Available online at www.sciencedirect.com
SCIENCE DIRECT
Journal of Urban Economics 58 (2005) 1–10
JOURNAL OF Urban Economics
www.elsevier.com/locate/jue

More evidence on the spatial scale of cities

Daniel T. McGrath

Associate Director, Institute for Environmental Science and Policy, University of Illinois at Chicago, 528 SPHW (MC 673), 2121 W. Taylor Street, Chicago, IL 60612, USA

Received 15 May 2003; revised 13 January 2005

Available online 17 March 2005

Abstract

This study re-examines the 1983 study by Brueckner and Fansler that empirically estimated the determinants of urbanized land areas, regressing land area on population, income, transportation costs, and agricultural land values. This study, however, utilizes a larger, more comprehensive data set of metropolitan statistical areas over a longer period of time. The estimation results, consistent with Bruecker and Fansler, confirm that the simple monocentric model is empirically robust and that the standard economic factors identified by urban economic theory explain the majority of spatial variation in the sizes the largest US metropolitan regions over the post-war period.

1. Introduction

The debate within the urban policy literature on the causes and consequences of urban sprawl is mature, with much of the more recent literature focused on a more precise definition and measurement of the phenomenon (Bento and Franco [2], Brueckner [3], Fulton, et al. [7], Galster, et al. [8], Lopez and Hynes [12], and Malpezzi [13]). Brueckner and Fansler's widely cited paper [5] is the first paper to place the topic of urban sprawl within standard urban economic theoretical framework. The authors empirically estimated the equation for urban spatial size derived from the standard monocentric urban model, regressing urban land area on population, income, transportation costs, and agricultural land values. In addition to finding that the parameter estimates of most of these variables, with

E-mail address: dmcgrath@uic.edu.

0094-1190/$ – see front matter
doi:10.1016/j.jue.2005.01.003

the exception of transportation costs, were significant and consistent with theory, their estimated model explained nearly eighty percent of the variance in urban spatial sizes within their data set. This strong result prompted the authors to conclude that a measured view of urban sprawl was warranted; urban structure was, consistent with theory, the result of an orderly economic process.

There were, however, significant limitations in the original data set upon which this early study was based. The authors used a data set of only 40 observations of small urban areas, each contained within one county, observed in only one year (1970). A re-exploration of the approach taken by these authors in determining whether economic factors explain the scale of metropolitan areas would help elucidate at least one issue in the current debate on urban sprawl.

This research note seeks to revisit the same monocentric urban framework as Brueckner and Fansler used in their original study while utilizing a larger, more comprehensive data set that includes observations over multiple years. The methodology employed is cross-sectional time-series analysis of urbanized land areas, as reported by the Census Bureau for the 33 largest US metropolitan statistical areas from 1950 through 1990. Metropolitan urbanized land areas are regressed on reported measurements of population, income, transportation costs, and agricultural land values for each city at each decade. In addition to the standard factors identified by urban economic theory as explaining urban land areas, a second model with a time trend variable is included. It is hypothesized that the time trend variable will capture the other possible systematic factors contributing to urban spatial size over time. The following section presents a short review of the economic theory relevant to the determination of urban land area and derives the functional form for the model used in this analysis. It is followed by a discussion of the data and the estimation results, before presenting conclusions for urban policy.

2. Theory

The fundamental urban economic theory relevant to the concept of urban sprawl is the monocentric urban model of Alonso [1], Mills [17], and Muth [20]. While total urban land area is not directly a parameter of interest, in the monocentric model it is most closely related to $\bar{x}$, the radius of the urbanized area. The equilibrium conditions of the monocentric model are expressed as:

$$\int_0^{\bar{x}} \frac{2\pi x}{q(x, y, t, u)}\,\mathrm{d}x = n, \tag{1}$$

$$r(\bar{x}, y, t, u) = r_a, \tag{2}$$

where the q function gives individual land consumption as a function of the model's parameters and the r function gives urban land rent.

Equations (1) and (2) require that the total urban population, n, fits inside the urban area and that the urban rent at $\bar{x}$ is equal to the agricultural rent, r_a, respectively. These equations determine the equilibrium values for u, the utility level for the metropolitan region, and for $\bar{x}$. The parameters y and t represent household income and transportation costs respectively. Wheaton [21] first presented a comparative static analysis of (1) and (2) to identify how the model's parameters influence urban spatial size. The results are expressed as:

$$\bar{x} = g\,[\underbrace{n, y}_{+}, \underbrace{t, r_a}_{-}]. \tag{3}$$

Equation (3) indicates that urban spatial size is increasing in population and income and decreasing in transportation costs and agricultural rent.

As discussed above, Brueckner and Fansler [5] present the only empirical exploration of Eq. (3). Analyzing land area data from approximately 40 small metropolitan regions from the 1970 census that fit inside one county, the authors find that population, income, and agricultural rent are statistically significant determinants of total urban land area. However, they do not find significance for transportation costs, indirectly measured both as the percentage of commuters using public transportation and as the percentage of households owning one or more automobiles. The authors argue that the significance of the empirical results support the view that urbanization is primarily the result of a rational market process and suggest that a dispassionate view of urban sprawl is warranted.

It is well known that the standard monocentric model of urban structure is not an accurate representation of reality and that there are other factors that influence the size of cities. Over the past several decades, nearly all of the US metropolitan regions have taken on a more complex, polycentric structure (Fujita and Ogawa [6], Garreau [9], McDonald [14], and McMillen and McDonald [15]), and it is likely that the appearance of multiple employment centers in urban areas has systematically contributed to the increase in urbanized land areas throughout the post-war period.

Additionally, fiscal and social disparities between central cities and suburbs have been identified as significantly impacting the spatial scale of metropolitan areas [19]. It is well known that US metropolitan areas are fiscally stratified. Higher income households reveal their preferences for housing and municipal services by creating fiscally autonomous suburban municipalities, which over the post-war years has accelerated central city decline as well as fueling the suburbanization process. Also recognized, however, is the difficulty in empirically capturing these effects [16].

Lastly, market failures in the urban development process may have contributed significantly to increased urban spatial scale. Brueckner [3,4] identifies three of these:

(1) the failure of land markets to internalize the social values associated with open space at the urban fringe,
(2) the failure of households to internalize the congestion costs generated in urban transportation, and
(3) the failure of municipal governments to accurately estimate the fiscal burden of development.

It is hypothesized that these combined effects—the emergence of polycentric employment subcenters, the fiscal stratification of metropolitan regions, and market failures in the development process—have all been in operation through the post-war period, and that they have systematically contributed to increases in urbanized land areas in the US. It is further hypothesized that the combined influence of these other factors is likely to have been a constant percentage of the urban area at any point in time. If this is true, then the inclusion of a time variable, τ, would serve as a proxy variable to control for these combined effects. Thus in this generalized model, the distance to the urban fringe for a given metropolitan region, i, at time τ is a function of the standard variables identified by Wheaton [21] and the additional time variable, τ:

$$\bar{x}_{i,\tau} = f\left(n_{i,\tau}, y_{i,\tau}, t_{i,\tau}, r^a_{i,\tau}, \tau\right). \tag{4}$$

While the economic literature does not specify a functional form for Eqs. (3) and (4), the form used for estimation in this study is consistent with logarithmic functional forms identified for metropolitan density gradients identified by Mills [18]. This is expressed as:

$$\bar{x}_{i,\tau} = n^{\alpha}_{i,\tau} \mathrm{e}^{\beta_0 + \beta \cdot \mathbf{Z}_{i,\tau} + \gamma\tau + \varepsilon} \tag{5}$$

where $\mathbf{Z}$ is the vector of the variables y, t, and r_a. Taking logs of both sides, the following expression is obtained:

$$\ln \bar{x}_{i,\tau} = \alpha \ln n_{i,\tau} + \beta_0 + \beta \cdot \mathbf{Z}_{i,\tau} + \gamma \cdot \tau + \varepsilon. \tag{6}$$

3. Discussion of data

The key variable of interest is the total urbanized land area in square miles for the 33 largest US metropolitan areas from the decennial US censuses from 1950 through 1990.[1] The urbanized land area measurements made by the US Census Bureau attempt to exclude non-urban (rural) land and measure contiguously developed land. This is, however, recognized to be an imprecise measure. The general definition of the US Census Bureau requires urbanized development to be at or above a density of 1000 persons per square mile. In the United States, the US Census Bureau defines metropolitan statistical areas by county boundaries in 44 states. Areas of urban development are further defined by major government jurisdictional boundaries, and it is possible that these jurisdictional areas can include comparatively large expanses of rural land.[2] Using the expression, $A = \Pi\bar{x}^2$, these observations have been converted to a proxy variable for city radius equal to $(A/\Pi)^{1/2}$. Additional census data include metropolitan area population and real per capita personal income for the years 1950 through 1990 for the 33 metropolitan areas. Table 1 identifies the urbanized land areas in square miles for the 33 metropolitan areas included in the data, ranked in descending order based on 1990 areas.

A proxy variable for transportation costs is obtained from two sources. The average annual consumer price index (CPI) for private transportation, published by the Bureau of

[1] The Phoenix, AZ MSA was the 15th largest metro region in 1990. However, due to the unavailability of transportation cost, and agricultural land value data, this MSA was excluded.

[2] The exception is metropolitan areas within the 6 New England states (the Boston metro area in this data set). These metropolitan areas are defined by the US Census Bureau by municipal boundaries and contain relatively little rural land.

Table 1
Total urban land area in square miles

	Urbanized area	1950	1960	1970	1980	1990
1	New York	1253	1892	2425	2808	2967
2	Los Angeles	871	1370	1572	1827	1966
3	Chicago	708	960	1277	1498	1585
4	Dallas-Fort Worth	262	920	1071	1280	1443
5	Houston	270	431	539	1049	1177
6	Philadelphia	312	597	752	1015	1164
7	Atlanta	106	246	435	905	1137
8	Detroit	423	732	872	1044	1119
9	Minneapolis-St. Paul	231	657	721	980	1063
10	Washington	178	341	495	807	945
11	Boston	345	516	664	857	891
12	San Francisco-Oakland	287	572	681	796	874
13	Pittsburgh	254	525	596	713	778
14	Kansas City	149	282	493	589	762
15	St. Louis	228	323	461	597	728
16	San Diego	133	276	381	611	690
17	Norfolk	62	109	299	418	664
18	Tampa-St. Petersburg	180	218	291	527	650
19	Cleveland	300	587	646	629	636
20	Baltimore	152	220	310	523	593
21	Seattle	123	238	413	485	588
22	Cincinnati	146	242	335	420	512
23	Milwaukee	102	392	457	496	512
24	Riverside-San Bernardino	61	169	310	359	460
25	Denver	105	167	293	439	459
26	San Antonio	90	192	223	354	438
27	Portland	114	192	267	349	388
28	Miami	117	183	259	340	353
29	San Jose	61	223	277	326	338
30	Sacramento	42	134	254	278	334
31	Fort Lauderdale	–	124	212	289	327
32	Buffalo	123	160	214	266	286
33	New Orleans	222	267	184	230	270
	Total	8010	14,457	18,679	24,104	27,097

Labor Statistics, was obtained for each metropolitan region for each census year. However, a major limitation of the regionally reported private transportation CPI measure is that it cannot capture regional differences; it only measures how much prices have changed in a given area. In an attempt to capture relative private transportation price differentials across US regions, comparative private transportation cost data for 1990, published by the American Chamber of Commerce, was used to rescale the 1990 CPI data. Arbitrarily setting Atlanta in 1990 to a value of 100, the 1990 values of the other 32 metropolitan regions were adjusted according to their deviation from the national average. The remaining decadal transportation cost figures were transformed to be consistent with adjusted 1990 values. Some of the metropolitan regions were missing CPI data for 1950 and 1960. This resulted in a total of 153 complete observations.

Table 2
Statistical summary of the metropolitan regional data

Variable	n	Min	Max	Mean	Std. Dev.
MSAULA Urbanized land area total in square miles for metro region	153	42	2967	588	498
XBAR Distance to urban fringe in miles calculated as the sqrt of MSAULA/π	153	3.66	30.73	12.70	5.07
MSAPOP Population in thousands for metropolitan region	153	136	16,207	2345	2883
RPINC Real per capita personal income for the metro region in 1990$.	153	3526	22,049	10,776	3674
APTCPI Regionally-adjusted private transportation consumer price index for the metropolitan region Atlanta in 1990 = 100	153	17.79	111.86	54.34	33.22
RAGVAL Real average per acre agricultural land value for state of metro region in 1989$	153	132	4056	914	650
DECADE $1 = 1950, \ldots, 5 = 1990$	153	1	5	3.06	1.42

Agricultural rent at the urban fringe, r_a, is proxied by nominal agricultural land per acre, as reported annually for all states by the USDA Economics and Statistics Office and converted to 1990 dollars. The agricultural land values for New York, Philadelphia, Pittsburgh, and Washington DC were calculated as an average for their respective adjoining states. For example, Washington DC's value is the average of values for Virginia and Maryland. Table 2 presents a statistical summary of the data used in this analysis.

4. Estimation results

The parameter estimates for two model specifications are presented in Table 3. Model 1 is based on Eq. (3), using the functional form of Eq. (6). The dependent variable is LNXBAR, the natural log of the transformed variable, $\bar{x}$, and the independent variables include the proxies for the four independent variables $\ln(N)$, y, t, and r_a. The t-statistics are in parentheses. Model 2 is based on Eq. (6), which includes the time variable, τ, hypothesized to capture the other possible systematic effects discussed above.

The OLS estimation results of Model 1 show that the intercept and the variables for population (LNPOP), real personal income (RPINC), and real agricultural land values (RAGVAL) are all significant with signs consistent with expectations. However, the co-

Table 3
OLS estimation results (dependent variable: LNXBAR)

	Model 1	Model 2
Intercept	−0.641 (6.25)	−0.672 (6.33)
LNPOP Natural Log of MSA population in thousands	0.376 (24.34)	0.382 (24.39)
RPINC Real personal income in 1990$	0.0000365 (5.83)	0.0000153 (1.74)
RAGVAL Real state average per acre agricultural land values in 1989$	−0.0000467 (2.01)	−0.0000547 (2.42)
APTCPI Regionally-adjusted private transportation consumer price index (Atlanta in 1990 = 100)	−0.000169 (0.24)	−0.00255 (2.58)
DECADE 1 = 1950, . . . , 5 = 1990	–	0.116 (3.33)
Adjusted R^2	0.871	0.879

Notes: The absolute values of the t-statistics are presented in parentheses. $N = 153$. The dependent variable is LNXBAR, which is the natural log of XBAR, the calculated distance to the urban fringe in miles.

efficient of the adjusted personal transportation cost index (APTCPI), is not significant, though its sign is consistent with theory.

In Model 2, which includes the time trend variable, DECADE, all the coefficients of the independent variables emerge as significant, but the explanatory power of the model improves only slightly. Importantly, the coefficient of the transportation cost variable, APTCPI, is significantly negative, as predicted, and the positive significance of the time variable, DECADE, suggests an independent systematic increase in urban land areas over time. The scale of this estimated effect is large, with the parameter indicating that urban land areas are on average 2.3% larger per year than can be explained by standard economic factors. While this result lends credibility to the view that market failures in the development process exist within US metropolitan regions, it is important to emphasize that the economic factors by themselves explain nearly 90% of the variation in the data, and that the addition of the time variable increases the explanatory power of the model by only 1 percentage point.[3]

The elasticities of urbanized land area[4] with respect to the model's exogenous variables, as identified by the parameters of Model 2, are evaluated at the sample means. The elasticity of urban land area with respect to population is 0.76 and shows that a one percent

[3] Unreported regressions utilizing a direct measurement of urban land area (the natural log of urban land area in square miles) as the independent variable produce nearly identical estimation results, with the exception of the significance of the constant term.

[4] Land area is determined from the expression $A = \Pi \bar{x}^2$.

increase in population leads to less than one percent increase in urban land area. This result differs substantially from the nearly unitary elasticity value derived by Brueckner and Fansler and suggests that, while average population density of US metropolitan regions has been declining in the post-war years, the maximum density may be increasing.[5] Population differences explain nearly 80% of the variation in the data.[6]

The elasticity of land area with respect to income is 0.33, which also stands in stark contrast to Brueckner and Fansler's estimate of 1.4. Lastly, the elasticity of land area with respect to transportation costs and agricultural land values are measured at 0.28 and 0.10 respectively. These low elasticity values indicate that these three factors, while significant, are clearly not as important as population growth in determining urbanized land areas.

5. Conclusion

Revisiting Brueckner and Fansler's approach, this paper explores empirically whether the spatial sizes of the 33 largest US metropolitan regions over the post-war years can be explained by the standard economic factors identified by urban economic theory. The estimation results presented here suggest that the answer is yes; fundamental economic factors are of primary importance in determining urban spatial sizes. Changes in population, income, transportation costs, and agricultural land values determine nearly 90% of the variation of urbanized land areas within the data. These results strongly support Brueckner and Fansler's assertion that "sprawl is the result of an orderly market process rather than a symptom of an economic system out of control" [5, p. 482].

However, it is important to emphasize that one cannot decisively conclude that US urban spatial size is socially optimal. The significance of the time variable in Model 2 indicates that factors beyond the "usual suspects" have made a systematic contribution to urbanized land areas, and the effect is quite large—about 2.3 percent per year or 26 percent per decade. It is not clear whether this unexplained effect is primarily the result of decentralized spatial employment patterns that have emerged within US metropolitan regions or whether market failures in the urbanization process are in operation.

The results presented here also provide some insight into policies for controlling externalities associated with the urbanization process. Most urban scholars concur that the correction of the current under-pricing of private transportation with the US would substantially improve the quality of life within US metropolitan regions. The negative significance of the transportation cost variable in Model 2 indicates that the adoption of policies that directly impact private transportation costs, such as congestion toll and fuel taxes, can have a direct impact on urban scale. Litman [11] provides convincing evidence that average US private transportation costs are under-priced by as much as 47%. Had private transportation costs been socially optimal, the results presented here suggest that US metropolitan

[5] The average population density figures for all 33 metropolitan regions in the data set declined from 5326 persons per square mile in 1950 to a low of 3156 persons per square mile in 1980. In 1990, average density in the data set actually rose slightly to 3189 persons per square mile.

[6] An unreported regression of LNXBAR on LNPOP has an adjusted R^2 value of 0.783.

regions would be as much 12% smaller on average.[7] To place this finding in perspective, it implies that over 3200 square miles of agricultural land may have been overdeveloped in these 33 metropolitan regions alone due to underpriced private transportation. Certainly, a similar discussion can be made as to under-pricing of agricultural land in land markets—prices that fail to internalize environmental costs of urbanization, specifically hydrological sustainability and ecosystem function damages [10]. While the long-term consequences of these losses are still unknown, it is because of these environmental social losses that the public debates over the need for effective policies to curtail urban expansion are likely to intensify.

The significance of this study for the debate on urban sprawl is twofold: First, it reinforces the fundamental conclusions of Brueckner and Fansler—urban expansion is primarily the result of orderly and predictable economic factors. Without question, any future policies designed to successfully curtail urban spatial growth in the US will need to engage these powerful economic factors. Second, this study presents evidence that there have been significant systematic increases in urbanized land areas that are not explained by the conventional economic factors identified by the monocentric model. Further research is needed to ascertain if the unexplained increases in urban land areas stem solely from non-market factors, such as market failures in the development process, or are the result of other economic factors such as emergence of multiple employment subcenters or a fundamental shift in consumer preferences.

Acknowledgments

I would like to express my thanks to Rachel Weber and Gib Bassett of the University of Illinois at Chicago, Jan Brueckner of the University of California, Irvine, and the two anonymous referees for their helpful comments and suggestions. I would also like to thank Joel Greenberg for his editorial assistance.

References

[1] W. Alonso, Location and Land Use, Harvard Univ. Press, Cambridge, MA, 1964.
[2] A. Bento, S. Franco, The economics of urban sprawl: Causes and optimal market-based instruments, Working paper, The University of California, Santa Barbara, Bren School of Environmental Science and Management, 2003.
[3] J.K. Brueckner, Urban sprawl: Diagnosis and remedies, International Regional Science Review 23 (2000) 160–171.
[4] J.K. Brueckner, Urban sprawl: Lessons from urban economics, Brookings-Wharton Papers on Urban Affairs (2001) 65–97.
[5] J.K. Brueckner, D. Fansler, The economics of urban sprawl: Theory and evidence on the spatial size of cities, Review of Economics and Statistics 55 (1983) 479–482.
[6] M. Fujita, H. Ogawa, Multiple equilibria and structural transition on non-monocentric urban configurations, Regional Science and Urban Economics 12 (1982) 161–196.

[7] Wheaton [22] shows that the imposition of a congestion toll in a simulated city reduces the distance to the urban fringe by 10%, which suggests that this externality is a significant cause of urban sprawl.

[7] W. Fulton, R. Pendall, M. Nguyen, A. Harrison, Who sprawls most? How growth patterns differ across the US, Brookings Institution, Washington, DC, 2001.
[8] G. Galster, R. Hanson, M. Ratchliffe, H. Wolman, S. Coleman, J. Freihage, Wrestling sprawl to the ground: Defining and measuring an elusive concept, Fannie Mae, Washington, DC, 2001.
[9] J. Garreau, Edge City, Doubleday, New York, 1991.
[10] M. Johnson, Environmental impacts of urban sprawl: A survey of the literature and a proposed research agenda, Environment and Planning A 33 (2001) 717–735.
[11] T. Litman, Transportation cost analysis for sustainability, Victoria Transport Policy Institute, 1999.
[12] R. Lopez, H. Hynes, Sprawl in the 1990s: measurement, distribution, and trends, Urban Affairs Review 38 (2003) 325–355.
[13] S. Malpezzi The measurement and determinants of sprawl, Working paper, University of Wisconsin Department of Real Estate, 1999.
[14] J.F. McDonald, The identification of urban employment subcenters, Journal of Urban Economics 21 (1987) 242–258.
[15] D.P. McMillen, J.F. McDonald, Suburban subcenters and employment density in metropolitan Chicago, Journal of Urban Economics 43 (1998) 157–180.
[16] P. Mieszkowski, E. Mills, The causes of metropolitan suburbanization, Journal of Economic Perspectives 7 (1993) 135–147.
[17] E. Mills, An aggregative model of resource allocation in a metropolitan area, American Economic Review, Papers and Proceedings 57 (1967) 197–210.
[18] E. Mills, Urban Economics, Scott Foresman, New York, 1972.
[19] E. Mills, R. Price, Metropolitan suburbanization and central city problems, Journal of Urban Economics 15 (1984) 1–17.
[20] R.F. Muth, Cities and Housing, Univ. of Chicago Press, Chicago, IL, 1969.
[21] W. Wheaton, A comparative static analysis of urban spatial structure, Journal of Economic Theory 9 (1974) 223–237.
[22] W. Wheaton, Land use and density in cities with congestion, Journal of Urban Economics 43 (1998) 258–272.