.85, far from 1. In some later years (1970 and 1990), a_t is lower than in some earlier years (1910–30) lending modest support to the view of increasing urban concentration in recent decades. The estimate of a_t is sensitive to the choice of sample size. In fact, other work typically estimates the Pareto parameter for just the larger cities in an economy, due to data limitations. To facilitate comparisons, we re-estimated the relationship for the top one-third in each decade. Results are in column (a) of Part 2 of Table 2. Now a_t is larger in every decade, starting at a value not significantly different than 1, exceeding 1 by 1920, and then rising to quite high levels (1.18). For this top one-third of cities, the rise in a_t would suggest decreasing urban concentration in the US over time, in contradiction of the conventional wisdom. But the fact that a_t in any decade is much larger for the top one-third of cities than for the whole sample suggests that the relationship in equation (1) is not log-linear.[8] That the rank size equation is not exactly linear in logs can be seen in Figure 2, which plots the log of rank against log urban population for 1990. Instead, a quadratic relationship is evident, as is estimated in column (b) of Parts 1 and 2 of Table 2.[9] Note the quadratic relationship holds throughout the size distribution, just as pictured in Figure 2, holding for the top one-third of cities (and holding, in all decades, for the top 100 cities in the decade). As Gabaix (1999b) hints, a non-linear relationship is likely to emerge if bigger and smaller cities are subject to different relative shocks, with the variance of shocks decreasing in city size. We next turn to a test of Gibrat's Law.

Gabaix (1999b) shows Zipf's Law is an outcome of Gibrat's Law, a stochastic process in which city i's share of national urban population in period $t+1$ is that in period t multiplied by γ^i_{t+1}, where the γ^i_{t+1} are identically and independently distributed across cities and time. This imposes the same relative shocks on big vs small cities; growth rates are uncorrelated with city sizes. As noted by Clark and Stabler (1991), who focus on a small sample of Canadian cities, testing for Gibrat's Law is equivalent to testing for unit roots in the evolution of city sizes. We estimate the following specification:[10]

$$\ln(n_{it+1}) - \ln(n_{it}) = \alpha + \delta_t + \gamma \ln(N_{it}) + \epsilon_{it}. \tag{2}$$

We test the null hypothesis, implied by Gibrat's Law, that $\gamma = 0$. Against the null, we are testing whether there is, say, mean reversion in the stochastic growth process, so that $\gamma < 0$. Note, this null hypothesis does not permit an auto regressive process to the ϵ_{it}, so simple OLS estimation suffices. Results are in column (1) of Table 3. For the overall sample in column (1), the specification in (2) is rejected. The process exhibits significant mean reversion (consistent with our hypothesized Markov process). Note the t-statistic for the basic test in column (1) is 7.60, far in excess of the value required in panel data to

8 Using an absolute cut-off point of 50,000 for the definition of metro area leads to fairly steadily declining estimates of a_t over the century, from −0.971 in 1900 to −0.842 in 1990. While, except for 1900 all coefficients are significantly less than one, there is a distinct pattern of decline in a_t over the century, using this sample criterion.

9 The specification for column (b) is $\ln R(n_{it}) = \ln A_t - a_t \ln n_{it} + b_t (\ln n_{it})^2 + \epsilon_{it}$.

10 While the specific formulation in Gabaix (1999a) concerns growth rates in city shares of national urban population (rather than relative sizes), the LHS reduces to that in (2), with a different interpretation of δ_t.

Table 2. Parametric distribution of city size (biased estimates of standard errors in parentheses)

	Part 1—full sample					
	(a) Pareto parameter			(b) Quadratic rank size equation		
	a_t	R^2	m_t	a_t	b_t	R^2
1900	0.861	0.981	194	−0.066	0.654	0.993
	(0.0087)			(0.0035)	(0.0803)	
1910	0.873	0.979	197	−0.072	0.841	0.993
	(0.0093)			(0.0037)	(0.088)	
1920	0.875	0.979	206	−0.081	1.08	0.996
	(0.0090)			(0.0026)	(0.062)	
1930	0.870	0.979	215	−0.080	1.08	0.997
	(0.0087)			(0.0022)	(0.054)	
1940	0.865	0.978	233	−0.078	1.04	0.997
	(0.0085)			(0.0022)	(0.055)	
1950	0.870	0.969	247	−0.099	1.59	0.996
	(0.0099)			(0.0021)	(0.059)	
1960	0.852	0.961	259	−0.116	2.06	0.996
	(0.011)			(0.0023)	(0.058)	
1970	0.836	0.957	267	−0.124	2.31	0.996
	(0.011)			(0.0025)	(0.063)	
1980	0.858	0.957	282	−0.133	2.54	0.995
	(0.011)			(0.0027)	(0.068)	
1990	0.842	0.952	282	−0.137	2.70	0.995
	(0.011)			(0.0027)	(0.070)	

	Part 2—Cities in top 1/3 of size distribution					
	(a) Pareto parameter			(b) Quadratic rank size equation		
	a_t	R^2	m_t	a_t	b_t	R^2
1900	1.01	0.975	64	−0.598	0.063	0.980
	(0.021)			(0.386)	(0.015)	
1910	1.03	0.971	65	−0.859	0.073	0.978
	(0.022)			(0.406)	(0.016)	
1920	1.07	0.986	68	−0.672	0.066	0.991
	(0.016)			(0.285)	(0.011)	
1930	1.07	0.989	71	−0.394	0.054	0.993
	(0.013)			(0.259)	(0.010)	
1940	1.05	0.985	77	−0.627	0.062	0.990
	(0.014)			(0.268)	(0.010)	
1950	1.11	0.984	82	−0.797	0.070	0.989
	(0.016)			(0.309)	(0.011)	
1960	1.14	0.976	86	−2.11	0.118	0.989
	(0.019)			(0.339)	(0.013)	
1970	1.14	0.965	88	−3.71	0.171	0.989
	(0.023)			(0.334)	(0.012)	
1980	1.18	0.966	93	−4.25	0.195	0.991
	(0.022)			(0.348)	(0.012)	
1990	1.18	0.964	93	−5.01	0.220	0.992
	(0.024)			(0.333)	(0.012)	

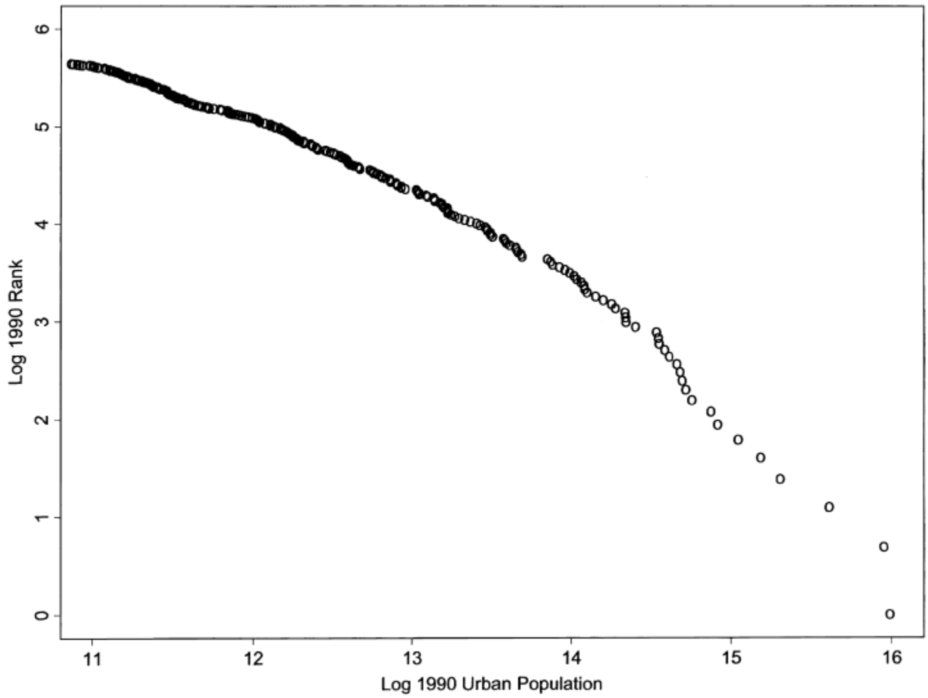

Figure 2. Plot of rank and city size, 1990

Table 3. Test of Gibrat's Law (SE in parentheses: Huber errors)

	(1) Overall sample	(2) Top 50 cities losers retained	(3) Top 1/3 cities losers retained
$\ln(N_{it})$	−0.022**	−0.039**	−0.031**
	(0.0034)	(0.011)	(0.008)
Time fixed effects	Yes	Yes	Yes
Sample size	2080	554	838
Adj R^2	0.193	0.208	0.203

compensate for the downward bias in estimation with finite samples (Dickey and Fuller, 1981; Levin and Lin's, 1992 critical t-value for panel data in this context is under 2.1).

This test does suffer from a serious potential flaw. In testing for mean reversion, we have used a sample of 'winners', so we are potentially biasing against it. Besides metro areas in 1900, added to the sample in each decade are counties in that grew into metro areas after 1900, but not those potential urban counties that did not. Our sampling method is appropriate for pickng metro areas out of the USA geography, given once a county qualifies as a metro area, in the data it does not exit. However, in testing for Gibrat's Law per se, we need to address the problem of over-sampling of winners. We conducted several experiments, all of which continued to reject Gibrat's Law. We report on two which focus on Gabaix's application to the upper tail of cities. For the first, we start just the top 50 cities in 1900, keeping those cities in our sample in all decades including those that declined severely in relative terms. In each subsequent decade, any city which newly enters the top 50 we add permanently to the sample. Besides winners, we keep permanently all losers who drop out of the top 50. Once added, cities never exit the sample even if they exit the top 50. Results are reported in column (2) of Table 3. In column (3), we repeat this procedure for the top one-third of cities in each decade, reporting the results for that as well.

The coefficients in columns (2) and (3) are actually larger than in column (1). Gibrat's Law is rejected for these samples of large winning and losing cities. While adding in losers works to lower the absolute value of γ, that value tends to be larger in any samples drawn from the upper tail of the city size distribution. For the top 50 cities, not retaining losers, it is -0.092! Other experiments such as adding in regional dummy variables, adding in MSA fixed effects, removing time fixed effects, and trying other winner and loser samples,[11] all reject Gibrat's Law, sometimes dramatically (e.g. no time fixed effects or adding in MSA fixed effects).

While Zipf's and Gibrat's Laws may give good first order descriptions of the data, they do not hold in any precise form for our sample of cities. We therefore turn to a more non-parametric approach to examining the evolution of the city size distribution. This framework will allow us to also examine the mobility of cities through the size distribution as well as to model the effects of geographic heterogeneity of cities.

2.2.2. Markov processes

We assume that distributions evolve over time according to a homogeneous stationary first-order Markov process from 1900–1990, accounting for entry. Stationarity is tested for and homogeneity will be relaxed in the next section. One can object to the Markov structure which implies that every city will at some point occupy each cell of the distribution. However, it is a convenient way to model fluidity and we do demonstrate the presence of substantial intra-distribution mobility. Who in 1900 would have predicted Phoenix would grow into a major metro area in our top size cell in 1990? And, can it be claimed that 1,000 years from now Champaign-Urbana couldn't possibly be a major metro area?

Our approach draws upon Dobkins and Ioannides (2000) and Eaton and Eckstein (1997), and we obtain some similar and some different results for our sample of conisistently defined US metro areas. We divide the relative size distribution of cities at a point in time into discrete cells, with cut-off points defined by relative sizes, $N_{it}/\overline{N}_t$ where $\overline{N}_t = \sum_i^{m_t} N_{it}/m_t$. Table 5 lists the upper cut-off points of cells 0.22, 0.47, 0.83, 2.2 and open, which in 1900 correspond to cell shares starting from the bottom of 35, 30, 15, 10, and 10% of the total number of cities. f_t denotes the vector of distributional shares for each cell and this distribution evolves over time according to

$$f_{t+1} = (1 - i_t)M_t f_t + i_t Z_t \tag{3}$$

M_t is the transition matrix of existing cities from time t to $t+1$. With a homogeneous, stationary process M_t equals a common M (see Table 4). i_t is the overall net entry rate of

11 In the sample in column (1) of Table 3, we tried adding to the sample all counties in a decade that met our relative cut-off point in terms of size to be a metro area in that decade. This adds in 'borderline' counties, which never grew into modern metro areas. For that larger sample of 2510, $\gamma = -0.0119$ with a standard error of 0.0036, which rejects $\gamma = 0$.

Table 4. The transition matrix (true (binomial) SE in parentheses)

		Cell in $t+1$					Decade net entry rate
		1	2	3	4	5	
	1	0.848 (0.0138)	0.151 (0.0148)	0	0.0015 (0.0022)	0	0.0414
	2	0.082 (0.0105)	0.797 (0.0166)	0.116 (0.0185)	0.0051 (0.0041)	0	0.00048
Cell in t	3	0	0.077 (0.011)	0.793 (0.0234)	0.130 (0.0193)	0	0
	4	0.0033 (0.0022)	0	0.013 (0.00656)	0.918 (0.0157)	0.066 (0.0172)	0
	5	0	0	0	0.034 (0.0103)	0.966 (0.0125)	0

Table 5. Projecting city size distributions based on 1900–90 evolutionary patterns (shares of cells in numbers of MSAs)

	(1) Cell cut-off point	(2) 1900 actual	(3) 1990 actual	(4) 1990 predicted	(5) Steady state*
f_1	0.223	34.5	35.1	33.7	32.2
f_2	0.472	30.4	24.8	25.9	23.5
f_3	0.832	14.9	13.5	13.4	11.8
f_4	2.22	9.8	15.6	16.0	17.6
f_5	Open	10.3	11.0	10.9	14.9

Note: * 95% confidence interval error bands for the five states are respectively 27.7–39.1, 20.7–28.4, 09.6–16.0, 12.4–22.5, and 08.2–17.4. These take as given our set of cities in each decade and are based on sampling with replacement to generate repeated samples to recalculate the transition process used to project the steady state.

cities between t and $t+1$: the total increase in number of cities divided by the total at $t+1$. Z_t is the distribution of entrants across the $(t+1)$ cells, almost exclusively in the lowest cell. This use of a discrete distribution differs from the continuous distribution approach in Ioannides and Overman (2000), but hypothesis testing is difficult when applying kernel estimation methods to the data.

If we start in 1900 with f_{00}, the evolution to τ decades after 1900 with constant M, i and Z is

$$f_{00+\tau} = (1-i)^{\tau} M^{\tau} f_{00} + [I-(1-i)M]^{-1}[I-((1-i)M)^{\tau}]iZ. \tag{4}$$

One can carry this to the extreme and calculate the steady-state distribution, f, where

$$f = [I-(1-i)M]^{-1} iZ \tag{5}$$

equations (4) and (5) are convenient ways of summarizing what tendencies towards future concentration the US city size distribution is displaying.[12] To do so below, we will

12 Strictly speaking we can't model the steady state in equation (5), with our data, since the calculation presumes indefinite steady-state entry of cities. Given a finite number of sites (counties), unending entry is impossible, although we note it would take five centuries to exhaust all remaining non-metro counties at this century's rate of entry.

use the calculated M and iZ for 1900–1990, testing for stationarity of the M matrix and using the average iZ. To get a sense of speed of transition of existing cities, we calculate mean first passage times, for a process starting at time zero. If ϕ_{jk}^t is the probability that a city in state (cell) j next visits state k at a time t decades later, then the mean first passage time, the expected time until the city next visits that state (in decades), fp_{jk}, from j to k, is

$$fp_{jk} = \sum_{t=1}^{\infty} t\phi_{jk}^t. \tag{6}$$

In our calculations fp_{jk} will be bounded. We define $[M^t]_{jk}$ as the j,k element of the transition matrix raised to the power t. Then for ϕ_{jk}^t, Markov chain theory gives the t-period transition probabilities as

$$[M^t]_{jk} = \sum_{s=0}^{\infty} \phi_{jk}^s [M^{t-s}]_{kk} \quad \forall t \geq 1.$$

Given $\phi_{jk}^1 = [M]_{jk}$ and $\phi_{jk}^0 = 0$, we can use this to recursively define ϕ_{jk}^t as (see Karlin and Taylor, 1975)

$$\phi_{jk}^t = [M^t]_{jk} - \sum_{s=0}^{t-1} \phi_{jk}^s [M^{t-s}]_{kk} \quad \forall t \geq 1. \tag{7}$$

For both the projected future size distributions and mean first passage times, we will report bootstrapped confidence intervals below, which account for variation in transition probabilities.

2.2.3. Evolution of the size distribution

In Table 5, we examine the predicted evolution of the size distribution of cities in the US. Columns (2) and (3) give the shares for the five cells of the MSAs in 1900 and 1990. They suggest an actual increase in urban concentration from 1900 to 1990. Column (4) gives the predicted 1990 shares, which are very close to the actual distribution, indicating good predictive power of the model. To see where the process is headed we calculate the steady-state distribution in column (5), assuming on-going evolution on the current path and we footnote the 2040 predicted distribution.[13] One obvious result is that the size distribution exhibits no tendency to completely flatten (spread), to converge on a middle cell (collapse to a common city size) nor to, say, go bimodal. However there are subtle, but definite changes in concentration. In comparing 1900 with the steady state, there is an increased share of the two highest cells, at the expense of cell shares of the three lowest cells. This increasing share of the f_4 and f_5 cells indicates increasing urban concentration. We bootstrapped confidence intervals, footnoted in the table, which indicate that certainly the f_4 steady-state (or, even the 2040 share) is significantly larger than the 1900 share. We also explored whether the point estimate gain in the f_5 cell is concentrated in the highest end by looking at a six cell distribution, splitting the f_5 cell in two. For this upper tail, the initial 1900 shares of 7.2% and 3.1% grow to 7.4% and 8.4% in the steady state indicating concentrated growth in numbers of cities at the very highest end.

What does increasing relative urban concentration mean? Between 1900 and 1990, our normalization factor in calculating relative size distributions, the mean city size, increases

13 Predicted shares 50 years out to 2040 are 33.2, 25.0, 12.8, 17.0, and 11.9.

fourfold. (Note the size of the largest or, say, eight largest cities increases by only 2.4–2.5 fold in the time period.) However typical city size, the median, has grown faster than the mean, increasing fivefold. This suggests the sizes of some medium-large cities have grown relatively faster than other cities pushing themselves over a threshold of, say being twice average size, so relatively more cities crowd into the highest cells. Recall that Figure 1 displayed loss in mass at the lower end of the distribution and the gain in the middle to upper range. Rather than simple mega-city growth of the primal cities, other cities are catching up, so by 1990 the US has over 30 MSAs with one million urban residents.[14]

In Black and Henderson (1998, 1999a), we explore the sensitivity of these results to the choices of normalization factor (e.g. mean vs median), whether to use absolute or relative cutoff points for city entry, the criteria used to establish the basic 1990 sample of cities, the set of initial cell shares, and the relevant time periods. Varying these choices does affect quantitative outcomes; however the basic qualitative results are consistent across different approaches.

It is tempting to conclude that this increased relative urban concentration is due to scale economies and changes in technology. As noted earlier, we definitely agree the enormous increase in *absolute* city sizes is due to technology—endogenous growth forces of local knowledge accumulation, modeled in Black and Henderson (1999a) and technological advances in commuting reducing the costs of congested urban living. However, we believe the *relative* shift in urban concentration has less to do with technology changes directly and more to do with changes in national output composition towards inter-city traded services. We explore this topic in Section 4.

2.2.4. Mobility

In Table 4, we present maximum likelihood estimates of the transition probabilities, P_{jk}, which are simply the total number of cities moving from cell j to k over nine decades divided by the total number of cities starting in j in the nine decades. We test for stationarity of this matrix against non-stationarity for the decades 1900–90, as well as subperiods such as 1900–50 and 1950–90. Stationarity is never close to being rejected (despite some decline in the rate of entry of new cities in the last decades).[15] The matrix has several features. Only the diagonal and immediately off-diagonal transition probabilities are significant, so cities in a decade generally only move up or down one cell. Diagonal elements, the probability of staying in the starting state, are much higher in the two highest states. For relevant starting states (2–4), the probability of moving up a state exceeds that of moving down (significantly in two of the three cases) and the probability of moving up out of state 1 is much higher than moving down out of state 5. Given a

14 Ehrlich and Gyourko (1999) reach a similar conclusion by examining changes shares of urban population in deciles of the city-size distribution.

15 The χ^2 statistic is

$$-2\log\left[\Pi_t\Pi_j\Pi_k\left\{\frac{\hat{P}_{jk}}{\hat{P}_{jk}(t)}\right\}^{m_{jk}(t)}\right]$$

with $(T-1)K(K-1)$ degrees of freedom. $\hat{P}_{jk}$ is the stationary estimate, $\hat{P}_{jk}(t)$ the decade by decade estimate, $m_{jk}(t)$ the number of cities moving from j to k in t, T the total number of years and K the number of cells. The χ^2 statistic is 122.16 which with 160 degrees of freedom has a p-value of 0.988. Similar p-values apply in testing 1900–50 or 1950–90 decade by decade estimates against the $\hat{P}_{jk}$.

Table 6. Mean first passage times in decades

		First passage state				
		1	2	3	4	5
Starting state	1	41.3	10.1	29.7	32.9	52.7
	2	261.7	31.5	21.6	26.6	46.4
	3	426.6	239.1	26.8	14.7	34.5
	4	516.4	372.5	185.8	3.2	19.8
	5	545.3	402.1	215.5	29.7	1.7

distribution with cell shares that are larger in lower cells (35% and 30%) compared to upper cells (10% and 10%), these transition probability patterns are unusual. If the distribution is stable, without entry the opposite pattern would hold.[16] Our patterns are driven by the fact that there is on-going entry, almost exclusively in the lowest cell (see below). Entrants 'push' existing cities in the lowest cell forward, creating a chain reaction as some cities in the next to lowest cell are themselves pushed into a higher cell.

To help characterize mobility, we use the transition matrix in Table 4 to calculate mean first passage times. We also bootstrap to estimate confidence intervals for these mean first passage times, some of which we footnote. Mean first passage times are calculated truncated at 3,000 periods, which is sufficient for convergence.[17] Confidence intervals are based on bootstrapping for 1,000 samples. In Table 6, the diagonal elements are mean first return times, where first return involves staying in the own cell one decade or first returning after exiting in the first decade. We focus on the off-diagonal elements, the mean first passage times.

In Table 6, movement up by cities is much quicker than movement down, significantly so.[18] A city can be expected to first visit the highest state from the lowest in 500+ years, with the upward mobility of Phoenix and Miami (albeit Miami is out of our data set) as modern examples of actual moves from state 1 to state 5. Movement downwards is extremely slow. The expected time to first move even from state 2 to state 1 is 2,600 years. Remember that these calculations account for the fact that starting from state 2, a site might visit states 3, 4, or 5 before going to state 1.[19] From state 5 it is 5,500 years on average to first revisit state 1.[20] This is a testament to the immobility created

16 For example in a two cell case, if cell shares are 75 and 25% and $p_{11}=0.8$ and $p_{12}=0.2$, then if cell shares are unchanging, it must be that $p_{22}=0.4$ and $p_{21}=0.6$, so $p_{21}>p_{12}$.

17 Convergence occurs for every j,k pair in equation (5) when $\sum_{t=0}^{t}\phi_{jk}^{t}\rightarrow 1$. At $\tau=3000$, $\sum_{t}^{\tau}\phi_{jk}^{t}\rightarrow 1$ for all j,k pairs so that $\phi_{jk}^{3001}\approx 0$ and $t\phi_{jk}^{t}\rightarrow 0$ in equation (6) for $t>3000$.

18 Bootstrapping takes as given the sample cities in each decade and samples from that (with replacement) to get different sample transition matrices used to calculate mean first passage times. Fixing the sample of cities means we take as given the pattern of entry. Five percent confidence intervals are fairly wide. For example, in going from state 1 to 1, 2, 3, 4, and 5 they are respectively 17–79, 6–25, 20–70, 27–42, and 42–70; and in going from state 5 to 1, 2, 3, 4, and 5, they are respectively 218–906, 185–900, 103–793, 17–104, and 1.2–2.4. Note the lack of overlap in intervals going up versus down and note the tight band on the first return time to state 5.

19 This result of very slow declines, also increased our comfort zone with our implicit assumption of no exit of metro areas—i.e., once born as a metro area, it survives 'forever'.

20 These much slower movements down than up, in part reflect entry of new cities, 'pushing' existing cities up through the distribution. But they also reflect pure downward immobility, especially among bigger cities. For example, if we eliminate entry and calculate for just the 1900 sample of 194 metro areas a transition matrix

by accumulated scale. Big cities stay big, perhaps more immune to forces generating potentially deleterious shocks.

We also estimated the transition process allowing for heterogeneity of urban sites. Using a multinomial logit model, we estimate the effects on transition probabilities of urban site characteristics, such as weather, location, and market potential (see next section). The estimation strains our data, since there are five models (one for each row of the transition matrix) with four sets of coefficients each, with an insufficient number of transitions in some cells. In the next section we take a simpler approach; but results from the multinomial logit are reported in Black and Henderson (1998). What is of interest is the mean first passage times for superior sites (enhancing each site characteristic by 1.5 SD from its mean value), compared to inferior ones (removing 1.5 SD from the mean). Inferior sites take forever (39,000+ years) to move up from state 1 to state 5, but move down quite quickly (630 years). Superior sites move up quickly (270 years), but 'never' move down. The big stay big; even grossly inferior sites are not that quick to move down.

3. Heterogeneity across urban sites

Until now we have treated the transition process of cities as involving homogeneous units. However, it is of interest to consider the role of various factors in determining the movement of individual cities within the distribution. We conceive of two basic sources of heterogeneity. First, is first nature geography with measurable characteristics such as climate, coastal location, and access to mineral deposits. To this we add and focus on second nature geography as represented by the force of what has evolved around a city, which we measure by market potential. The form we use is normalized market potential for MSA j in time t. Market potential and normalized market potential are respectively

$$mp_j(t) = \sum_{i \neq j} \frac{\text{Pop}_i(t)}{d_{it}}.$$

$$MP_j(t) = \frac{mp_j(t)}{\overline{mp}(t)}.$$

Market potential is the sum is over all US counties excluding counties in MSA j where d_{ij} is the distance from the middle of county i to the middle of MSA_j and $\text{Pop}_i(t)$ is the total population of county i in time t. This sum, the cross-sectional measure of market potential, is normalized by average market potential (otherwise later decades could overpower effects in earlier ones through absolute growth in market potential). Relative market potential gives a measure of each city's relative demand or competitive position in each decade.[21]

One might expect that better climate (warmer and drier), coastal location, and better access to resources all lead to increased relative growth. Greater amenities increase a location's benefits relative to the costs of urban congestion. In a growth context where technological change and knowledge accumulation enhance absolute city sizes (Black

and mean first passage times for a five cell size distribution with the same initial 1900 shares as in Table 4, the mean first passage time to move up from cell 1 to 5 (1146 years) is still quicker than to move down from 5 to 1 (1740 years) for these cities, despite a discrete distribution with higher shares in lower cells (see footnote 14).

21 Other forms of $MP_j(t)$, such as deflating by $\sqrt{d_{ij}}$ or not normalizing by $\overline{mp}(t)$, provide very similar results to those reported.

and Henderson, 1999b), cities with better amenities may grow more quickly, because the better amenities accentuate the gains from technological advances. However, the expectation may simply arise from the contextual situation in the US in the twentieth century. The US moved away from raw material based production in the Midwest to footloose service and high-tech manufacturing production. Location decisions became more heavily influenced by consumer amenities such as climate and coastal location (in the West and parts of the South). Hence the growth enhancement of these geographic variables in this time period.

Expectations about the impact of market potential are less clear. Trade theory (Davis and Weinstein, 1997; Hanson, 1997) suggests greater market potential should foster growth, since more and nearer neighbors offer better markets for a city's products and enhanced chances of growth. Location theory (Fujita et al., 1999) and hierarchy models (Beckman, 1968; Dobkins and Ioannides, 2000) suggest spatial competition forces place the biggest cities far from each other, such that the bigger the city the smaller its nearest neighbors. This suggests a market potential variable may have negative effects on growth. It is possible that market potential effects differ by city size, being important for small cities but adverse for big cities. However, we do not find that this is the case.

3.1. The determinants of relative growth

As noted in the previous section, to parameterize the effect of heterogeneity on the transition process utilized in the paper involves estimation of up to 20 different sets of coefficients. In this section, we use a simpler, continuous growth specification based on equation (2) which involves only one set of coefficients and suffices to exposit the forces at work:[22]

$$\ln(N_{it+1}) - \ln(N_{it}) = \alpha + \delta_t + \beta X_{it} + \gamma \ln(N_{it}) + u_i + \epsilon_{it}. \tag{8}$$

where X_{it} is a vector containing both time varying and time invariant variables.

We estimate equation (8) with and without the lagged dependent variable. Although that variable is meant to capture mean reversion and we have not modelled city growth processes here (cf. Black and Henderson, 1999b), it does imply a tendency of convergence towards a 'long run' size for each city. However that size is conditional on measured city characteristics, including time-varying market potential variables, time (national technological shocks), and unmeasured city characteristics reflected in the ϵ_{it} terms. In no way does this imply that all cities are converging towards a single long run steady-state. Even after conditioning on measured characteristics, ongoing shocks will maintain dispersion in the conditional distribution of city sizes.

We begin by pooling all decades in an OLS formulation assuming there are no fixed effects and that the ϵ_{it} are identically and independently distributed. We estimate with and without lagged values of city size. We then incorporate fixed effect considerations to account for the u_i and estimate the model by GMM to deal with other sources of endogeneity discussed below.

Basic OLS results are presented in Table 7(a). In column (i) explanatory variables are geographic variables—heating degree days and annual precipitation in logs and

22 It might be more consistent with the transition process utilized in this paper to have $N_{it+1}/\overline{N}_{t+1} - N_{it}/\overline{N}_t$ as the dependent variable. Results on this are reported in Black and Henderson (1998) and are qualitatively the same.

Table 7. City growth

	(a) OLS Results ($N = 2068$) (SE in parentheses)			
	(i)	(ii)	(iii)	(iv)
ln(heat° days)	−0.095**	−0.102**	−0.105**	−0.113**
	(0.015)	(0.015)	(0.015)	(0.015)
ln(precipitation)	−0.075**	−0.074**	−0.087**	−0.089**
	(0.016)	(0.015)	(0.017)	(0.017)
Coastal dummy	0.034**	0.049**	0.031**	0.046**
	(0.010)	(0.011)	(0.010)	(0.010)
Market potential			0.127**	0.141**
			(0.030)	(0.030)
Market potential sq.			−0.027*	−0.028**
			(0.0065)	(0.0065)
$\ln(N_{it})$		−0.023**		−0.025**
		(0.0033)		(0.0034)
Regional and time dummies	Yes	Yes	Yes	Yes†
adj R^2	0.373	0.385	0.378	0.392

	(b) Fixed effect results and GMM results fixed effects	GMM
Market potential	0.651**	0.805**
	(0.147)	(0.055)
Market potential sq.	−0.117**	−0.085**
	(0.024)	(0.0061)
$\ln(N_{it})$	−0.194**	−0.192**
	(0.018)	(0.0091)

Notes:
* Significant at 10% level.
** Significant at 5% level.
† Relative to the West, in column (iv), regional dummies on city growth rates for the NE, Midwest, and South, are −1.9, −1.3, and −0.095 respectively. With time-regional dummies, the time pattern is that regional differences before 1950 are very modest, with only the NE growing consistently, but modestly slower. From 1950 to 1970 or 1980, all regions grow much more slowly than the West, although the differential for the South is smaller. The differentials for the NE and Midwest are distinctly smaller for 1990.

a dummy variable for coastal. Coastal is a city on the seacoast or Great Lakes (where generally separate sea and Great Lakes impacts were similar). Cities in warmer (less heating degree days), drier (less precipitation) climates on the coast indeed grow faster. The utilization of regional dummy variables in the formulation potentially controls for other geographic differences and has little effect on the coefficients of geographic variables. Similarly, controlling for base period size has little effect. A 1% decrease in heat degree days or precipitation increases a city's growth rate per decade by about 0.1 and 0.08 percentage points respectively. Surprisingly regional dummies only have a modest impact on the first nature variables.[23]

The effects of market potential are shown in columns (iii) and (iv). Market potential's impact on growth has a quadratic shape, starting with a high marginal effect for cities

23 In column (iv), without regional dummies the coefficients on variables in the column going from top to bottom are −0.125, −0.125, 0.067, 0.017, −0.004, and −0.047. The main effect is on market potential, where the low market potential West grows quickly, diminishing market potential effects.

with low market potential. However, the benefits of greater market potential diminish as market potential rises. The mean, standard deviation and maximum value of relative market potential are respectively 1.0, 0.48, and 5.0. Total market potential effects are maximized at about 2 to 2.5 SD above the mean market potential, and decline thereafter. They are exhausted around the maximum market potential in the sample. Around mean market potential, a 1-SD increase in market potential raises the decade growth rate by three percentage points, a big impact. The general positive marginal effect conforms to trade theory expectations. The need to control for regional dummies and the diminution of marginal effects as market potential rises suggests the role of location theory. If a city is in a very high market potential area, it suffers from competition as location theory suggests. In terms of regional controls, the rapid growth in the West could be viewed as the benefits of strategic location—Los Angeles benefits by being far from New York. We also checked if larger cities benefit more or less from market potential effects by interacting market potential with city size, but found no significant impact.

The results in general help explain why big cities stay relatively big as found in Section 2. The additions to cities in the top, say, 10th percentile over time are fast growing cities in the West and South, with good climates and coastal locations. But traditional large cities in the Midwest and Northeast hold their places in the top 10th percentile by virtue of having high market potential.

There are problems with the OLS specification of (8) used in Table 7(a). First, while geographic characteristics for city i are exogenous to city i, a fixed effect, u_i, affects city growth which in turn affects, say, $MP_j(t+1)$, market potential for other cities. Contemporaneous errors ϵ_{it} also affect growth rates which in turn affect $MP_j(t+1)$. The introduction of lagged city size creates a major endogeneity problem. In addition, the potential for spatial correlation between cities exists. To deal with these issues, we start with fixed effects estimation and then turn to GMM estimation.

To estimate the model by GMM, we first difference equation (8) eliminating fixed effects. For each decade, there is then a separate equation (i.e., $[(\ln N_{it+1} - \ln N_{it}) - \ln N_{it-1})] \equiv \ln N_{it+1} - 2\ln(N_{it}) + \ln N_{it+2}$ is the dependent variable for year $t+1$). We use predetermined values of market potential and, lagged dependent variables, as well as the exogenous regional dummies and weather and coastal variables as instruments.[24] After first differencing and instrumenting there are seven equation years. Having separate decade equations, as well as instrumenting with predetermined variables, breaks the cross observation correlation, both across time and space,[25] and deals with endogeneity of own lagged size. In estimation of the GMM model, we impose equal coefficients across decades and use an unbalanced panel, incorporating cities into the sample, once they enter as a metro area. Given differencing and instrumenting, to be in the sample an MSA must have entered by 1960. In the estimation procedure we use the 1998 version of the DPD program of Arellano and Bond (1991), which corrects for heterogeneity.

Fixed effect and GMM results are reported in Table 7(b). Fixed effect results display greater mean reversion and more pronounced quadratic market potential effects than in Table 7(a). Marginal market potential effects in column (i) are much higher at the mean (of 1) than under OLS, but they do still become negative at high values of MP (about 3

24 For year $t+1$, time varying instruments start are for $t-s$, where $s \geq 2$.
25 See Anselin (1988).

SD above mean MP). In column (ii), we examine results under GMM.[26] Marginal market potential effects are greater and persist throughout the relevant range of market potential. In general, the trade theory version of market potential for city growth is vindicated. But there is evidence that market potential effects diminish as market potential gets high, indicating the role of competition from location models.

4. Urban specialization and transformation

Part of the reason why there is a stable size distribution of cities, where cities have very different sizes, lies beyond natural geography. In a context with non-differentiated geography, urban theory predicts absolute or relative urban specialization, whether based on external scale economies as in Henderson (1974) or on the transport costs of serving an exogenous agricultural population as in Fujita et al. (1999). In this section we examine specialization and its relation to city sizes. We also discuss the link between relative growth rates of cities and changes in the industrial composition of cities. These connections provide an explanation for the increase in urban concentration in recent decades.

4.1. Typing of cities

There is no consistent data on detailed industrial composition of cities prior to 1950. However, from County Business Patterns, we have county data for recent years. The common method of typing cities is to use cluster analysis on employment data (see Bergsman et al., 1972 and Henderson, 1988). These prior cluster analyses based on data from around 1970 show very clear typing of cities by manufacturing activity. Since 1970, manufacturing has declined from over 28% of US non-government employment to under 19%, with many cities losing their manufacturing bases. So many cities are in transition and classifying the industrial base of transforming cities at a point in time is problematical. Nevertheless, clear patterns emerge.

We classify cities based on their degree of specialization by two-digit SIC activities for private employment in all industries in 1992. We do this for 317 metro areas. The degree of specialization in a particular industry *i* in city *j* is measured by the share of industry *j* in total local private employment, s_{ij}. There are 80 two-digit industries (after dropping the share of the last (99-), non-independent industry). Cities are grouped on the basis of similarity of production patterns, indicated by employment shares of different industries. In clustering cities we use Ward's criterion, grouping cities to minimize the error sum of squares within clusters summed across all n clusters. The criterion function is

$$\sum_{c=1}^{n} \delta_c \left[\sum_{j=1}^{m_c} \sum_{i=1}^{80} \left(s_{ijc} - \frac{\sum_{j=1}^{m_c} s_{ijc}}{m_c} \right)^2 \right]$$

where m_c is the number of cities in cluster c and δ_c takes a value 1 if the city is in cluster c and zero otherwise. The clustering algorithm is hierarchical (step-wise) and the number

26 The tests on serial correlation indicate the assumption that the ϵ_{it} are identically and independently distributed overtime is fine. However, the GMM models under all formulations and choice of instruments fail the Sargan test for over-identifying restrictions. While the test tends to falsely reject, this is an unsettling diagnostic.

of clusters n is set by the researcher (there generally being no globally optimal number of clusters, nor general algorithms that are non-hierarchical).

Full cluster results are reported in Appendix 1 and a sample of results are reported in Table 8. In forming clusters, we originally specified 50 clusters but broke apart three of the last five clusters formed (in the hierarchy) to form eight sub-clusters among those three, to better distinguish the largest metro areas. We conducted an F-test on whether the 55 individual clusters are distinct from each other, and strongly reject the null that clusters are similar.[27] In the Appendix the 55 clusters are organized into eight groups of broad product-city categories for expositional purposes. These are clothing and food, wood products, electronics, heavy manufacturing, oil and chemicals, market centers, health services, and other services. In Table 8 we report on the cluster results for two of these groups, a typical one, electronics, and an unusual one, market centers. These cover clusters 10–14 and 40–45 of the 55 listed in the Appendix.

To explain the contents of Table 8, consider cluster 10, computers and interlinked industries. Column (3) tells us that there are three metro areas in the cluster. Column (4) indicates the average population size of metro areas in the cluster, about 700,000, and the average percent of the population completing college (28%). Compared to other electronics clusters, the innovative computer industry is in larger size metro areas with highly skilled labor forces. In column (5), the share (s_{ij}) of the dominant industry [which SIC number is listed in column (2)] is 22% of local private employment, compared to a 6% share in national employment ($\bar{s}_i$) of computers and inter-linked industries. Finally, the last column names two MSAs belonging to the cluster.

In general Table 8 and Appendix A tell us that there is incredible diversity across cities in terms of production patterns. As Appendix A reveals most cities and clusters are like the electronics cities. Some cities are intensely specialized with 15–30% of their labor force in just one industry (or set of inter-linked industries). Note that normally in the US at least 65% of local labor forces are in 'non-traded' good activity, whereas these industries of specialization are traded goods. In most clusters even if the dominant industry in a city only accounts for 5% of local employment, that s_{ij} is multifold the national share of the industry, $\bar{s}_i$. There are two exceptions. First are market centers which are much larger than other cities with above average education. These can be strictly diverse centers or have very mixed bases, emphasizing a particular range of service and non-service activities. In many cases they have no dominant industry and hence column (5) has 'n.a.' listed. Second, clusters are also based on what cities don't produce—e.g., an absence of heavy manufacturing. Despite the national move away from manufacturing, some cities are still strongly specialized in manufacturing activity. But there are a number of emerging or strengthened services centers for health, insurance, business, transport, education, hotels and recreation and eating and hotels for military and government and in some cases a dominant industry is just emerging.

From Table 8 and the Appendix average city sizes tend to vary by cluster. Within most clusters, cities have very similar sizes with one or two outliers. Across clusters, if we use the cluster with 20 market centers in diverse manufacturing in Table 8 having an average

27 The F-test examines whether the model sum of squares increases significantly (relative to the overall residual sum of squares) from a model where the s_{ij} differ only by industry compared to a model where they differ by industry clusters. $N = 317 * 80 = 25{,}360$. For 4,320 and 20,960 degrees of freedom (noting there are $55 * 80 = 4{,}400$ industry clusters) $F = 11.3$, with a critical value near 1.

Table 8. Examples of clusters

Cluster number	Cluster name (SIC)	No. of MSAs	Average size (1000s); % college +	Share of dominant industry*		Comments
(B) Electronics:						
10	Computers (87 + 35 + 36) (engineering, R&D, electronics)	3	712; 27.7	0.220	(0.06)	(San José, CA, Huntsville, AL)
11	Electrical machinery (36)	2	114; 12.6	0.245	(0.016)	(Madison County, IN Kokomo, IN)
12	Electronics (36)	4	159; 29.7	0.101	(0.016)	(Binghamton, NY, Bloomington, IN)
13	Diverse machinery (electronics)	6	232; 26.4	n.a.	n.a.	(Boulder, CO, Cedar Rapids, IA)
14	Instruments (38)	5	291; 18.2	0.077	(0.008)	(Rochester, NY, Sherman, TX)
(G) Market centers:						
40	Diverse services (health, educ, eng and mgmt)	7	2,471; 25.8	n.a.	n.a.	North-East post industrial metro areas (Boston, MA Philadelphia, PA)
41	Financial services (60 + 62)	1	8,547; 24.6	0.086	(0.010)	(New York, NY)
42	Mixed base metro areas (high tech, wholesale, transportation and bus serv)	9	2,673; 26.0	n.a.	n.a.	Western (and SW) metro areas (Orange County, CA, Denver, CO)
43	New mixed base (high tech, eating places, eng and mgmt)	8	1,405; 27.5	n.a.	n.a.	Western (and SW) cities (Phoenix, AZ, San Diego, CA)
44	Business services (73), transport, eating	7	1,403; 24.0	0.093	(0.051)	Southern cities (Houston, TX, Tampa, FL)
45	Diverse manufacturing	20	1,285; 21.5	n.a.	n.a.	Mid-west MSAs, (Cleveland, OH, Chicago, IL)

* Normal share in parentheses.

size of 1.3 m as the base, 36 out of the 55 clusters in Appendix 1 with average sizes under 350,000 have significantly smaller average city sizes.[28] Correspondingly the two clusters with average sizes over 2.7 m have significantly larger city sizes. Education also varies by city type. Using the cluster with 20 market centers as the base again, it has a little higher than the national average percent of adults with 4+ years of college in 1990, or 21.5%. Seventeen other clusters have significantly lower percent college, especially traditional manufacturers, with only 11–12% college educated. Eight other clusters have significantly higher percentages, especially high-tech centers (electronics, instruments, and computers), as well as some of the bigger market centers, with up to 30% college educated.

In the last 20 years, with the decline of manufacturing and rise of services nationally, many cities have undergone transformations of their industrial bases. Black and Henderson (1998) demonstrate that there is a strong correlation between changes in city size and changes in industrial composition over the period 1980–1990, with both the relatively fastest and slowest growing cities experiencing the greatest change in their industrial compositions. This suggests the fastest and slowest growing cities are both changing sizes because they are changing type (i.e., industrial composition), moving respectively up and down the urban 'hierarchy'.

This link between city sizes and output mix relates to our earlier analysis of increased urban concentration. We note that the largest city types are those producing modern, inter-city traded services such as financial, engineering and management, business, education, and some transport. These are all sectors of the economy that have been experiencing the very highest growth rates over the last two decades, especially business services. Given that, the increase in relative urban concentration in the post-World War 2 period should not be surprising. The relative numbers of very large service oriented cities are increasing.

5. Conclusions

This paper has documented the following aspects of urban evolution in the US, based on patterns over this century: (1) Urban evolution, as predicted by an urban endogenous growth model, involves on-going increases in absolute metro area sizes due probably to technological change represented by local human capital accumulation. Evolution has also involved on-going increases in numbers of metro areas; (2) The system of cities supports a wide fairly stable relative size distribution of cities over time. Cities of different types, or industrial compositions, tend to have different absolute sizes, as well as educational attainments, with relative size changing as industrial composition changes; (3) The transition of cities through the size distribution is stationary over the century, with no tendency for mobility speeds to dampen. With entry of new cities, existing cities tend to move up the size distribution 'fairly quickly', but to move down extremely slowly. The slow moves down are a testament to the effects of established urban scale; (4) While the size distribution is largely stable, there is a some tendency in the USA towards increasing urban concentration, with a greater proportion of cities in large relative size categories.

28 This is based on a regression of log (1990 SMSA) population on cluster dummy variables, looking for significant percentage variables. Two other larger city size clusters also have sizes significantly smaller than the 1.3 m.

This probably reflects the change in US industrial composition towards modern traded services, which tend to be concentrated in the very largest types of cities; (5) Urban sites are heterogeneous and superior sites tend to grow more quickly than inferior sites. Superior sites are those with better geography—better climate and on the coast—and better market potential. As trade theory predicts better market potential—larger neighbors nearer by—enhances a city's demand and growth potential. However, positive marginal market potential effects do die out as market potential becomes very large, suggesting location theory forces of nearby competitors eroding the market start to dominate at some point.

Acknowledgements

Support of the National Science Foundation for this research is gratefully acknowledged. Some of this research was completed while Duncan Black was a visiting research fellow at the Center for Operations Research and Econometrics (CORE) in Louvain-la-Neuve, Belgium. Mark Beardsell, Gilles Duranton, Ed Glaeser, Henry Overman, Yannis Ioannides, and Justin Tobias provided helpful comments.

References

Anselin, L. (1988) *Spatial Econometrics: Methods and Models*. Boston: Kluwer Academic Publishers.

Arellano, M., Bond S. (1991) Some tests of specification for panel data: Monte Carlo evidence and an application to employment equations. *Review of Economic Studies*, 58(2): 537–559.

Arthur, B. (1990) 'Silicon Valley' locational clusters: when do increasing returns imply monopoly. *Mathematical Social Sciences*, 19(3): 235–251.

Beckmann, M. (1968) *Locational Theory*. New York: Random House.

Bergsman, J., Greenston, P., Healy, R. (1972) The agglomeration process in urban growth. *Urban Studies*, 9(3): 283–88.

Black, D., Henderson, V. (1998) Urban evolution in the USA. Working Paper No. 98-121, Brown University.

Black D., Henderson, V. (1999a) Spatial evolution of population and industry. *American Economic Review* (AEA Papers and Proceedings), 89(2): 321–327.

Black, D., Henderson, V. (1999b) A theory of urban growth. *Journal of Political Economy*, 107(2): 252–284.

Bogue, D. (1953) Population growth in standard metropolitan areas 1900–1950. Oxford, OH: Scripps Foundation in Research in Population Problems.

Clark, S.J., Stabler, J.C. (1991) Gibrat's Law and the growth of Canadian cities. *Urban Studies*, 28(4): 635–639.

Davis, S., Weinstein, D.E. (1997) Economic geography and regional production structure: an empirical investigation. Working Paper No. 6093, NBER.

Davis, S., Haltiwanger, J., Schuh, S. (1996) *Job Creation and Destruction*. Cambridge, MA: MIT Press.

Dickey, D.A., Fuller, W.A. (1981) Likelihood ratio statistics for autoregressive time series with a unit root. *Econometrica*, 49(4): 1057–1072.

Dobkins, L., Ioannides, Y. (2000) Dynamic evolution of the US city size distribution. In J.M. Huriot and J.F. Thisse (eds) *The Economics of Cities*. Cambridge: Cambridge University Press.

Duranton, G., Puga, D. (2000) Nursery cities: urban diversity, process innovation, and the life-cycle of products. Mimeo, University of Toronto.

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

Ehrlich, S., Gyourko, J. (1999) Changes in the scale and size distribution of metropolitan areas in the United States during the 20th century. Mimeo, The Wharton School, University of Pennsylvania.
Fujita, M., Ishii, T. (1994) Global location behavior and organizational dynamics of Japanese electronics firms and their impact on regional economics. Paper prepared for Prince Bertil Symposium on the Dynamic Firm, June 1994.
Fujita, M., Kurgman, P., Mori, T. (1999) On the evolution of hierarchical urban systems. *European Economic Review*, 43(2):209–251.
Gabaix, X. (1999a) Zipf's Law and the growth of cities. *American Economic Review (AEA Papers and Proceedings)*, 89:129–132.
Gabaix, X. (1999b) Zipf's Law for cities: an explanation. *Quarterly Journal of Economics*, 114: 759–767.
Hamilton, B., Mills, E. (1994) *Urban Economics*. Glenview, IL: Scott-Forseman.
Hanson, G. (1997) Market access and economic geography. Mimeo, University of Texas.
Heckman, J. (1982) Branch plant location in the product cycle in computer manufacturing. Mimeo, University of North Carolina.
Henderson, J.V. (1974) The sizes and types of cities. *American Economics Review*, 64: 640–656.
Henderson, J.V. (1988) *Urban Development: Theory, Fact and Illusion*. Oxford: Oxford University Press.
Ioannides, Y., Overman, H. (2000) Zipf's Law for cities: an empirical examination. Working Paper No. 2000-06, Tufts University.
Ioannides, Y., Overman, H. (2001) Cross-sectional evolution of the US city size distribution. *Journal of Urban Economics*, 49: 545–566.
Karlin, S., Taylor, H. (1975) *A First Course in Stochastic Processes*. San Diego: Academic Press.
Kolko, J. (1999) Can I get some service here? Information technology, services industries, and the future of cities. Mimeo, Harvard University.
Krugman, P. (1993) *Geography and Trade*. Cambridge, MA: MIT Press.
Krugman, P. (1996) *The Self-Organizing Economy*. Cambridge, MA: Blackwell.
Levin, A., Lin, C.F. (1992) Unit root tests in panel data: asymptotic and finite-sample properties. Working Paper 92-14, Department of Economics, University of California, San Diego.
Rosen, K.T., Resnick, M. (1980) The size distribution of cities: an examination of the Pareto law and primacy. *Journal of Urban Economics*, 165–186.

Appendix 1

Table A1. City types*

Cluster number	Cluster name	No. of MSAs	Average size (1000s); % college +	Share of dominant industry*		Comments
(A) Health and hotel services						
1	Health services (80)	3	152 29.8	0.278	(0.115)	(Iowa City, IA, Rochester, MN)
2	Health services (80)	8	148 17.3	0.190	(0.115)	Former heavy industry (Alexandria, VA, Wichita Falls, TX)
3	Health services services (80) and mixed industry	20	311 21.4	0.136	(0.115)	Transforming industrial centers (Jackson, MS, Springfield, MO)
4	Health services (80) with electronics	5	160 28.2	0.172	(0.115)	State govt and former heavy machinery (Charlottesville, VA, Gainesville, FL)
5	Eating places, health services	16	458 20.7	n.a.	n.a.	Govt. military (Bremerton, WA, Fort Walton, FL)
6	Hotels	2	530 15.2	0.330	(0.016)	(Las Vegas, NV, Cape May, NJ)
7	Hotels (70) and recreation (79)	1	255 20.7	0.192 0.042	(0.016) (0.012)	(Reno, NV)
8	Eating places (58), some industry	4	159 21.9	0.113	(0.079)	Military (Fayetteville, NC, Lawton, OK)
9	Eating places (58) and hotels	13	293 21.9	0.121	(0.079)	State govt, military (Jacksonville, NC, Bryan, TX)
(B) Electronics						
10	Computers (87 + 35 + 36) (engineering, R&D, electronics)	3	712 27.7	0.220	(0.06)	(San José, CA, Huntsville, AL)
11	Electrical machinery (36)	2	114 12.6	0.245	(0.016)	(Madison County, IN Kokomo, IN)
12	Electronics (36)	4	159	0.101	(0.016)	(Binghamton, NY, Bloomington, IN)
13	Diverse machinery (electronics)	6	232 26.4	n.a.	n.a.	(Boulder, CO, Cedar Rapids, IA)
14	Instruments (38)	5	291 18.2 29.7	0.077	(0.008)	(Rochester, NY, Sherman, TX)

Continued

Table A1. (*continued*)

Cluster number	Cluster name	No. of MSAs	Average size (1000s); % college +	Share of dominant industry*		Comments
(C) Clothing and food processing						
15	Textiles (22)	3	121 12.5	0.188	(0.006)	(Danville, VA, Washington County, RI)
16	Textiles (22)	3	915 19.0	0.069	(0.006)	(Charlotte, NC, Greenville, SC)
17	Textiles and food processing (22 + 20)	4	236 19.3	0.105	(0.025)	(Athens, GA, Columbus, GA)
18	Apparel (23) and mixed industry	10	370 14.9	0.054	(0.007)	Declining industrial centers (Brownsville, TX, industry Anniston, AL)
19	Food processing (20)	6	165 15.5	0.106	(0.019)	(Merced, CA, Sioux City, IA)
20	Food processing, machinery, transport	2	156 12.4	n.a.	n.a.	(Fort Smith, AR, Joplin, MO)
21	Food processing (20), wholesale, agriculture, transport	11	293 17.6	0.046	(0.019)	(Visalia, CA, Fresno, CA)
22	Food processing (20), and diversified heavy machinery	9	248 14.5	0.050	(0.019)	Declining heavy industry (Gadsden, AL, Sheboygan, WI)
(D) Heavy manufacturing						
23	Machinery (35)	11	175 16.7	0.080	(0.020)	Declining machinery, steel centers (Peoria, IL, Waterloo, IA)
24	Transforming steel cities (33)	9	260 14.0	0.049	(0.010)	Former steel giants (Huntington, WV, Reading, PA)
25	Diverse heavy machinery	6	212 13.6	n.a.	n.a.	Transport, machinery, metals (Jackson, MI, Rockford, IL)
26	Primary metals (33)	3	290 12.4	0.165	(0.010)	(Gary, IN, Steubenville, OH)
27	Transport equipment (37)	4	363	0.149	(0.016)	(Wichita, KS, Flint, MI)
28	Transport equipment (37)	13	965 19.3	0.067	(0.016)	Declining transport (Kenosha, WI, Saginaw, MI)

			24.5			
29	Transport equipment (37) and lumber (24)	1	156	0.155	(0.016)	(Elkhart, IN)
			14.2	0.074	(0.006)	
30	Transport equipment (37)	1	115	0.284	(0.016)	(Jackson, MI)
			14.4			
(E) Oil and chemicals						
31	Non-metallic minerals (32)	1	138	0.141	(0.006)	(Vineland, NJ)
			10.8			
32	Oil and gas (13), mining and water transport	2	145	0.106	(0.003)	(Houma, LA, Midland, TX)
			18.0			
33	Oil and gas (13), transport	5	231	0.050	(0.012)	(Lafayette, LA, Anchorage, AK)
			18.3			
34	Chemicals (28), petro chemicals	9	246	0.073	(0.011)	(Johnson City, TN, Lake Charles, LA)
			15.5			
35	Chemicals (28)	2	171	0.179	(0.011)	(Brazooria, TX Richland, WA)
			18.1			
(F) Wood products						
36	Furniture (25)	1	222	0.178	(0.004)	(Hickory, NC)
			16.1			
37	Pulp and paper (26)	3	208	0.076	(0.008)	Diversifying (Green Bay, WI, Oshkosh, WI)
			16.1			
38	Pulp and paper (26), leather	4	110	0.057	(0.008)	Declining non-heavy industry (Pine Bluff, AR Bangor, ME)
			14.1			
39	Lumber (24) and non-durable wholesale	6	185	0.054	(0.006)	NW Declining lumber cities (Salem, OR, Eugene, OR)
			17.7			
(G) Market centers						
40	Diverse services (health, educ, eng and mgmt)	7	2,471	n.a.	n.a.	North-East post industrial metro areas (Boston, MA, Philadelphia, PA)
			25.8			
41	Financial services (60 + 62)	1	8,547	0.086	(0.010)	(New York, NY)
			24.6			
42	Mixed base metro areas (high tech, wholesale, transportation and business serv)	9	2,673	n.a.	n.a.	Western (and SW) metro areas (Orange County, CA, Denver, CO)
			26.0			
43	New mixed base (high tech, eating places, eng and mgmt)	8	1,405	n.a.	n.a.	Western (and SW) cities (Phoenix, AZ, San Diego, CA)
			27.5			

Continued

Table A1. (*continued*)

Cluster number	Cluster name	No. of MSAs	Average size (1000s); % college +	Share of dominant industry*		Comments
44	Business services (73), transport, eating	7	1,403 24.0	0.093	(0.051)	Southern cities (Houston, TX, Tampa, FL)
45	Diverse manufacturing	20	1,285 21.5	n.a.	n.a.	Mid-west MSAs, (Cleveland, OH, Chicago, IL)
(H) Service centers						
46	Insurance (63)	2	627 27.7	0.142	(0.017)	(Bloomington, IL, Hartford, CT)
47	Insurance (63) and business services	5	369 23.3	0.063	(0.017)	(Des Moines, IA, Lincoln, NE)
48	Business services (73)	2	154 13.7	0.216	(0.051)	(Nanatee County, FL, Kankakee, IL)
49	Air transport (45) and oil	1	57 17.3	0.084	(0.005)	(Enid, OK)
50	Trucking and transport services (42 + 47)	1	133 11.1	0.163	(0.020)	(Laredo, TX)
51	Private education services (82), manufacturing	2	295 27.9	0.113	(0.018)	(Provo, UT, Trenton, NJ)
52	Diverse services (business, insurance, education, health)	8	556 20.1	n.a.	n.a.	NE cities transforming from heavy industry (Springfield, MA, Syracuse, NY)
53	New Jersey transport, wholesale, chemicals	4	1,169 26.2	n.a.	n.a.	(Bergen, Somerset)
54	Diverse services and light manufacturing	15	796 19.2	n.a.	n.a.	Govt and military (Miami, FL, Sacramento, CA)
55	Manufacturing NEC (3999), chemicals, diverse manufacturing	4	373 24.4	0.075	(0.010)	(Racine, WI, Wilmington, DE)

Note: * A number in parentheses in column (2) represents an SIC code for which city-industry share, s_{ij}, is reported in column (5). In column (5), the number in parenthesis is the national share s_i of the industry.