The Negative Exponential Decline of Density in Large Urban Areas in the U.S., 1950-2010

John R. Ottensmann
Indiana University-Purdue University Indianapolis
john.ottensmann@gmail.com
urbanpatternsblog.wordpress.com
October 2016

Abstract

The monocentric model predicts the negative exponential decline of density with distance from the center, but employment in urban areas has become increasingly dispersed in multiple employment centers. Using census tract data from 1950 to 2010, the parameters of the negative exponential model are estimated for 43 large urban areas. Density gradients have declined consistently. Overall, central densities likewise declined, with exceptions for some areas that have seen more rapid growth. The average fit of the model across the areas, measured by R^2, fell by nearly half after 1970, consistent with more polycentric employment patterns. Nevertheless, even for areas with extremely poor fit, negative exponential parameter estimates continued to be reasonable.

Introduction

Densities of population and housing units within urban areas have been shown to decline as a negative exponential function of distance from the center of the city for many years. Muth (1969) and Mills (1972) provided an economic foundation for understanding this with what has come to be known as the monocentric model, assuming a city in which employment is concentrated at the center. Numerous studies have shown that densities at the center have tended to fall over time, while the density gradient, the rate of decline with distance, has also decreased. The monocentric model continues to be advanced as a reasonable basis for examining urban structure.

From at least from the middle of the twentieth century up to the present, urban employment has tended to move out from the center to form a more dispersed pattern, with multiple outlying employment centers. This raises questions regarding the continued viability and performance of the monocentric model and its prediction of the negative exponential decline of density.

This paper examines the negative exponential decline of housing unit density from 1950 to 2010 in 43 large urban areas in the United States. A dataset with census tract housing unit counts for consistently defined urban areas is used for the estimation of the negative exponential model for each census year. This enables examination of the trends in both the estimated model parameters over time and the fit of the model as a measure of the extent to which urban densities conform to the negative exponential pattern.

The Monocentric Model and the Polycentric City

Clark (1951) showed that population densities decline with distance from the center of cities and that this holds for many cities at very different times. But it was Muth (1969) and Mills (1972) who provided an economic basis for understanding this phenomenon. They assumed that all employment was concentrated in the central business district (CBD), to which residents commuted. In selecting a location, residents were required to make a tradeoff between the costs of commuting and space. Commuting costs were reduced as one moved closer to the CBD, so residents were willing to pay more for those locations. This led to a decline in land rents with distance and the lower densities of development made possible by those lower rents. With some simple but reasonable assumptions regarding functional forms and parameters, the negative exponential decline of density, observed by Clark and others, was the result.

Mills (1972) estimated the parameters of the negative exponential decline of density for a number of metropolitan areas in the United States over an extended period of time, showing the tendency of density gradients to decline over time. He used an ingenious 2-point method for estimating the parameters using only the central city and total urban area populations and the approximate central city radius. As a result, he could not confirm the fit of the model or even the existence of a pattern of negative exponential decline. That had to be assumed.

Guest (1975) used census tract data to estimate the negative exponential model for 37 cities in the United States from 1940 or 1950 (depending on data availability) to 1970. He showed steady declines in both the density gradient and the central density. But his work also reported very large decreases in the fit of the negative exponential model, as measured by R^2, over the period.

The monocentric model and the assumption of the negative exponential decline of density has continued to receive attention up to the present. In a general discussion of urban spatial structure, Anas, Arnott, and Small (1998) cited numerous studies showing the decline of density gradients in the previous century. In his comprehensive study of urban patterns around the world, Angel (2012) extensively discussed the negative exponential decline of densities.

While work with the monocentric model continued, the fundamental assumption that all employment was concentrated in the center of the city was becoming increasingly less tenable. Indeed, Mills himself (1972) looked at the increasing decentralization of employment. The problems this posed for the monocentric model were also noted in the estimation of hedonic house price models. The monocentric model also predicts the negative exponential decline of land rents with distance from the center. This means that distance to the center should be significant in hedonic house price models, but numbers of studies found a lack of significance. Heikkila, *et al.* (1989) reviewed these studies, discussed the problem, and found that distances to multiple employment centers (except the CBD) were significant in a hedonic model for Los Angeles.

Directly addressing population density, Song (1994, 1996) found that standard measures of accessibility to all employment in an urban area predicted population densities better than distance to the center. Returning to hedonic house price models, Ottensmann (2008) found such measures performed better in that context than both distance to the center and distances to multiple centers. In a very different approach from these, Arribas-Bel and Sanz-Gracia (2014) considered the applicability of monocentric versus polycentric urban models in U.S. metropolitan areas from 1990 to 2010 using local indicators of spatial association (LISA). Their conclusions were mixed.

Data and Methods

This research uses a dataset that was developed with data on numbers of housing units in census tracts for large urban areas in the United States from 1950 to 2010. The tracts for urban portions of metropolitan areas were identified within the Combined Statistical Areas (CSAs) as delineated by the Office of Management and Budget for 2013 (U.S. Bureau of the Census 2013). CSAs were used rather than the more commonly employed Metropolitan Statistical Areas (MSAs) as it was felt they more properly represented the full extent of the metropolitan areas, including those instances in which 2 or 3 MSAs should more properly be considered to be parts of a single area. For those MSAs which were not incorporated into a CSA, the MSA was used.

The 59 CSAs and MSAs with 2010 populations over one million were selected for the creation of the dataset. A number of these areas had multiple large centers associated with separate urban areas than had grown together. This posed the issue of identifying those cases in which a second or third urban area could be considered sufficiently large in relation to the largest area to be considered as an additional center. The decision was made by comparing the population of census Urbanized Areas (either from the current census or the last census in which the areas were separate) with the largest area. A center was considered to be an additional center if its population were greater than 28 percent of the population of the largest area. The three areas included with the lowest percentages were Akron (with Cleveland), Tacoma (with Seattle), and Providence (with Boston). Sixteen of the areas had multiple centers, which made them unsuitable for estimating the monocentric negative exponential model, leaving the 43 areas with a single center to be used in this analysis.

The primary data source for this research was the Neighborhood Change Database developed by the Urban Institute and Geolytics (2003). This unique dataset provides census tract data from the 1970 through 2000 censuses, with the data for 1970 through 1990 normalized to the 2000 census tract boundaries. Population and housing unit data from the 2010 census were added by aggregating the counts from the 2010 census block data (U.S. Bureau of the Census 2012).

Housing unit densities—the numbers of housing units divided by the land areas of the tracts in square miles—are used in this research rather than the more commonly employed population density measure for two reasons. Housing units better represent the physical pattern of urban development as they are relatively fixed, while the population of an area can change without any changes in the stock of housing. Other studies of urban patterns have made similar arguments for choosing housing units over population, for example Galster, *et al.* (2001); Theobald (2001); Radeloff, Hammer, and Stewart (2005); and Paulsen (2014).

Using housing units also allows the extension of the analysis to census years prior to 1970. The census includes data on housing units classified by the year in which the structure was built, and these data are included in the Neighborhood Change Database. The 1970 year-built data can be used to estimate the numbers of housing units present in the census tracts for 1940, 1950, and 1960. Several prior studies have used the housing units by year-built data to make estimates for prior years in this manner, though they have used more recent census data to make the estimates, not the earlier 1970 census data (Radeloff, *et al.* 2001; Theobald 2001; Hammer, *et al.* 2004; Radeloff, Hammer, and Stewart 2005).

Sources of error in these housing unit estimates for earlier years from the year-built data arise from imperfect knowledge of the year in which the structure was built and from changes to the housing stock due to demolitions, subdivisions, and conversion to or from nonresidential uses. These errors increase for estimates farther back in time. Numbers of housing units for 1970 to 1990 were estimated from the 2000 year-built data and compared with the census counts in the Neighborhood Change Database. The judgment was made that estimates 2 decades back involved acceptable levels of error, but this was not the case for 3 decades back. As a result, the decision was made to use the housing unit estimates for 1950 and 1960 but not for 1940.

Urban areas were defined for each census year from 1950 to 2010 consisting of those contiguous tracts meeting a minimum housing unit density threshold. For the definition of Urbanized Areas for the 2000 and 2010 censuses, a minimum population density of 500 persons per square mile was required for a block or larger area to be added to an Urbanized Area (U.S. Bureau of the Census 2002, 2011). Using the ratio of population to housing units for the nation in 2000 of 2.34 persons per unit, a density of 500 persons per square mile is almost exactly equivalent to 1 housing unit per 3 acres or 213.33 units per square mile. This was used as the minimum urban density threshold.

Note that this is a measure of gross density, not lot size, as the areas of roads, nonresidential uses, and vacant land are included.

The location of the CBD must be specified to measure distances. One of the only efforts by the Census to do so came in a report for the 1983 economic censuses (U.S. Bureau of the Census 1983). This lists the census tracts comprising the CBD for many larger cities. This information was used to identify the CBD tracts for those urban areas included and for which the tract numbering and boundaries were the same for 2000. For the other urban centers, the tract or tracts for the CBD were identified by determining the location of the city hall or other major government buildings and examining the pattern of major roads, which generally converge on the CBD. The centroid of the CBD tract or tracts was taken as the center. Distances to the center were measured in miles to the centroids of each of the census tracts in the urban area.[1]

The monocentric model posits the negative exponential decline of density with distance from the center:

$$D_i = D_0 e^{-\beta s_i}$$

where D_i is the density in tract i, s_i is the distance from the center to tract i, D_0 is the density at the center, β is the density gradient, and e is the base of the natural logarithms. For estimation, this is transformed by taking the natural logarithm of both sides,

$$\ln(D_i) = \ln(D_0) - \beta s_i$$

producing a linear relationship that can be estimated using ordinary least squares regression. Log of density is regressed on distance to estimate the central density and density gradient parameters. This was done using the tracts within the urban areas for each census year for each of the 43 urban areas.

Negative Exponential Patterns Over Time

The parameters and goodness-of-fit of the negative exponential model predicting housing unit density by census tract were estimated for the 43 large urban areas with a single center out of the 59 CSAs or MSAs with populations over one million in 2010. Estimates were made for each census year from 1950 through 2010, including those tracts identified as part of the urban area for each year. This section reports the results of

---

[1] More detail on the construction of the dataset and the delineation of the urban areas is provided in Ottensmann (2014).

the estimates of the density gradient, the central density, and the coefficient of determination, R^2.

Table 1 summarizes the estimates of the housing density gradients, showing the mean value and minimum and maximum across the 43 areas for each year. The mean density gradient declined sharply and steadily, from 0.31 in 1950 to just 0.07 by 2010. This is consistent with findings by many others. The minimum and maximum density gradients also declined. The minimum value in 1950 was almost the same as the mean value in 2010. The maximum plummeted, from 0.77 to 0.24. All but one of the areas experienced a decline in the density gradient over the period. The exception was Birmingham, which had a small increase from 0.06 to 0.07.

Table 1. Negative Exponential Model Density Gradients, 1950-2010.

Year	Mean Density Gradient	Minimum Density Gradient	Maximum Density Gradient
1950	0.309	0.065	0.765
1960	0.226	0.061	0.610
1970	0.176	0.053	0.345
1980	0.116	0.027	0.291
1990	0.090	0.027	0.264
2000	0.074	0.020	0.268
2010	0.067	0.020	0.242

The consistency of the decline and the rapidly decreasing range of density gradient values is shown in Figure 1, which presents box plots for the density gradients for each of the years. It shows the decline accompanied by the convergence of the density gradients into ever-narrower ranges of values.

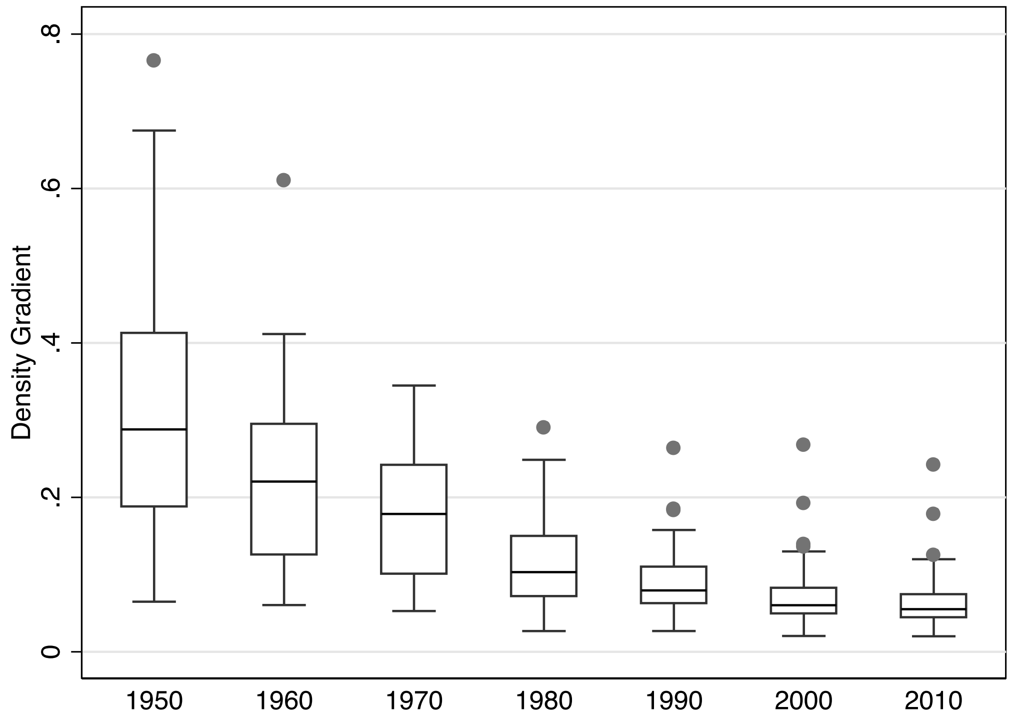

Figure 1. Box plots of Density Gradients, 1950-2010.

The mean central densities likewise declined steadily and significantly over the period from 1950 to 2010, from about 4,200 housing units per square mile to 2,900, as shown in Table 2. The maximum central density dropped from 18,100 to the 12,000 range during the first part of the period, remaining relatively steady after that. New York, of course, had the highest central density each year, followed by Chicago. The minimum values for the central density, on the other hand, showed only small variation, with no particular pattern over time. Smaller areas in the South consistently had the lowest central densities.

Table 2. Negative Exponential Model Central Densities, 1950-2010.

Year	Mean Central Density	Minimum Central Density	Maximum Central Density
1950	4,182	1,045	18,110
1960	3,958	1,266	11,753
1970	3,975	1,405	13,565
1980	3,360	1,256	12,730
1990	3,122	1,299	11,727
2000	2,956	1,227	12,584
2010	2,894	1,304	12,820

Unlike the density gradient, which declined for nearly all areas over the period, the pattern for the central densities showed far more variation. Eleven of the areas saw their central densities increase from 1950 to 2010, some by very substantial amounts. These areas and the estimated beginning and ending central densities and percentage change are listed in Table 3. These are areas that have grown rapidly and are located either in the South or especially the West. A previous analysis of these data examined the changes in the overall housing unit densities for these urban areas over the same period (Ottensmann 2015) and showed a similar pattern. The mean densities and maximum densities declined steadily over the period. But a substantial number of areas saw increases in overall density or relative stability. Of the 11 areas listed in Table 3 with increasing central densities, 9 also saw their overall housing unit densities increase over the period, while the two that did not (Austin and Birmingham) were essentially stable, with trivial declines in density of 9 and 29 units per square mile.

Table 3. Urban Areas with Increased Central Densities, 1950-2010.

Area	Central Density 1950	Central Density 2010	Percent Change 1950-2010
Las Vegas	1,687	3,217	90.7
Tucson	1,212	2,254	86.0
Austin	1,588	2,609	64.2
Albuquerque	1,382	2,240	62.1
Birmingham	1,045	1,557	49.0
Orlando	1,328	1,888	42.2
Sacramento	2,369	3,022	27.6
Portland	3,232	3,943	22.0
San Diego	2,670	3,005	12.6
Phoenix	1,947	2,169	11.4
Los Angeles	4,547	4,787	5.3

The R^2 values provide measures of how closely the distribution of census tract densities conforms to the negative exponential model. Table 4 presents the mean, minimum, and maximum R^2 values across the areas for each year. The mean value held fairly steady from 1950 through 1970, varying between 0.33 and 0.36. After 1970, the mean R^2 steadily declined, to a low of 0.19 in 2010, a drop of nearly 50 percent from the highest mean. The minimum and maximum R^2 values show some slight and inconsistent decline over time. The magnitudes of the changes for individual areas varied greatly, with some areas holding fairly steady. Nevertheless, only 6 areas did not experience some decline in their R^2 values, and in half of those cases, the increase was less than 0.05. So it seems that some areas continued to have density patterns more consistent with the negative exponential model, while for other areas the negative exponential model does not now predict the density pattern nearly as well as it did in the past.

Table 4. Negative Exponential Model R^2 Values, 1950-2010.

Year	Mean R^2	Minimum R^2	Maximum R^2
1950	0.334	0.072	0.619
1960	0.355	0.083	0.627
1970	0.358	0.111	0.568
1980	0.276	0.010	0.513
1990	0.226	0.027	0.454
2000	0.202	0.019	0.558
2010	0.190	0.023	0.504

Looking only at summary statistics limits the understanding of the nature of the changes. Considering a single area can provide further insight. Denver serves as an appropriate example because it is the most typical in terms of its pattern. In 1950, the density gradient, central density, and R^2 value were all at or within 3 places of the median values. In 2010, 2 of the 3 values for Denver were within 2 places of the medians (the estimated central density was somewhat higher). The density gradient for Denver fell from 0.284 to 0.057 from 1950 to 2010. Denver's central density declined only from 3,622 units per square mile to 3,134, but as noted, Denver was well above the median on this measure in 2010. Finally, the R^2 value decreased from 0.338 to 0.166.

Scatter plots of the census tract densities versus distance from the center for Denver in 1950 and 2010 illustrate the change These plots, in Figures 2 and 3, use the same axes to facilitate comparison and show the magnitude of the changes. The estimated negative exponential curve is also shown.The first thing that is striking about the comparison is the expansion of the Denver urban area. In 1950, the most distant tract was less than 10 miles from the center. By 2010, the urban area had expanded out to a distance of nearly 28 miles. The number of (year 2000 boundary) census tracts in the urban area increased from 124 to 506.

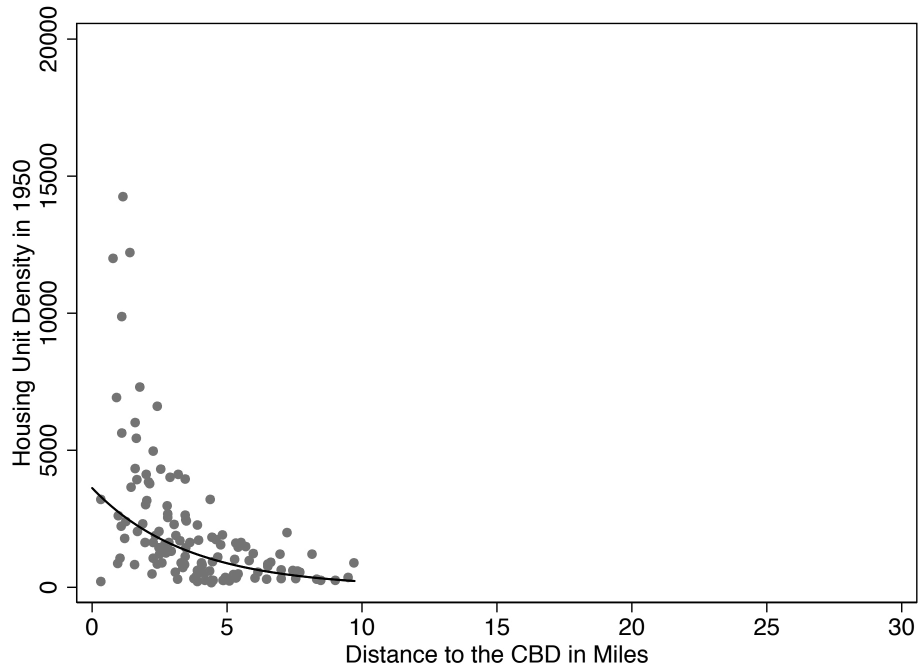

Figure 2. Negative Exponential Decline of Density, Denver, 1950.

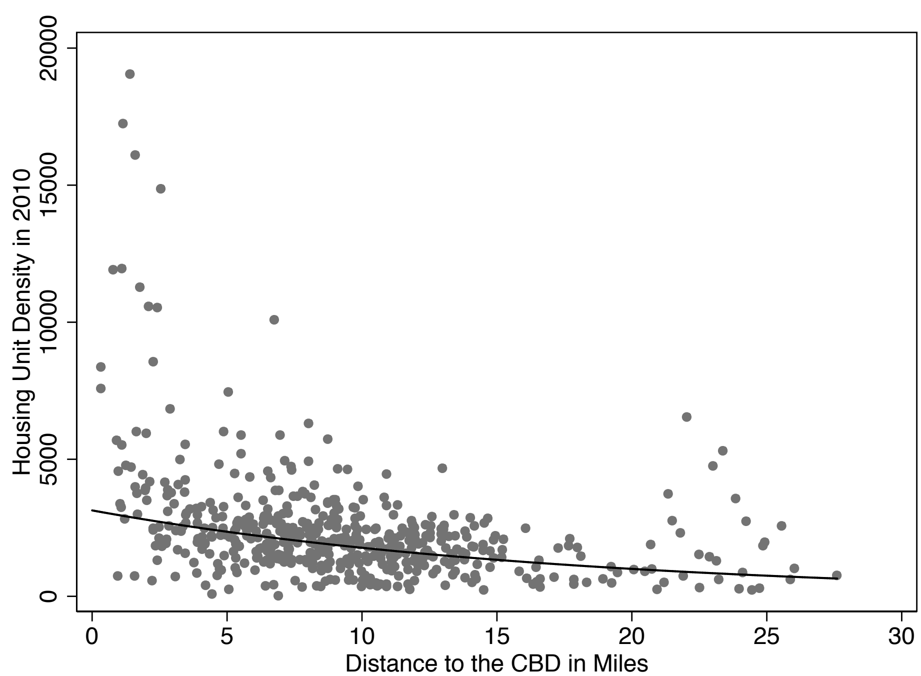

Figure 3. Negative Exponential Decline of Density, Denver, 2010.

In 1950, the pattern of the negative exponential decline in the densities is clearly evident. By 2010, the density gradient has become much flatter, and while there was still somewhat of a pattern of decline, the scatter of the densities around the line is clearly greater. Note in particular the higher densities out beyond the 20-mile distance.

If Denver is a typical example, what about are the areas with the highest and lowest R^2 values, the areas that conform most closely and the least to the negative exponential model? Table 5 lists the 8 areas with the highest R^2 values in 2010. It's a mix of very different areas, large and small, Northeast, Midwest, but a few in the South. These seem to be areas in which the CBD has remained quite dominant—New York and Chicago, of course with their extremely large CBDs. But this might also explain the presence of Atlanta and Charlotte on the list of areas conforming most closely to the negative exponential pattern predicted by the monocentric model. These areas all had similar R^2 values in 1950, so their monocentric pattern has remained reasonably constant.

Table 5. Areas with the Highest R^2 Values in 2010

Area	R^2
Grand Rapids	0.504
Rochester	0.472
Chicago	0.424
New York	0.366
Louisville	0.354
Atlanta	0.318
Milwaukee	0.297
Charlotte	0.278

Table 6 shows the 8 areas with the lowest R^2 values in 2010. All of the areas are within or near the Southwest. These are predominantly areas that have experienced much of their growth relatively recently. Most of these areas had high R^2 values in 1950, with all but one exceeding 0.30 in that year. So they started the period as small areas oriented around a small CBD. But their subsequent growth produced very different distributions of densities by 2010. This regional pattern is not limited to these 8 areas with the least monocentric pattern. The 8 areas with the next lowest values are, with only one exception, in the South or the West.

Table 6. Areas with the Lowest R^2 Values in 2010

Area	R^2
Oklahoma City	0.023
Phoenix	0.027
Kansas City	0.031
El Paso	0.033
Las Vegas	0.047
San Antonio	0.064
Albuquerque	0.078
Fresno	0.083

Conclusions

From 1950 to 2010, the negative exponential density gradient declined, often very dramatically, for nearly all of the large urban areas, similar to what has been reported by others. This decline is often attributed to a combination of declining transportation costs and increasing real incomes, causing households to place higher value on having more space relative to transportation costs. The mean central densities estimated for all of the areas likewise showed the expected, large decline over time. In this case, however, the decline did not occur uniformly across all areas. Eleven areas in the West and South saw increases in central densities over the period. These were areas that experienced higher rates of growth. Apparently the increased demand in these areas overcame the decentralizing tendencies.

The mean goodness-of-fit of the negative exponential model, measured by R^2, remained relatively constant from 1950 through 1970. After that time, mean R^2 declined steadily to a value just over half the largest value by 2010. Some areas retained higher values, likely areas in which the CBD retained greater dominance. The areas with the smallest, extremely low goodness-of-fit values were smaller, newer urban areas in or near the Southwest.

The decreasing fit of the negative exponential model is consistent with the shift to more polycentric urban areas, with multiple centers of employment. But despite extremely low R^2 values for some areas, the negative exponential model still seems to "work," in the sense that the estimates of the density gradients and central densities are reasonable. The density gradients for these low-fit areas have mostly declined over time and while they are not large, they are similar to the gradients for areas for which the fit of the model is much better. Likewise, estimated central densities for the low-fit areas fall in the same range as the densities for areas with much higher R^2 values.

Without considering the goodness-of-fit, the monocentric model and its prediction of the negative exponential decline of density with distance from the center appears to continue to serve for describing the pattern of settlement in urban areas. And certainly people continue to use the model for this purpose. But the fit of the model has declined dramatically for many areas over that past 4 decades, to levels, accounting for as little as 2 or 3 percent of the variation in density in some areas. This is entirely consistent with the development of polycentric urban areas with multiple employment centers. This raises the question as to how it can be true that the monocentric model continues to work as areas become more polycentric.

References

Anas, Alex, Richard Arnott, and Kenneth A. Small. 1998. Urban spatial structure. *Journal of Economic Literature* 36, 1 (September): 1426-1464.

Arribas-Bel, Daniel, and Fernando Sanz-Gracia. 2014. The validity of the monocentric city model in a polycentric age: US metropolitan areas in 1990, 2000 and 2010. *Urban Geography* 35, 7: 980-997.

Angel, Shlomo. 2012. *Planet of Cities*. Cambridge, MA: Lincoln Institute of Land Policy.

Clark, Colin. 1951. Urban population densities. *Journal of the Royal Statistical Society. Series A (General)* 114, 4: 490-496.

Galster, George, Royce Hanson, Michael R. Ratcliff, Harold Wolman, Stephen Coleman, and Jason Freihage. 2001. Wrestling sprawl to the ground: Defining and measuring an elusive concept. *Housing Policy Debate* 12, 4: 681-717.

Guest, Avery M. 1975. Population suburbanization in American metropolitan areas, 1940-1970. *Geographical Analysis* 7, 3 (July): 267-283.

Hammer, Roger B., Susan I. Stewart, Richelle L. Winkler, Volker C. Radeloff, and Paul R. Voss. 2004. Characterizing dynamic spatial and temporal residential density patterns from 1940-1990 across the North Central United States. *Landscape and Urban Planning* 69, 2-3: 183-199.

Heikkila, E., P. Gordon, J. I. Kim, R. B. Peiser, H. W. Richardson, and D. Dale-Johnson. 1989. Whatever happened to the CBD-distance gradient?: Land values in a polycentric city. *Environment and Planning A* 21, 2 (February): 221-232.

Mills, Edwin S. 1972. *Studies in the Structure of the Urban Economy.* Baltimore: Johns Hopkins Press for Resources for the Future.

Muth, Richard F. 1969. *Cities and Housing: The Spatial Pattern of Urban Residential Land Use*. Chicago: University of Chicago Press.

Ottensmann, John R. 2014. Urban patterns dataset description. Available at https://urbanpatternsblog.files.wordpress.com/2016/12/urban-patterns-dataset-description.pdf.

Ottensmann, John R. 2015. Density of large urban areas in the U.S., 1950-2010. Available at https://urbanpatternsblog.files.wordpress.com/2016/12/density-of-large-urban-areas.pdf.

Paulsen, K. 2014. Geography, policy or market? New evidence on the measurement and causes of sprawl (and infill) in US metropolitan regions. *Urban Studies* 41, 12 (September): 2629-2645.

Radeloff, Volker C., Roger B. Hammer, Paul R. Voss, Alice E. Hagen, Donald R. Field, and David J. Mladenoff. 2001. Human demographic trends and landscape level forest management in the northwest Wisconsin Pine Barrens. *Forest Science* 47, 2 (May): 229-241.

Radeloff, Volker C., Roger B. Hammer, and Susan L. Stewart. 2005. Rural and suburban sprawl in the U.S. midwest from 1940 to 2000 and its relation to forest fragmentation. *Conservation Biology* 19, 3 (June): 793-805.

Song, Shunfeng. 1994. Modelling worker residence distribution in the Los Angeles region. *Urban Studies* 31, 9 (November): 1533- 1544.

Song, Shunfeng. 1996. Some tests of alternative accessibility measures: A population density approach. *Land Economics* 72, 4 (November): 474-482.

Theobald, David M. 2001. Land-use dynamics beyond the urban fringe. *Geographical Review* 91, 3 (July): 544-564.

U.S. Bureau of the Census. 1983. *1982 Economic Censuses. Reference Series. Geographic Reference Manual.* EC82-R-1. Washington, DC: U.S. Government Printing Office, Table 9, pp. 309-322.

U.S. Bureau of the Census. 2002. Urban area criteria for the 2000 Census. *Federal Register* 67, 51 (Friday, March 15): 11663-11670.

U.S. Bureau of the Census. 2011. Urban area criteria for the 2010 Census. *Federal Register* 76, 164 (Wednesday, August 24): 53030-53043.

U.S. Bureau of the Census. 2012. Census block shapefiles with 2010 Census population and housing unit counts. Downloaded from http://www.census.gov/geo/www/tiger/tgrshp2010/pophu.html on September 29, 2012.

U.S. Bureau of the Census. 2013. Core based statistical areas (CBSAs) and combined statistical areas (CSAs) 2013 delineation file. February 2013. Downloaded from https://www.census.gov/population/metro/data/def.html on February 13, 2014.

Urban Institute and Geolytics. 2003. *Census CD Neighborhood Change Database (NCDB): 1970-2000 Tract Data*. Somerville, NJ: Geolytics, Inc.