JOURNAL OF REGIONAL SCIENCE, VOL. 47, NO. 2, 2007, pp. 273–287

CHANGES IN THE NATURE OF URBAN SPATIAL STRUCTURE IN THE UNITED STATES, 1890–2000*

Sukkoo Kim

Department of Economics, Washington University in St. Louis and NBER, St. Louis, MO 63130. E-mail: soks@artsci.wustl.edu

ABSTRACT. This paper documents the long-run trends in the average densities and density gradients of urban areas in the United States. The data show that between 1890 and 2000 the average densities of cities and metropolitan areas rose and fell but that the density gradients of urban areas generally declined monotonically over time. While it is beyond the scope of this paper to estimate the causes of these changes, this paper argues that a complete understanding of the changes in the nature of US urban spatial structures is likely to go beyond the standard explanations based on the monocentric city model such as decreases in transportation costs and increases in household incomes.

1. INTRODUCTION

One of the most distinguishing characteristics of a city is its density. Indeed, an urban area is defined as a densely populated place with a sizeable number of inhabitants. Yet, despite the fact that the defining element of an urban area is its density, few scholars have systematically examined the long-run changes in the average densities of US cities. Instead, the majority of the empirical work, motivated by the monocentric city model, examines the density gradient which measures the changes in the density of an urban area as one moves outward from a city's central business district.[1] For most urban economists, the decline

*Comments from Will Strange, Andrew Haughwout, three referees, Marcus Berliant, Rebecca Menes, Chuck Moul, Fred Smith, Brian Spicher, and the seminar participants at the Economic History Association meetings in Los Angeles, Regional Science Association meetings in Charleston, ASSA Meetings in Atlanta, SEA meetings in NEW Orleans, and workshops at University of British Columbia, Federal Reserve Bank of Kansas City, UCLA and Washington University are gratefully acknowledge. I thank Nate Baum-Snow for generously providing data for a previous version of this paper and Shilpi Kapur for valuable research assistance.

Received: April 2005; revised: January 2006 and May 2006; accepted: May 2006.

[1]Density is measured simply as persons per square mile. The density gradient is usually estimated using a negative exponential function: $D(x) = D_0\, e^{-\gamma x}$ where $D(x)$ is population density at distance x from the center, D_0 is the density at the center, and γ, the density gradient, is the proportional rate at which population density falls with distance from the center (see Mills and Hamilton (1989)).

Blackwell Publishing, Inc., 350 Main Street, Malden, MA 02148, USA and 9600 Garsington Road, Oxford, OX4 2DQ, UK.

in the density gradient rather than that of the average density has been used to define the process of suburbanization.[2]

While the long-run historical data on spatial structures of urban areas are relatively sparse, there is some evidence that urban densities rose and fell between the late nineteenth and the twentieth centuries whereas density gradients fell monotonically over time. Mills (1972) finds that the densities of four cities (Baltimore, Milwaukee, Philadelphia, and Rochester) rose and fell while their density gradients declined over time.[3] McDonald (1997) finds similar patterns for Chicago. Edmonton (1975) finds that the density gradients of 41 metropolitan districts declined between 1900 and 1950, but provide no information on their densities.

This paper documents the long-run trends in the densities and density gradients of US urban areas between the late nineteenth and the twentieth centuries. The data for US cities and metropolitan areas for a reasonably large sample of observations between 1890 and 2000 confirm that urban densities rose and fell over this period whereas density gradients fell monotonically over time. For cities, densities rose modestly between 1890 and 1950 but then fell sharply between 1950 and 2000. For metropolitan areas, densities peak slightly later in time, but the overall trend mirrors that of cities. Density gradients, calculated using the Mills' two-point method for a panel of 87 consistent cities, generally declined monotonically over time.

The standard explanations for why density gradients fell over time, such as a fall in transportation costs or a rise in household incomes, are based on the monocentric city model. However, these factors alone are unlikely to explain why average densities of urban areas rose and fell over time. It is quite likely that spatial structures of cities are also determined by location decisions of firms as well as households. While it is beyond the scope of this paper to estimate the causes of the rise and decline of average densities and the secular decline of density gradients of urban areas, this paper suggests that a complete understanding of these trends is likely to involve estimation strategies which take into account the joint location decisions of firms and households.

2. THE DENSITY OF URBAN AREAS

This section presents data on the densities of cities and metropolitan areas, and density gradients of urban areas between the late nineteenth and the twentieth centuries. The density gradient is calculated using the Mills'(1972)

[2]The three widely used textbooks in urban economics all define suburbanization as a decline in the density gradient: McDonald (1997), Mills and Hamilton (1989, p.374) and O'Sullivan (1996). Also see Mieszkowski and Mills (1993) and Anas, Arnott and Small (1998).

[3]The average density gradients of these four cities declined from 1.22 to 0.31 between 1880 and 1930. The decline was more or less monotonic except for Philadelphia whose density gradient remained flat between 1880 and 1910, rose slightly between 1910 and 1930, and then fell over the second half of the twentieth century (Mills (1972)).

two-point method which utilizes information on central city population, total metropolitan population and central city land area. Unlike the central city, whose political boundary changes only with annexation, consolidation or secession, the metropolitan area is defined by government officials. The metropolitan area is generally composed of a group of counties that contain a central city with a population of 50,000.[4] While the criteria for inclusion are rather complicated and have changed over time, a county is generally included if it has sufficient population density or sufficient commuters to the central city.

Density of Cities, 1890–2000

For each decade between 1890 and 2000, the densities of cities are calculated for cities whose populations are greater than 25,000 inhabitants. Because the number of cities in the sample changes over time, densities for a consistent sample of 119 cities are also examined. The data are from the *Social Statistics of Cities*, 1890, *Financial Statistics of Cities*, 1901–1940, the *County and City Data Book*, 1949–1994, and www.census.gov/prod/www/ccdb.html.

The densities of cities on average declined slightly between 1890 and 1930, but then fell sharply over the second half of the twentieth century. Between 1890 and 1930, the densities of cities fell from 7,748 to 7,366 persons per square mile; however, between 1930 and 1990, the figure fell almost by half to 3,789 and then rose to 3,917 in 2000. However, as shown below, the decline in the densities of cities is significantly influenced by the changes in the composition of cities over time. Data in Table 1 show that the number of cities in the sample rose almost tenfold from 122 to 1079 between 1890 and 2000. Since newer cities are generally less dense than older cities, the inclusion of newer cities tend to bias the density calculations downwards.

Since density is defined as population divided over land area, its change in an accounting sense can be attributed to changes in the numerator or the denominator. Between 1890 and 1930, urban population density remained relatively flat as the land area and population of cities in the sample remained

[4]The Census Bureau's first attempt to define the metropolitan area was in 1910. For cities with populations of at least 200,000, the "metropolitan district" was defined as the political city boundary plus any contiguous minor civil divisions and incorporated places having population density of at least 150 persons per square mile. For cities with populations between 100,000–200,000, it was defined as the municipal city plus all cities, towns, villages or other divisions located within 10 miles of the central city (See Thompson (1948)). Because the "metropolitan district" did not coincide with county boundaries, the amount of useful information that could be reported for the "metropolitan district" was extremely limited. In 1949, the Bureau of the Budget (Office of Management and Budget) redefined the metropolitan area based on units of counties, except for in New England where it was based on cities and towns. For the years between 1940 and 1980, the metropolitan area data were categorized at the smallest metropolitan unit designated as the standard metropolitan area (SMA) or standard statistical metropolitan area (SMSA). In 1983, the Census Bureau introduced three different metropolitan areas: metropolitan area (MSA), primary metropolitan area (PMSA), and consolidated metropolitan area (CMSA). Some CMSA's contain several cities and metropolitan areas.

TABLE 1: Population Densities of Cities, 1890–2000

	Number of Cities	Average Population	Average Land Area (sq. miles)	Average Population Density
1890	122	113,835	16.7	7,648
1900	160	123,243	20.2	7,377
1910*	184	148,442	23.1	7,176
1920*	252	145,966	20.9	7,597
1930*	310	152,890	21.9	7,366
1940	412	128,051	19.2	6,742
1950	481	128,811	19.5	6,536
1960	673	112,400	22.8	5,340
1970	835	104,785	28.5	4,673
1980	944	97,756	32.8	3,998
1990	1,068	98,108	34.9	3,783
2000	1,079	108,991	42.4	3,917

Note: The data, except for years 1910–1930, are for cities with population over 25,000. In 1890, two cities were omitted due to lack of data on land area. Cities in Alaska and Hawaii are excluded.

*Data for 1910–1930 are for cities with population over 30,000.

Sources: Social Statistics of Cities, 1890; Census of Population, 1900; Financial Statistics of Cities, 1910, 1920, 1930; County and City Data Book, 1949, 1952, 1962, 1972, 1982, 1988, 1994.

relatively constant. However, the decline in the densities of cities between 1920 and 1960 was primarily due to a fall in the central city population whereas the sharp decline between 1960 and 2000 can be attributed to an increase in city land area as well as a continued decline in the central city population.

When the densities of cities are examined by census regions, the data indicate that the long-run trends in the densities of cities in each region were similar to that of the overall sample. However, as indicated in Figure 1, there were some important regional variations. First, cities in the Middle Atlantic region were significantly and consistently denser than those in other regions. Second, cities in the two southern regions, South Atlantic and West South Central, were the densest in 1890 and 1900, but their densities fell precipitously over the twentieth century. Finally, the cities in the two Western regions, despite their reputation for low densities, were just as dense as cities in most other regions.

Since the number of cities in the sample changes over time, the changes in the overall average in the density of cities can be driven by secular changes or by changes in the composition of cities. Table 2 presents data on the average density of 119 consistent sample of cities, which was obtained by taking the intersection of cities over the entire period. Cities are persistent creations. Of the 122 cities whose populations were above 25,000 in 1890, the populations of only 3 cities fell below the cutoff value and fell out of the sample over the next hundred years.

The data for the consistent sample, unlike that of the full sample, show that the densities of cities rose and fell between the nineteenth and the twentieth

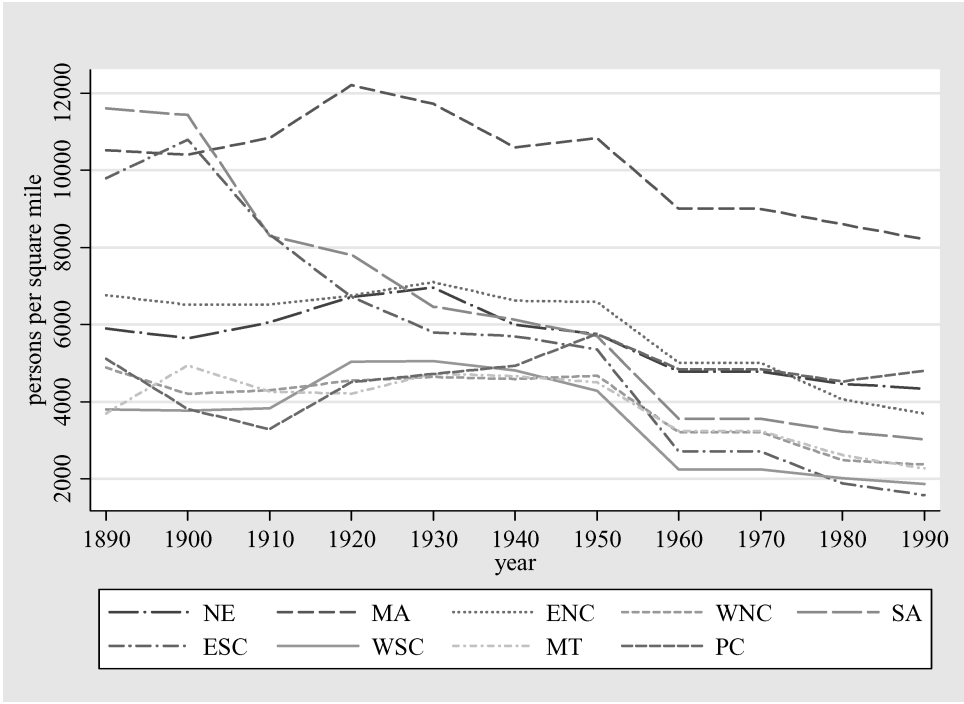

FIGURE 1: Density of Cities by Region, 1890–1900.

centuries. In 1890, the average of the population densities of the consistent sample was 7,203 persons per square mile; the figure rose to 8,876 in 1950, but then fell sharply to 5,659 by 2000. As noted above, the main cause of the differences in the trends in the average densities of the full and the consistent sample is due to the fact that, at any given point in time, new cities that came into the sample were less dense than older existing cities. Thus, rather than rising, the average density of the full sample of cities between 1890 and 1930 fell slightly as the addition of new cities lowered the aggregate average density. In addition, the continual addition of newer and less dense cities in the full sample over the second half of the twentieth century contributed to a sharper reduction in average densities for the full sample as compared to that of the consistent sample.

Urban densities rose between 1890 and 1950 despite the fact that cities during this period annexed significant amounts of new land from surrounding areas. Data on land area presented in Table 2 show that, between 1890 and 1910, cities increased their land area by over 20 percent per decade.[5] However,

[5]According to Jackson (1985), annexation or consolidation was prevalent in the late nineteenth century. Most annexation occurred in a discontinuous fashion: Philadelphia's annexation of

TABLE 2: Population Densities of Cities, 1890–2000
(Consistent Sample of 119 Cities)

	Average Population	Average Land Area	Average Population (sq. miles)	Average Growth Rates		
				Pop.	Land	Density
1890	117,124	19.1	7,203	0.25	0.20	0.05
1900	154,266	23.9	7,762	0.30	0.28	0.02
1910	206,711	29.5	7,626	0.22	0.10	0.13
1920	258,182	32.5	8,697	0.15	0.12	0.02
1930	311,661	36.8	8,751	0.03	0.01	0.01
1940	324,865	37.2	8,751	0.08	0.07	0.01
1950	359,297	40.6	8,876	0.04	0.19	−0.15
1960	371,478	50.8	7,660	−0.01	0.15	−0.16
1970	375,600	64.2	6,744	−0.10	0.06	−0.16
1980	344,608	69.9	5,890	0.01	0.05	−0.04
1990	350,320	72.8	5,647	0.02	0.05	−0.03
2000	367,680	78.0	5,659			

Sources: Social Statistics of Cities, 1890; Census of Population, 1900; Financial Statistics of Cities, 1910, 1920, 1930; County and City Data Book, 1949, 1952, 1962, 1972, 1982, 1988, 1994.

city densities rose as population growth outpaced land acquisition during this period. Over these two decades, the population growth of cities averaged over 25 percent per decade. While city density stabilized during the interwar years, it fell markedly over the second half of the twentieth century. Between 1950 and 1970, density declined because land area annexation increased whereas population growth remained relatively flat. On the other hand, the decline in city density between 1970 and 2000 was caused by a significant decline in population growth.

Older cities were more densely organized than younger cities. Figure 2 presents data categorized by dates of incorporation, 1653–1800, 1801–1830, 1831–1840, 1841–1850, 1851–1860, and 1861–1889, for the consistent sample

its county occurred in 1854, Chicago in 1889, and New York in 1898. However, some cities such as Minneapolis, Cleveland, Cincinnati and Pittsburgh expanded their boundaries in a series of smaller additions.

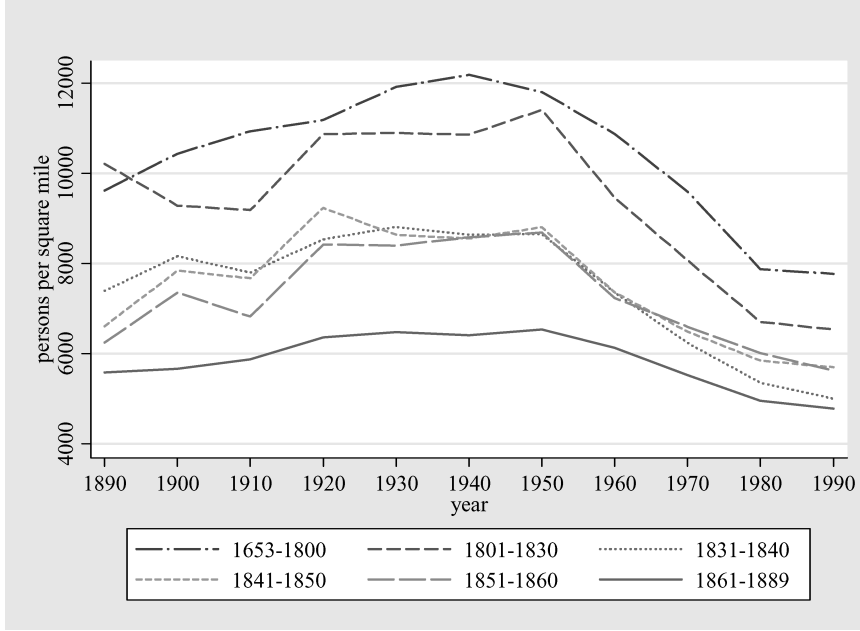

FIGURE 2: Density of Cities by Dates of Incorporation, 1890–1990.

TABLE 3: Population Densities of Metropolitan Areas, 1940–2000

	Number of Metro Areas	Average Population	Average Land Area (sq. miles)	Average Population Density
1940	156	439,143	1,138	386.9
1950	170	498,325	1,226	513.1
1960	212	632,774	1,463	589.4
1970	243	573,737	1,597	539.7
1980	305	558,993	1,888	383.1
1990	281	686,860	2,067	288.0

Sources: County and City Data Book, 1949, 1952, 1962, 1972. State and Metropolitan Area Data Book, 1982, 1991.

of cities. Like the trends observed in Table 3, cities in each cohort by dates of incorporation exhibited a rise and decline in their population densities over time. However, at any given point in time, the densities of older cities were consistently higher than those of younger cities.

Density of Metropolitan Areas, 1940–1990

For each decade between 1940 and 1990, the densities of metropolitan areas are estimated using the concurrent definitions for all metropolitan areas

TABLE 4: Population Densities of Metropolitan Areas, 1940–1980
(Consistent Sample of 149 Metropolitan Areas)

	Average Population	Average Land Area (sq. miles)	Average Population Density	Average Growth Rates		
				Pop.	Land	Density
1940	448,495	1,161	389.6	0.23	−0.01	0.23
1950	545,626	1,216	492.2	0.25	0.10	0.15
1960	664,643	1,336	547.0	0.17	0.23	−0.06
1970	783,523	1,589	500.9	0.19	0.31	−0.12
1980	885,569	2,100	442.1			

Sources: County and City Data Book, 1949, 1952, 1962, 1972. State and Metropolitan Area Data Book, 1982, 1991.

in each decade as well as for a consistent sample of metropolitan areas over time. The data are from a variety of sources such as the *County and City Data Book*, 1949–1977, *State and Metropolitan Area Data Book*, 1980–1990, *Census of Population*, 1990, and *Census of Housing and Population*, 1990.

The densities of metropolitan areas for the full sample rose and fell on average between 1940 and 1990. Data in Table 3 indicate that the number of metropolitan areas almost doubled from 156 to 281 over the sample period. Between 1940 and 1960, the densities of metropolitan areas rose from 387 to 589 persons per square mile; however, by 1990, the figure fell to 288.[6] Unlike the patterns observed for cities, the trends for a consistent sample of 149 metropolitan areas for the period between 1940 and 1980 are similar to that of the full sample.[7] As shown in Table 4, densities on average rose from 389 to 547 between 1940 and 1960, but then fell to 442 in 1980. Metropolitan population densities grew between 1940 and 1960 as population growth outpaced the modest increase in land area considered metropolitan. However, in the last few decades of the twentieth century, as metropolitan land area grew much faster than population, the estimated metropolitan population densities declined precipitously.

Density Gradients of Urban Areas, 1890–2000

The density gradient is calculated using the Mills' two-point method for a consistent sample of 87 metropolitan areas between 1890 and 2000.[8] The

[6]The 1990 data for metropolitan areas reported in Table 3 use CMSA's.

[7]Due to major changes in the definition of metropolitan areas, it was difficult to extend the data to 1990.

[8]We use the following equation from McMillen (2004): $P_c/P = [1 - (1 + \beta d)e^{-\beta d}]$ where P_c is the population of central city, P is population of the entire metropolitan area, d is the radius of

data on central city population and land area come from the same sources used to study average urban density. Because the Census Bureau did not report metropolitan area population for the period prior to 1940, it is necessary to estimate metropolitan area population for urban areas between 1890 and 1930. For reasons of consistency, this paper uses the 1950 definition of metropolitan areas for estimating metropolitan area population for the entire sample period (1890–2000). The data on county level population are from the *Population of States and Counties of the United States: 1790–1990* and www.census.gov/population/cen2000/pht-c3/.

While the use of Mills' two-point method is extremely attractive given its minimal data requirements, the estimates presented below are somewhat fictitious and must be interpreted with caution. First, the assumption that all cities are circular is likely to introduce cross-sectional distortions as cities vary considerably in their shapes; however, this assumption is less likely to introduce bias over time. Second, in a few cases where the metropolitan area contained more than one central city, a fictitious composite city was created by adding the two central city populations and land areas. Third, the use of the 1950 metropolitan area definition is likely to bias the estimates downwards for the period prior to 1950 and upwards for the period after 1950.[9]

Despite the heroic assumptions used to estimate density gradients, the estimates accord fairly well with the few long-run estimates reported in the existing literature. The density gradients of the 87 urban areas declined monotonically over the long-run, but did so with increasing pace over the second half of the twentieth century. On average, the gradients declined modestly from 0.93 to 0.82 between 1890 and 1920 and then fell to 0.72 by 1940. However, between 1940 and 2000, the gradients flattened sharply to 0.32. These long-run trends are similar to those found by Mills (1972), Edmonston (1975), Macauley (1985) and others surveyed by McDonald (1989).

Spatial Structures of Urban Areas, 1890–2000

The long-run historical data on urban densities and density gradients present a more complete picture of the changes in the urban spatial structures over time. Between 1890 and 1950, urban densities rose but density gradients declined. As Figure 3 illustrates, the intercept of the density gradient rose

---

the central city, and β is the estimated density gradient. The radius is estimated using the identity $A = (\varphi/2)d^2$ where A is the land area of the central city and $\varphi = 2\pi$ if the city is circular or $\varphi = \pi$ if the city is a half-circle bordered by lakes, oceans and other geographic barriers to settlement. Also see Mills (1972), Edmonston (1975), and McDonald (1989),.

[9]The estimates of urban density gradients for urban areas prior to 1950 are likely to be biased downwards since some counties in the 1950 definition may not have passed the metropolitan criterion prior to 1950. Thus, the estimates of the suburban county populations are likely to be over-estimated. On the other hand, the estimates of density gradients after 1950 using the 1950 definition are likely to be biased upwards. Since the 1950 definition excludes new counties that are considered as metropolitan after 1950, the suburban population is likely to be under-estimated.

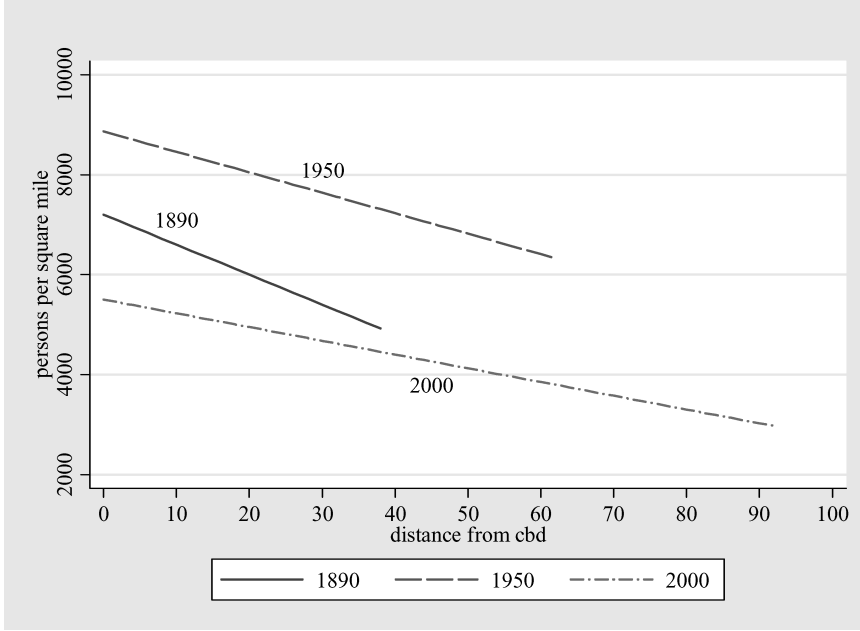

FIGURE 3: Urban Spatial Structure, 1890–2000.

even as its slope declined. Between 1950 and 2000, however, urban densities and density gradients both declined as the intercept of the density gradient curve shifted downwards as its slope declined further. Clearly, the standard use of the density gradient to define the process of suburbanization provides an incomplete picture of the changes in the spatial structures of urban areas over time.

The examination of the changes in the central city and suburban populations provides additional clues to the causes of urban spatial structures over time. Between 1890 and 1950, urban density gradients flattened but central city population grew absolutely and relatively to that of the suburban area. Data presented in Table 5 and Figure 4 show that the central city share of the total metropolitan area population rose from 0.57 to 0.64 over this period. How did the density gradient decline when the central city population rose absolutely and relatively to the suburban population? The answer must lie in the changes in the relative land area of central cities and suburbs. During this period, the relative density of cities as compared to that of the suburban area must have fallen as central cities acquired relatively more amounts of land than suburban areas.

Between 1950 and 2000, the average densities of cities fell even as the density gradients flattened significantly. During this period, as shown in Figure 4, the central city population declined absolutely and relatively to that of the suburban population. While these changes in population are consistent with the

TABLE 5: Population, Land Area, Densities and Density Gradients of Cities, 1890–2000 (Consistent Sample of 87 Cities and their Metropolitan Areas)

	Central City Population	Cental City/Metro Pop. Ratio	Central City Land Area	Average Density	Density Gradient
1890	120,275	0.57	20.4	7,039	0.93
1900	158,370	0.59	25.8	7,621	0.88
1910	212,994	0.62	31.9	7,484	0.84
1920	265,988	0.64	34.9	8,600	0.82
1930	320,711	0.64	39.2	8,650	0.76
1940	334,176	0.61	39.7	8,654	0.72
1950	371,084	0.57	43.3	8,794	0.64
1960	382,379	0.50	53.6	7,567	0.50
1970	386,713	0.46	67.3	6,661	0.42
1980	358,163	0.42	73.5	6,111	0.37
1990	360,646	0.40	76.5	5,572	0.34
2000	378,170	0.38	81.9	5,581	0.32

Note: Except for the cities in New England, metropolitan area is defined using the 1950 definition. The metroplitan population includes the populations of the central city and of any county included in metropolitan area definition. In New England, the definition is based on the 2000 metropolitan definition. Density gradient is calculated using the Mill's two-point method (see Mills (1972) and McMillan (2004)).

Sources: Social Statistics of Cities, 1890; Financial Statistics of Cities, various years; County and City Data Book, various years; Census of Population, various years.

decline in the density gradient over this period, it is likely that the expansion of suburban land area relative to that of the central city also contributed to the accelerated flattening of the density gradient over the second half of the twentieth century.

3. SOME POTENTIAL CAUSAL FACTORS OF URBAN SPATIAL STRUCTURES

What factors account for the long-run historical changes in the spatial structures of cities and metropolitan areas in the United States between the late nineteenth and the twentieth centuries? For economists, the standard explanations for the changes in urban spatial structures, especially the decline in the urban density gradients, are based on the monocentric city model.[10] Most scholars believe that the density gradients flattened over time because of

[10]Fujita (1989) derives the comparative statics properties of the closed-city model under absentee land ownership. In this model, two factors, agricultural land rent and population, affect average population density everywhere but not the slope of the density gradient. An increase in agricultural land rent or an increase in population results in an increase in population density everywhere. Two factors, transportation costs and income, affect both the average density and the slope of the density gradient. A reduction in transportation costs flattens the density gradient by

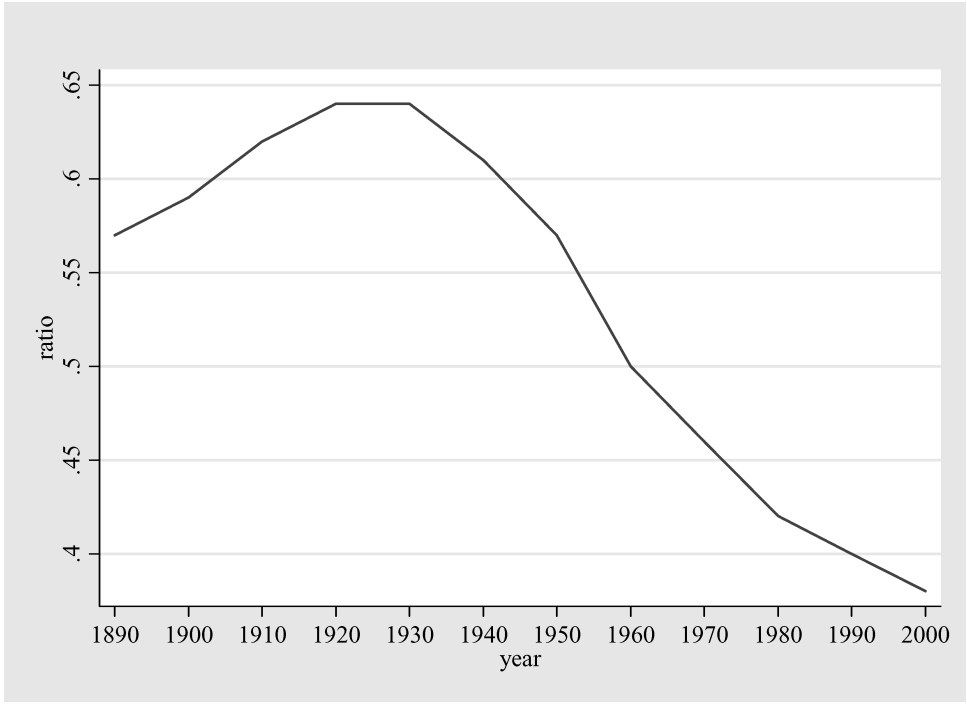

FIGURE 4: Share of Central City Population in Metropolitan Areas, 1890–2000.

declining transportation costs and rising incomes. In the late nineteenth century, the costs of commuting fell with the introduction of railroad street cars whereas in the twentieth century, they declined with the widespread adoption of automobiles (Warner (1978), Muth (1968), Jackson (1985), Glaeser and Kahn (2004)). In the second half of the twentieth century, Margo (1992) finds that rising incomes of American households also contributed to suburbanization.

Yet, falling commuting costs and rising household incomes cannot fully explain the long-run trends in the spatial structures of urban areas. Despite the usefulness of the monocentric city model, urban densities are unlikely to be fully characterized by household commuting and housing consumption decisions. In particular, these explanations cannot explain why the average densities of cities

lowering the population density at the city center and raising density in the suburbs; however, the average density of the city declines as the urban boundary moves outwards holding population constant. The predictions for income depend upon the supply of land in the suburbs. If there is a sufficient supply of land in the suburbs, then an increase in income will have the same effect as a fall in transportation costs; on the other hand, if there is an insufficient supply, then an increase in income will simply raise land rent everywhere. The monocentric model ignores the general equilibrium effects of rural development associated with changes in transportation costs. Kilkenny (1998) finds that there is a nonlinear relationship between transport costs and rural development.

and metropolitan areas rose between the late nineteenth and the mid-twentieth centuries. Moreover, contrary to the assumptions of the monocentric city model, firms were not always concentrated in the urban city center. In fact, throughout the twentieth century, the location decisions of manufacturing firms have been an important contributor of suburbanization (Kitagawa and Bogue, 1955). Because firms locate near each other to take advantage of various agglomeration economies, and because these economies attenuate over distance, urban densities are likely to depend upon the geographic nature of these economies (Rosenthal and Strange, 2004).

In sum, a complete understanding of the changes in the spatial structures of urban areas are likely to involve estimation strategies based on urban models where the location decisions of households and firms are both assumed to be endogenous. The simultaneous equations approach, developed by Steinnes and Fisher (1974), which estimate whether "people move to jobs" or "jobs move to people" implicitly assumes that the location decisions of households and firms are both endogenous, but these empirical estimates are difficult to interpret since the estimated equations are not based on a rigorous model. While developing an empirical estimation strategy based on the joint household and firm location models developed by Fujita and Ogawa (1982) and Lucas and Rossi-Hansberg (2002) is extremely challenging, pursuing such a strategy may yield important insights on understanding the changes in the nature of urban spatial structures.

4. CONCLUSION

This paper examines the long-run trends in US urban densities and gradients between the late nineteenth and the twentieth centuries. Despite the fact that a city is defined as a densely populated place with a sizeable number of inhabitants, most urban economists have focused their attention on examining the density gradient and define the process of suburbanization based on its decline. However, this paper argues that the spatial structure of an urban area is characterized by its density as well as its density gradient. The findings of this paper suggest that a more complete understanding of the long-run changes in urban spatial structures involves an understanding of the changes in urban densities as well as in its density gradients.

The paper finds that the average densities of cities and metropolitan areas rose and fell between 1890 and 2000. Between 1890 and 1950, the average densities of cities rose despite the fact that cities annexed large tracts of their hinterlands or consolidated with other cities. However, in the second half of the twentieth century, even as the pace of land annexation or consolidation slowed, the population density of cities fell as population growth declined sharply. While metropolitan areas exhibited similar patterns as cities, the substantial decline in the metropolitan density in the second half of the twentieth century was accompanied by a significant increase in land area that was newly defined and considered as metropolitan.

Unlike the long-run trends in the average densities of cities and metropolitan areas, the density gradients calculated for the consistent sample of 87 cities using the Mills two-point method indicate that density gradients declined monotonically over the long-run. Between 1890 and 1950, the paper finds that urban density gradients fell even as average densities of cities rose because the relative densities of cities fell compared to that of the metropolitan suburbs. While the municipal population rose during this period, this population growth was accompanied by a significant growth in municipal land area due to annexation and consolidation. However, in the second half of the twentieth century, density gradients flattened even more sharply as the population growth shifted significantly from cities to suburbs.

The most popular explanations for the long-run secular decline in urban densities, especially the density gradients, are falling commuting costs and rising incomes (Mills and Hamilton, 1989; Margo, 1992; Anas, Arnott and Small, 1998; O'Sullivan, 2007). Yet, despite the dominance of these two explanations based on the monocentric city model, firm agglomeration economies are also likely to play an important role in explaining urban spatial structures. Kim (2000, 2002) finds that urban densities are systematically associated with specializations in manufacturing, wholesale trade and retail trade.[11] While other factors such as white flight may also explain the trend toward suburbanization (Boustan (2006)), a more complete understanding of the long-run trends in urban densities awaits an empirical estimation strategy based on an urban spatial structural model where location decisions of households as well as firms are both taken to be endogenous.

REFERENCES

Anas, Alex, Richard Arnott, and Kenneth Small. 1998. "Urban Spatial Structure," *Journal of Economic Literature* 36, 1426–1464.

Bogue, Donald J. 1953. *Population Growth in Standard Metropolitan Areas 1900–1950*. Washington, DC: Housing and Home Finance Agency.

Boustan, Leah. 2006. "Inside the Black Box of 'White Flight': The Role of Suburban Political Autonomy and Public Goods," mimeo.

Edmonston, Barry. 1975. *Population Distribution in American Cities*. Lexington, D.C.: Heath and Company.

Fujita, Masahisa. 1989. *Urban Economic Theory*. Cambridge: Cambridge University Press.

[11]The empirical link between industrial activity and urban concentration has been well known since Weber (1899). Weber notes that specialization in extractive industries such as agriculture led to dispersion whereas specialization in distributive service industries and manufacturing contributed to urban concentration. Kim (2000) finds that, in 1900, city specialization in trade and transportation increased population density; in 1920, specialization in transportation and clerical services increased density; and in 1940, specialization in business services and government employment increased density. In all years, specialization in agriculture led to a decline in urban density. In general, the regression results indicate that specialization in manufacturing did not contribute to an increase in the density of cities. Specialization in wholesale trade was also significantly correlated with city size.

Fujita, Masahisa and H. Ogawa. 1982. "Multiple Equilibria and Structural Transistion of Non-monocentric Urban Configuration," *Regional Science and Urban Economics*, 12, 161–196.
Glaeser, Edward and Matthew Kahn. 2004. "Sprawl and Urban Growth," in Vernon Henderson and Jacques-Fran coise Thisse, (eds.), *Handbook of Regional and Urban Economics*, Volume 4, North-Holland/Elsevier.
Jackson, Kenneth T. 1985. *The Crabgrass Frontier*. New York: Oxford University Press.
Kilkenny, Maureen. 1998. "Transport Costs and Rural Development," *Journal of Regional Science*, 38, 293–312.
Kim, Sukkoo. 2000. "Urban Development in the United States, 1690–1990," *Southern Economic Journal* 66, 855–880.
Kim, Sukkoo. 2002. "The Reconstruction of the American Urban Landscape in the Twentieth Century," NBER Working Paper #8857.
Kitagawa, Evelyn M. and Donald J. Bogue. 1955. *Suburbanization of Manufacturing within Standard Metropolitan Areas*, Miami: Scripps Foundation.
Lucas, Robert E. and Esteban Rossi-Hansberg. 2002. "On the Internal Structure of Cities," *Econometrica*, 70(4), 1445–1476.
Macauley, Molly. 1985. "Estimating Recent Behavior of Urban Population and Employment Density Gradients," *Journal of Urban Economics*, 18, 301–310.
Margo, Robert A. 1992. "Explaining the Postwar Suburbanization of Population in the United States: The Role of Income," *Journal of Urban Economics*, 31, 301–310.
McDonald, John F. 1989. "Econometric Studies of Urban Population Density: A Survey," *Journal of Urban Economics*, 26, 361–385.
McDonald, John F. 1997. *Fundamentals of Urban Economics*. Upper Saddle River: Prentice Hall.
McMillen, Daniel P. 2004. "Testing for Moncentricity," in Richard Arnott and Daniel McMillen, (eds.), *A Companion to Urban Economics*, forthcoming.
Mieszkowski, Peter and Edwin S. Mills. 1993. "The Causes of Metropolitan Suburbanization," *Journal of Economic Perspectives*, 7, 135–147.
Mills, Edwin S. 1972. *Studies in the Structure of the Urban Economy*. Baltimore: Johns Hopkins Press.
Mills, Edwin S and Bruce W. Hamilton. 1989. *Urban Economics*. 4th Edition. Glenview: Scott, Foresman and Company.
Muth, Richard. 1968. *Cities and Housing*. Chicago: University of Chicago Press.
O'Sullivan, Arthur. 2007. *Urban Economics*. Sixth Edition. Chicago: McGraw-Hill.
Rosenthal, Stuart and William Strange. 2004. "Evidence on the Sources and Nature of Agglomeration Economies," in Vernon Henderson and Jacques-Fran coise Thisse, (eds.), *Handbook of Regional and Urban Economics*, Volume 4, North-Holland: Elsevier.
Steinnes, Donald N. and Walter D. Fisher. 1974. "An Econometric Model of Intraurban Location," *Journal of Regional Science*, 14(1), 65–80.
U.S. Bureau of the Census. *Census of Manufactures*, various years.
——. *Census of Population*, various years.
——. *Census of Population and Housing*, various years.
——. *County and City Data Book*, various years.
——. *Financial Statistics of Cities*, various years.
——. *Population of States and Counties of the United States: 1790–1990*.
——. *Social Statistics of Cities*, various years.
——. *State and Metropolitan Area Data Book*, various years.
Warner, Sam Bass Jr. 1978. *Streetcar Suburbs*. Cambridge: Havard University Press.
Weber, Adna F. 1899. *The Growth of Cities in the Nineteenth Century*. Ithaca: Cornell University Press.