Urban Centralization in U.S. Cities prior to World War II: Experience with a New Metric

Kevin Skadron

28 April 1994

Abstract

Effective urban management and planning requires an understanding of the forces driving cities' development. Trends in population density and distribution are of particular importance, but are difficult to measure, and recent studies have come to conflicting conclusions regarding the rate of decentralization in U.S. cities. We use a new metric to obtain new evidence regarding decentralization in U.S. cities prior to World War II. The new evidence is not, however, conclusive, and although the new metric may potentially be useful, we remain skeptical of the its ability to yield reliable results.

1 Introduction

The structure and evolution of city organization has long interested economists. Changes in urban density patterns over time yield insight into residents' preferences for housing, neighborhood, and commuting, and how these vary as technology and the economy change.

Drawing conclusions about general trends in U.S. cities is, however, difficult. Cities are complex and not easily modeled, and a variety of city types exist in the U.S. Consider New York and Houston as examples: New York is dense and has evolved over several hundred years; Houston is sparse, with almost uniform density, and has seen its most significant development only in the postwar era.

While there is general agreement that in the postwar era cities have decentralized, studies for the pre-WWII era draw conflicting conclusions. Clark in 1951 [2] and later Mills, Muth, and Barr (cited in [3]) find significant decentralization of pre-war U.S. cities. These studies use a monocentric model, with a single central business district (CBD), an exponential falloff in population density with distance from the CBD, and equilibrium adjustments to urban spatial structure as employment, commuting costs, and income change over time. The model's defining equation is:

$$D(x) = D_0 e^{-\theta x} \tag{1}$$

Harrison and Kain, in 1973, emphasize in [3] the "disequilibrium" nature of urban growth due to durability of urban structures. They propose that new buildings reflect equilibrium conditions in commuting costs, income, etc. which exist at the time of building, but that existing buildings are torn down or converted at a negligible rate. They conclude that pre-war U.S. cities exhibit little or no decentralization, and may in fact exhibit centralization.

Mills in 1991 [6] proposes a new metric for measuring urban decentralization which does not assume a monocentric city structure and is better suited for examination of cities with multiple CBD's and varying central city land areas. He proposes the following relationship:

$$P_c = AL_c^\alpha P^\beta \tag{2}$$

$$\alpha \leq 1, \beta > 0, \alpha + \beta > 1$$

Except for a study of the postwar Germanys in [7], to our knowledge this measure has only been used to study determinants of suburbanization, such as race, income, and crime, and then only in modern years [4, 6]. We therefore believe application of (2) to be an untested technique for examining trends in U.S. urban density.

We report results from application of (2) to search for trends in urban density over the period 1910–1940 in pre-WWII U.S. cities. Section 2 of this paper describes our method, Section 3 presents the findings, and Section 4 draws some conclusions. Section 5 suggests avenues for further work in this area.

2 Method

We use city data provided in *The Growth of Metropolitan Districts in the United States: 1900–1940*, a Census publication authored by Thompson [8], and in *Population Growth in Standard Metropolitan Areas: 1900–1950*, by Bogue [1]. Thompson provides central city population (P_c), central city land area (L_c), and central city population density (ρ_c) for the years 1910, 1920, 1930, and 1940. Bogue provides SMSA and central city populations for these cities for the same years plus 1950, but provides central city land areas and population densities only for 1950.

We choose cities with only a single central city within the SMSA and which were SMSA's by 1910. This gives a sample of 55 cities, listed in Appendix A. We run two sets of regressions: one in which all P_c's are normalized to 1910 boundaries and results are compared across all four decades; and one in which decades are compared pairwise, with P_c's normalized to the earlier decade, and results compared only within pairs. The samples used in each case are varied in order to search for trends limited to subsets of cities and in order to verify stability of the results.

Central city boundaries vary over the period 1910–1940, and in order to eliminate distortions from jurisdictional changes, normalization of P_cs is required. Normalization is achieved between two decades x and y by using the following formula, which normalizes y to x:

$$P_c^y(corr) = P_c^y - \rho_c^y(L_c^y - L_c^x) \tag{3}$$

This assumes the average central city population density ρ_c holds in the area which has been added or removed over the decade, and subtracts this much population from P_c in order to get the P_c^y which would hold in decade y were the cities' boundaries unchanged from x. Observe that this is likely to subtract more population from the parcel of land in question than actually resides there, as jurisdictional changes are likely to occur at the fringe of a central city, where density is lower than nearer the center. This error, however, makes $P_c(corr, estimated) \leq P_c(corr, actual)$, and hence makes centralization harder to prove. Any conclusion of centralization or lack of decentralization is thus not weakened by this error induced in normalization.

For regression purposes, logs of (2) are taken to obtain

$$\ln P_c = (\ln A)\alpha L_c \beta P_{SMSA} \tag{4}$$

Regressions are performed using Microsoft's Excel 5.0 for Windows.

3 Results

The first set of regressions operates on the set of 35 cities for which Thompson has complete L_c data for 1910–1940. Regressions are run for all four decades on the entire set, on a subset consisting of only cities whose L_c's vary by less than 5% over the period, and on a subset of this consisting of only cities whose P_{SMSA}'s are below 500,000 for the period. Calculated regression statistics appear in Tables 1, 2, and 3, respectively. The first set of values are plotted in Figures 1 and 2; Figure 1 plots P_c vs. P_{SMSA} for values from Table 1 for P_{SMSA} out to 5 million, and Figure 1 also plots P_c vs. P_{SMSA} for values from Table 1 but for P_{SMSA} out to only 1 million.

Figure 1 appears to suggest a strong decentralizing trend, but the curves beyond P_{SMSA} are fairly unreliable due to a small sample size in this region. Figure 2 is more reliable, but unfortunately not very informative; plots for values of L_c other than 100 sq. mi. are equally uninformative. Plots for the data in Table 2, not presented here, are again uninformative. The results in Table 3 are unreliable.

An alternative method to search for trends is to compare the decades pairwise. Regressions are run for the pairs 1910–1920, 1920–1930, and 1930–1940; regression statistics are presented for the set of cities whose L_c's vary by less than 10% over the decade in question, and for subsets of these cities whose P_{SMSA}'s are less than 500,000, less than 600,000, and greater than 500,000 at the beginning of the decade in question. The data appear in Tables 4–7. Data from Table 4 is plotted in Figures 3–5.

These plots suggest mild centralization over the period in question, and the R^2 values and t-values are encouraging. But the subsample results are discouraging. The figures in Tables 4, 5, 6, and 7 vary widely, indicating that the 1920–1930 results may not be stable.

4 Conclusions

The first set of results, obtained by examining all the decades together, are inconclusive. The second set of results, obtained by examining the decades pairwise, suggest possible decentralization in the pre-war era, especially during the Roaring Twenties. Such a finding confirms those of Harrison and Kain and contradicts those of Clark, Mills, Muth, and Barr. Yet the regression results are suspect due to lack of stability when the sample is varied.

In general, we think this approach holds some promise for resolving the debate over centralization versus decentralization in the pre-war era, if the technique can be sufficiently refined in order to obtain significant and stable results. We remain skeptical, however, of the likelihood of obtaining reliable results using (2).

5 Future Work

The P_{SMSA} values here contain both urban and rural ring populations. It is possible that the very low densities in the rural rings are distorting the results; this may be particularly true in the case of southern cities, which have a large percentage of rural population in their SMSA's [5]. One unexplored avenue is to remove these rural ring populations from the P_{SMSA}'s and redo the regressions.

U.S. cities do vary widely in structure and density; this suggests analyzing cities of only a certain type. For example, it may be that only older cities like Philadelphia, Baltimore, etc. exhibit centralization.

Another potential avenue, perhaps to be explored in conjunction with either of the above, is to analyze subsamples of cities in more detail than has been performed in this study, searching for reliable trends among subsamples. Groups can be formed according to city size, density, age, and so forth.

A more farfetched possibility is to apply a different function to the data. Finding an interpolating polynomial or a spline for each decade, or for each city, may yield trends across the fits which shed light on urban evolution.

6 Acknowledgements

I would like to express deep thanks to Dr. Peter Mieszkowski of Rice's Department of Economics for his guidance and assistance both in choosing this area of study and in obtaining and analyzing the results.

References

[1] D. J. Bogue. Population growth in standard metropolitan areas, 1900–1950. Housing and Home Finance Agency, Washington: 1953.

[2] C. Clark, Urban population densities. *Journal of the Royal Statistical Society Series A*, **114**, 490–496 (1951).

[3] D. Harrison, Jr. and J. F. Kain. Cumulative urban growth and urban density functions. *Journal of Urban Economics*, **1**, 61–98 (1974).

[4] R. Kawal. Effects of central city problems on suburbanization, 1950–1990. Draft paper, Rice University Dept. of Economics, August 1993.

[5] P. Mieszkowski. Personal communication, April 1994.

[6] E. S. Mills. The measurement and determinants of suburbanization. *Journal of Urban Economics*, **32**, 377-387 (1992).

[7] K. Skadron. A new method for measurement of ubran decentralization. Class paper, Rice University Dept. of Economics, May 1993.

[8] W. S. Thompson. Population; the growth of metropolitan districts in the united states: 1900–1940. Publication of the Bureau of the Census, 1947.

	1940	1930	1920	1910
A	53.5942	41.4626	8.9532	1.7810
	(4.4401)	(4.1074)	(2.2110)	(0.8466)
α	0.4253	0.3784	0.1911	0.1427
	(4.0065)	(3.5406)	(1.7071)	(1.9216)
β	0.5325	0.5657	0.7396	0.8838
	(6.6152)	(6.9671)	(8.2631)	(14.2133)
Adjusted R^2	0.8068	0.7986	0.7910	0.9108
sample size	35			

Table 1: *Regression statistics for the four decades taken together, with a sample consisting of the 35 cities for which full data for 1910–1940 is available.*

	1940	1930	1920	1910
A	3.1537	3.2705	1.7094	1.1479
	(1.3606)	(1.4576)	(0.6267)	(0.1475)
α	0.3871	0.3786	0.3260	0.2737
	(3.2769)	(3.4145)	(2.7617)	(2.1667)
β	0.7652	0.7660	0.8302	0.8722
	(9.4819)	(10.0546)	(10.3006)	(9.9288)
Adjusted R^2	0.9463	0.9467	0.9424	0.9312
sample size	15			

Table 2: *Regression statistics for the four decades taken together, with a sample consisting of only those cities for which L_c's vary by less than 5% over 1910–1940.*

	1940	1930	1920	1910
A	44.9240	89.0428	207.1575	1004.9830
	(1.3432)	(1.7428)	(2.2297)	(2.7846)
α	0.4040	0.3602	0.2645	0.2158
	(1.8601)	(1.8959)	(1.5285)	(1.3270)
β	0.5502	0.5083	0.4604	0.3354
	(2.3013)	(2.4281)	(2.3765)	(1.6748)
Adjusted R^2	0.6005	0.5614	0.4805	0.2712
sample size	9			

Table 3: *Regression statistics for the four decades taken together, with a sample consisting of only those cities for which L_c's vary by less than 5% and whose SMSA populations are less than 500,000 over 1910–1940.*

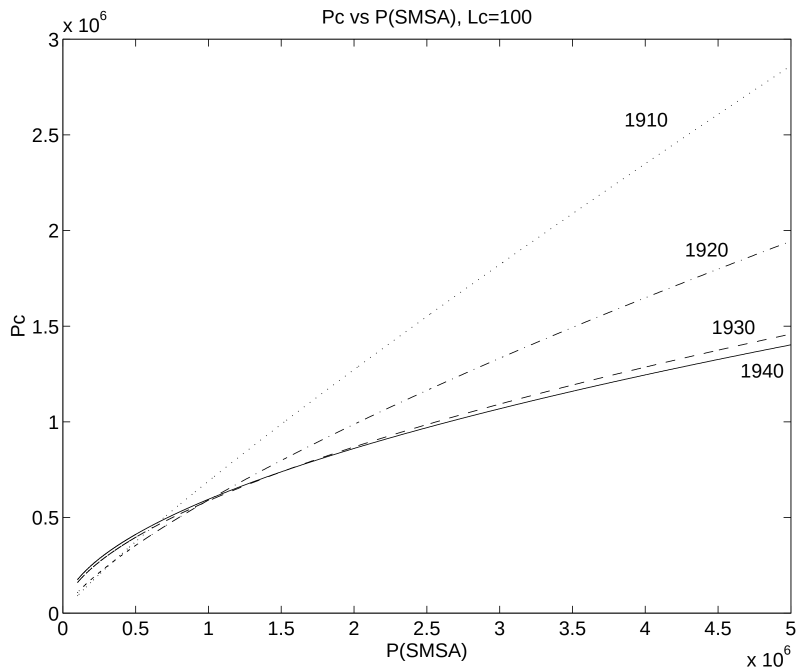

Figure 1: P_c vs. P_{SMSA} for values from Table 1 for P_{SMSA} out to 5 million. Here $L_c = 100$ sq. mi.

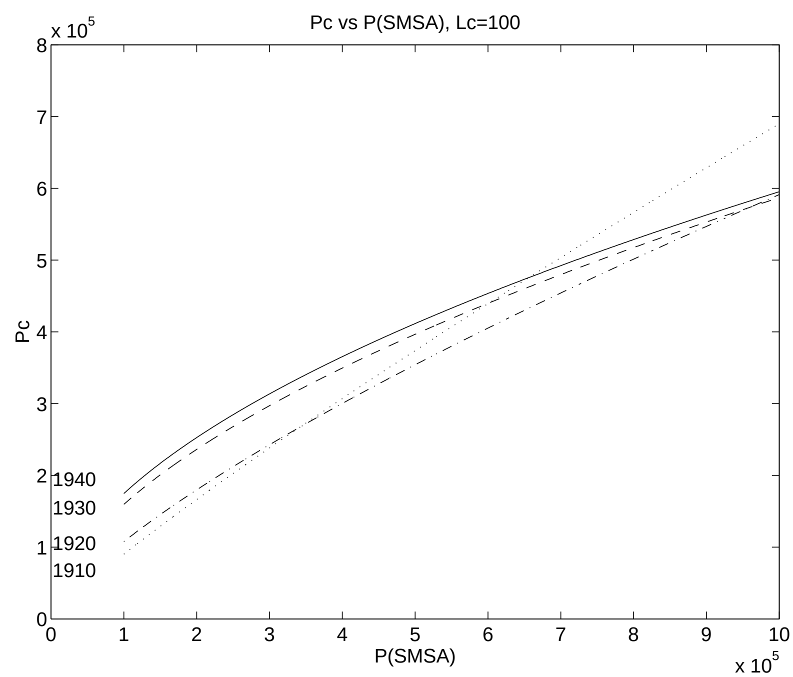

Figure 2: P_c vs. P_{SMSA} for values from Table 1 for P_{SMSA} out to 1 million. Here $L_c = 100$ sq. mi.

	1940	1930	1930	1920	1920	1910
A	3.9868	4.0189	3.0796	2.5572	4.3872	3.0195
	(2.3136)	(2.3992)	(1.9326)	(1.2843)	(2.1743)	(1.5660)
α	0.3112	0.3038	0.2395	0.1196	0.3016	0.2738
	(6.1590)	(6.2196)	(4.0918)	(1.6233)	(3.1182)	(2.7419)
β	0.7677	0.7713	0.8120	0.8555	0.7683	0.8013
	(14.1630)	(14.7196)	(15.0890)	(12.6720)	(11.2759)	(11.2794)
$Adjusted R^2$	0.9405	0.9391	0.9493	0.9077	0.9398	0.9342
sample size	45		31		25	

Table 4: *Regression statistics for the decades taken pairwise: 1910–1920, 1920–1930, and 1930–1940. A pair's sample consists of cities for which data exists for the endpoints of the decade and whose L_c's vary by less than 10% over the decade.*

	1930	1920
A	87.7507	291.9551
	(3.1004)	(4.2739)
α	0.2274	0.1048
	(3.6999)	(1.9536)
β	0.5472	0.4749
	(4.5556)	(4.3255)
Adjusted R^2	0.7109	0.5436
sample size	22	

Table 5: *Regression statistics for the 1920–1930 pair, with the sample reduced to include only cities with $P_{SMSA} < 500,000$ in 1920.*

	1930	1920
A	30.8643	28.4999
	(3.2101)	(3.0325)
α	0.2369	0.1260
	(4.1311)	(2.1496)
β	0.6289	0.6597
	(7.0842)	(7.2317)
Adjusted R^2	0.8041	0.7327
sample size	25	

Table 6: *Regression statistics for the 1920–1930 pair, with the sample including only cities with $P_{SMSA} < 600,000$ in 1920.*

	1930	1920
A	1.7909	0.0654
	(0.5167)	(-1.0093)
α	0.3445	0.0012
	(3.0374)	(0.0045)
β	0.8211	1.1575
	(7.7949)	(4.5688)
Adjusted R^2	0.9741	0.8707
sample size	9	

Table 7: *Regression statistics for the 1930–1940 pair, with the sample reduced to include only cities with* $P_{SMSA} > 500,000$ *in 1920.*

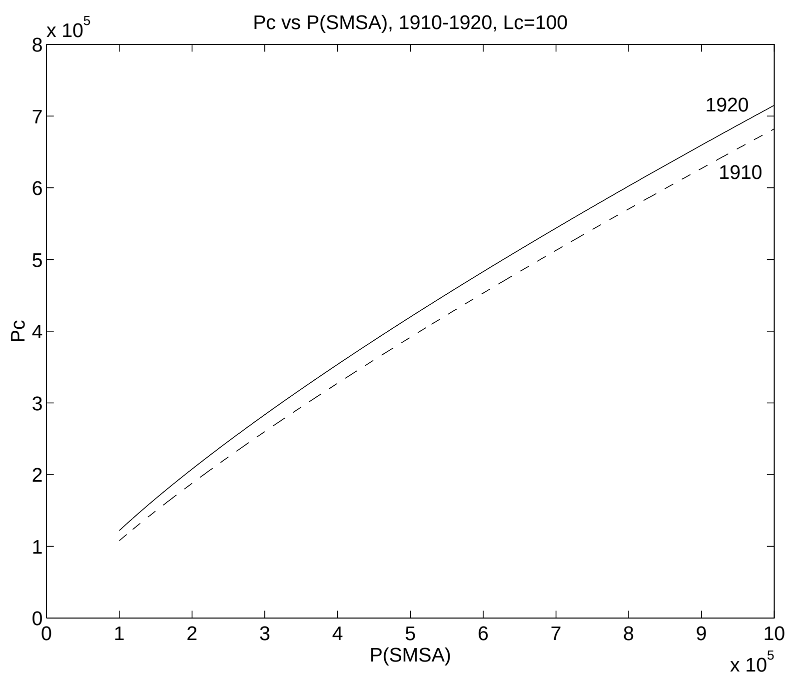

Figure 3: P_c vs. P_{SMSA}, 1910–1920, for values from Table 4 for P_{SMSA} out to 1 million. Here $L_c = 100$ sq. mi.

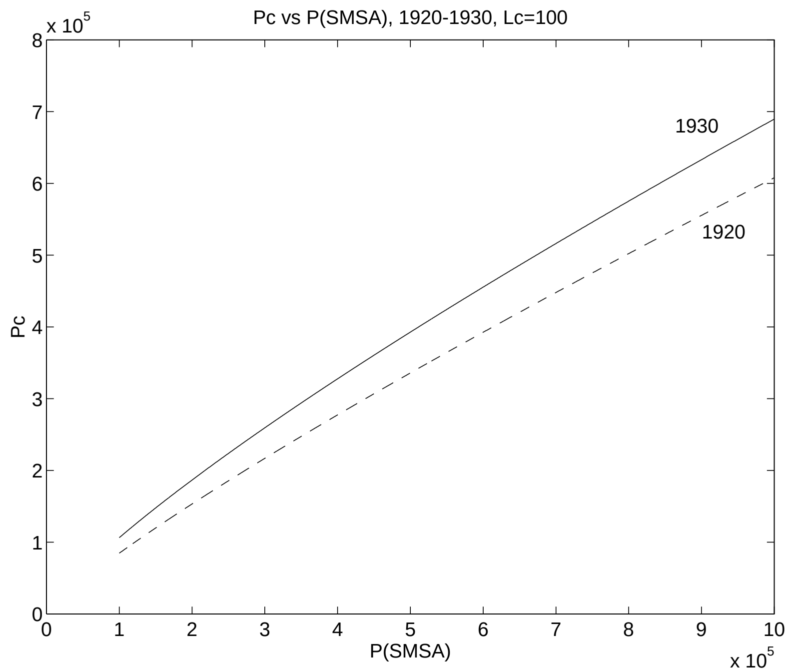

Figure 4: P_c vs. P_{SMSA}, 1920–1930, for values from Table 4 for P_{SMSA} out to 1 million. Here $L_c = 100$ sq. mi.

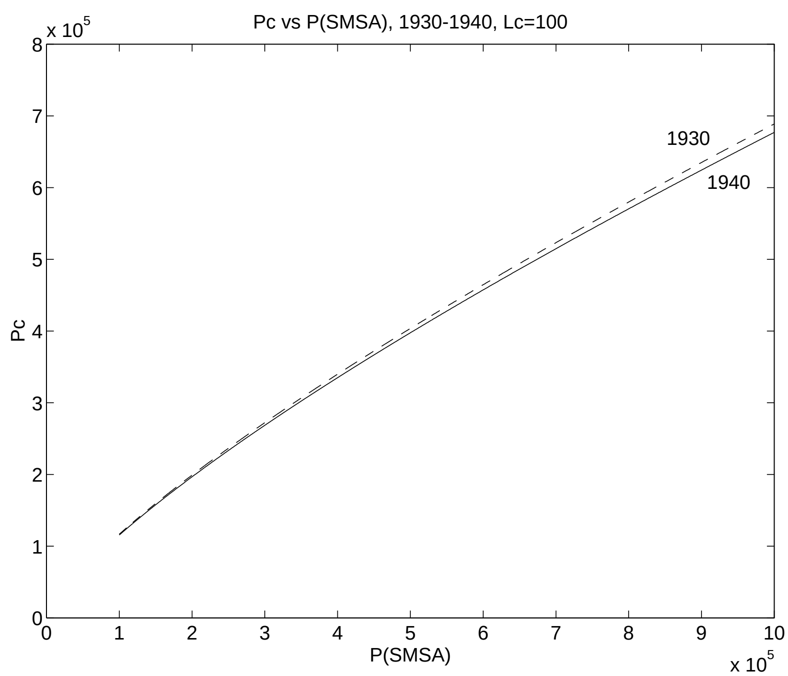

Figure 5: P_c vs. P_{SMSA}, 1930–1940, for values from Table 4 for P_{SMSA} out to 1 million. Here $L_c = 100$ sq. mi.

A Cities Examined

Akron	OH	Milwaukee	WI
Altoona	PA	Nashville	TN
Atlanta	GA	New Orleans	LA
Baltimore	MD	Omaha	NE
Birmingham	AL	Peoria	IL
Brockton	MA	Philadelphia	PA
Buffalo	NY	Pittsburgh	PA
Canton	OH	Portland	ME
Chicago	IL	Portland	OR
Cincinnati	OH	Providence	RI
Cleveland	OH	Reading	PA
Columbus	OH	Richmond	VA
Dallas	TX	Rochester	NY
Dayton	OH	St. Louis	MO
Denver	CO	Salt Lake City	UT
Des Moines	IA	Scranton	PA
Detroit	MI	Seattle	WA
Erie	PA	Spokane	WA
Fort Worth	TX	Syracuse	NY
Grand Rapids	MI	Tacoma	WA
Harrisburg	PA	Toledo	OH
Houston	TX	Trenton	NJ
Indianapolis	IN	Washington	DC
Johnstown	PA	Wilmington	DE
Kansas City	MO	Worcester	MA
Los Angeles	CA	Youngstown	OH
Louisville	KY		
Manchester	NH		
Memphis	TN		

Table 8: *Cities used in this study. These are cities which were SMSA's in 1910 and only have a single central city in the SMSA. Brockton, Manchester, and Portland, ME were not used in any regressions due to lack of available data for them.*