Exploring Urban Density

John R. Ottensmann
Indiana University-Purdue University Indianapolis
john.ottensmann@gmail.com
urbanpatternsblog.wordpress.com
July 2021

Abstract

Density in urban areas is not a simple concept. It can be defined in a variety of ways. This paper presents alternative measures of urban density and compares how they perform across a set of 59 large urban areas in the United States in 2010. Population and housing unit densities show similar patterns, though with differences for some areas with extreme population-housing unit ratios. The housing unit densities do perform better in the estimation of the negative exponential model of density decline. The conventional measure of total population divided by the total land area of an urban area differs from various measures of the density of residential areas. The absence of data precludes the calculation of residential densities for large numbers of urban areas, but using block data and excluding blocks that are completely or largely nonresidential makes possible alternative measures coming closer to residential density. When used to estimate the negative exponential model, these measures produce higher estimates of the parameters and a significantly greater goodness-of-fit. Weighted density calculated as the mean of the densities of small areas weighted by their population is another, completely different measure. These are higher than convention densities, with the differences being larger for those urban areas with greater internal variation in density, such as areas with high density cores and low density peripheries.

Introduction

Louis Wirth, in his seminal essay "Urbanism as a Way of Life," (1938) identified density as one of three characteristics that mark an area as urban (population size and heterogeneity were the others). The U.S. Bureau of the Census (2011) defines as urban those areas in which the population density exceeds minimum thresholds.

The importance of urban density is demonstrated by the wide range of studies examining its relationship to other characteristics of urban areas (Boyko and Cooper 2011). A few examples (the number of studies that could be cited is huge): The density of urban areas has been shown to be related to motor vehicle use and the attendant energy consumption and greenhouse gas emissions (Brownstone and Golob 2009; Lee and Lee 2014). The cost of providing public services was associated with urban density

(Ladd 1992). Economists have shown urban density to be related to the level of innovation as measured by numbers of patents issued (Carlino, Chatterjee, and Hunt 2006; Knudsen, *et al.* 2008). A great deal of attention has been given to the relationship between density and measures of health, most recently to the spread COVID-19 (Baser 2020; Carozzi, Provenzano, and Roth 2020; Garland, *et al.* 2020). Survey research has shown perceived quality of life to be related to density (Cramer, Torgersen, and Kringlen 2004; Walton, Murray, and Thomas 2008).

Urban sprawl represents a prominent example of the importance of urban density. Low-density development is nearly always cited as one aspect of sprawl, and scattered, leapfrog development contributes to low densities when looked at over larger areas (Ewing 1997). A wide range of negative effects have been attributed to sprawl (though not all agree, e.g., Gordon and Richardson 1997). Some of these costs of sprawl include vehicle miles traveled, energy consumption, and air pollution; higher infrastructure and public service costs; loss of resource lands; impacts on central cities and downtowns; and psychic and social costs (Ewing 1997). An early study of these effects was *The Costs of Sprawl* (Real Estate Research Corporation 1974), but see the critique by Windsor (1979). More recent works revisiting this topic are Burchell (1997) and Burchell, *et al.* (1998, 2002).

Urban area density is a seemingly simple concept, the population of the urban area divided by its area—a ratio. But various choices are possible for the values used in both the numerator and denominator. Data other than the totals may be used with measures of density being calculated in a variety of ways.

This paper explores a variety of measures of density for urban areas. The next section discusses different ways in which density has been conceptualized and measured. This is followed by the description of the specific measures of urban density being considered in this paper and the 59 large urban areas to which these measures are applied. The results are presented in two sections. The first examines overall measures of the density of the large urban areas, comparing population and housing unit density, densities using different measures of area, and weighted densities. The use of different measures to examine the variation in density within the urban areas follows. Brief consideration is given to the differences in density between the older cores and newer peripheries of the urban areas. More extensive treatment is given to the effect of using alternative measures of density for the estimation of the pattern of density decline with distance from the center.

Conceptions of Density

In the most general sense, density can be seen as the relationship between the count of some entities and the area in which those are located. Most often when thinking about urban area density, the focus is on the number of people residing in the area. The density of many other types of phenomena within urban areas can also be of interest (Boyko and Cooper 2011). Examples include daytime population density (Boeing 2018), employment density (Glaeser and Kahn 2004), road network density (Borruso 2003), trip density (Kaddoura and Schlenther 2021), and air pollution density (Sifakis 1998). Hawley (1972) made a distinction between physical density such as population per land area and social density, the number of interactions or messages exchanged per unit of time (and this before the time of email, text messages, and the other forms of electronic messaging taking place today). He also stressed the relationship between the two types of density. The volume of communications by electronic means within an area could be the basis for another type of density. Rapoport (1975) goes further in discussing perceptions of both physical and social density.

This paper considers the density of the population living in urban areas and the closely related but distinct density of housing units. The most common relationship used to express urban area density is the ratio of the count, such as population, to the area of the urban area, such as population per square mile. Of course area could be expressed in other units such as kilometer, acres, or hectares. But this difference is not important as areal units can readily be converted. This paper will use population and housing units per square mile.

Since the most basic concept of urban area population density is the total population divided by the total area of the urban area, how the area is defined plays a critical role. It is important to consider an area that corresponds to the actual extent of urban development. Overbounded areas that include significant rural territory obviously increase the area and, since they add only small numbers of people, reduce the density. This is a particular problem with Metropolitan Statistical Areas (MSAs), which are constructed using counties as building blocks. Counties can have significant commuting to the core of an MSA and be added when only a small portion of the county is actually urban. The problem is further exacerbated for counties with large areas, which can have more territory that are not urban.

Other approaches consider the density of population or housing units within those portions of the urban area that are considered to be residential. These measures of density vary in the extent of the area included in the calculation (Alexander 1993; Churchman 1999). The most restrictive measure uses only the area of the parcels of land with residential uses—lots for single-family houses and the parcels on which multi-family structures are located. Since this includes the smallest quantity of land, it produces the highest measure of density. Next comes the measure of residential density that also includes the areas of streets. This could be the entire area of a single-family subdivision, including the streets and any common areas. By convention, such a measure includes one-half the area of streets on the periphery of the residential area. Most generous with respect to the land area included are measures that include the area of broader residential neighborhoods including the area of other uses typically found in such neighborhoods such as schools, places of worship, neighborhood parks, and perhaps neighborhood-serving commercial. With greater area included, such measures produce lower levels of density. Of course all of these densities are greater than the overall density of the urban area as measured by population or housing units divided by the total area, which includes nonresidential as well as residential areas.

A few comments on these measures of density. Residential density is more often measured and expressed as housing units per acre (or hectare) (Ellis 2004), while the overall density for entire urban areas is more often given in terms of population per square mile (or square kilometer). But of course either population or housing units and any measure of area can be used for any of these measures of density. In comparing these different measures of density, the measure using the more restricted definition of area is often referred to as *net* density and the less restrictive density measure as *gross* density. When comparing the residential density using only the area of residential parcels with that using the total land area of the urban area, the former would be referred to as net density and the latter as gross density. Unfortunately the intermediate measures have been referred to by different sources as both net and gross density, depending on the density to which is was being compared (Alexander 1993). Therefore, the net and gross density distinction is not very helpful.

The measures of urban area density considered thus far have considered a total quantity such as the number of people or housing units in the urban area divided by some measure of area—either the total land area or the area of residential land however defined. But if one begins with the densities for small areas such as census tracts within the urban area, the possibility arises for creating measures of density in different ways. Craig (1984, 1985) described various approaches for calculating population-weighted densities from such subarea densities. Most common is the population-weighted mean of the densities of the tracts or other subareas. Population-weighted density is the density experienced by the average person. Craig (1984) described this as a better measure "at which the population lives." Ottensmann (2018) is an extensive discussion of population-weighted density.

When one is considering the densities of subareas of the urban area, it becomes natural to expand the discussion to the consideration of the variation in densities within the urban area. Simple comparisons can be made such as between densities in the central cities and the suburban portions of the urban areas. The greatest attention has been given to the variation of densities with distance from the center, the central business district (CBD). Clark (1951) observed that population densities have tended to follow a regular pattern of negative exponential declines with distance for cities in different part of the world and at different times. The standard monocentric model of urban economics predicts such decline, providing one explanation for the pattern (Muth 1969; Mills 1972). Many studies of the density gradients estimated for the negative exponential decline have followed, with reviews of that literature from very different perspectives provided by Thrall (1988), McDonald (1989), and Smith (1997).

Measures of Density

This research compares numbers of different measures of density in 2010 for 59 large urban areas in the United States. This section describes the various measures and how their values are calculated.

As mentioned earlier, how the urban area is defined is very important, especially for the basic measures of total count divided by total land area. The area included should approximate the extent of the area that is actually developed and urban in character and should not include extensive rural areas like many Metropolitan Statistical Areas. To accomplish this, the urban area should be delineated using small areas, which can then be included (or not) based on their meeting some criterion of urban character, generally having a density exceeding some threshold. The Census (U.S. Bureau of the Census 2011) defined the areas of the country that are urban and Urbanized Areas using census tracts and blocks with population densities exceeding a minimum value (1000 persons per square mile for the initial, very small urban core, and 500 persons per square mile for tracts and blocks to be added to this). To allow certain questions to be addressed, some researchers have chosen to define urban areas using block groups (Theobald 2012) or census tracts, which I have used. These are the urban areas used here. Details on their delineation are provided in the following section.

Both population and housing unit densities are being considered. (The population densities are the residential population—people residing in an area—as reported by the Census.) While the population and housing unit densities obviously area closely related, occasionally interesting differences can arise. The source of the basic data being used for the calculation of all of the density measures is the block data from the 2010 census, which includes counts of both population and housing units (U.S. Bureau of the Census). Blocks are the smallest units for which the Census reports data. There were over 11 millions blocks in the United States for the 2010 census (U.S. Bureau of the Census 2021a). Within cities, the census blocks are essentially the same as what is generally meant by a block—a small area bounded by streets on each side. Farther from the center census blocks can be larger but are generally the smallest areas that can be delineated using features visible on the ground—roads, railroads, water, and so on.

The first measures of density are the conventional population and housing unit density for the urban area. These are the total population and the total number of housing units divided by the total land area of the urban area, also as reported by the Census.

The various measures of residential density require the area of the territory considered to be residential. The Census does not include information on the type of land use. No other consistent national source of data on land use for small areas is available either. (The National Land Cover Database provides some information on land cover (not exactly the same thing as land use) for small pixels including the general intensity of development in 3 classes (U.S. Geological Survey 2021). However, for calculating measures of residential density, this data is not sufficiently specific or accurate and fails to include areas that are residential but classified as having other land cover (such as yards for large-lot single-family housing).

Use of the census block data provides an opportunity to devise measures of density that move towards measures of residential density and 2 of these are considered here. The first measure excludes the areas of blocks with no population or no housing units and uses the total land area of the nonzero blocks for calculating the density. Obviously the blocks without population or housing units (depending on the density type) are totally nonresidential and include none of the area that would be considered to be residential. This will produce measures of density higher than obtained using the total area of the residential area. This is far from being a measure of residential density, however, as blocks containing just a small number of people or housing units would not be considered to be residential. Rae used an analogous approach in calculating the densities of countries in Europe using data for 1 square kilometer grid cells. Densities were calculated using only the areas of the cells with nonzero populations. This was referred to as "lived density."

Given the limitations of using the nonzero-density blocks, an attempt is made to get closer the area of residential use by excluding additional blocks with very limited presence of population or housing units. The cutoff is the minimum density threshold used to identify areas as urban. Blocks having a density above the minimum of 500 persons per square mile or the roughly equivalent 213 housing units per square mile are included in the area for the calculation of density. Blocks having densities below these levels are excluded as being primarily nonresidential. Of course this means that some population and housing units are included in blocks with densities under the minimum, but the amounts are surpassingly small, as will be shown. Only the total of the population and housing units in the included blocks exceeding the density threshold are used in the calculation of the density measure, divided by the area of those blocks.

Weighted population and housing unit densities constitute the next set of measures considered. The formula for population-weighted density D_W is

$$D_W = \frac{1}{P} \sum p_i d_i$$

where the sum is over all of the subareas within the urban area, p_i is the population of subarea i, d_i is the density of the subarea (population divided by area), and P is the total population. The formula for housing unit-weighted density is, of course, analogous.

In the United States, census tracts have been the subareas most often (but not always) used for the calculation of population-weighted density, so the population and housing unit weighted densities are calculated using census tract populations, housing units, and areas. The values of weighted densities depend on the subareas used. Smaller, more numerous subareas will produce higher densities as the variation across the smaller subareas is greater and the weighted density increases with the variance (Craig 1984, 1984; Morton 2015; Ottensmann 2018). To examine this effect, the weighted densities are also calculated using the blocks as the subareas.

In a study of the incidence of COVID-19 at the country level, Garland, *et al.* (2021) used population-weighted density calculated using data for 30-arc-second grid cells (approximately 1 kilometer square at the equator). They compared the relationships of measures of COVID-19 with conventional density and the population density for grid cells with nonzero population using the data from Rae (2018) to illustrate the differences using alternative measures of density.

The variation in density within the urban areas is also considered. The first step is the basic comparison of densities between the urban cores and the suburban peripheries of the urban areas. The definitions of these areas and the rationale for their use is described in the following section. Both the conventional densities and the densities using the areas of blocks with nonzero populations and housing units are considered.

More detailed examination is given to the negative exponential decline of density with distance from the center. This model can be expressed as

$$d_i = D_0 e^{\beta s_i}$$

where d_i is the density in subarea i and s_i is the distance from that subarea to the center. The density gradient β and the central density D_0 are parameters to be estimated in fitting the model, and e is the base of the natural logarithms. When estimating the parameters using data for subareas, the method nearly always used is to take the natural logarithm of both sides of the equation,

$$ln\left(d_i\right) = ln\left(D_0\right) - \beta s_i$$

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.

When estimating the negative exponential model using subarea data, the densities and distances for census tracts are nearly always used. The first set of negative exponential models estimated use tract data. The negative exponential model is estimated using conventional tract densities, densities for tracts using the areas of blocks with nonzero populations and housing units, and the tract densities using populations and housing units and areas for blocks in which the block density exceeds the minimum urban density threshold.

Increasing attention has been devoted during the past several decade to what has been termed the modifiable areal unit problem (MAUP). This refers to the issue that when using data that has been aggregated to areas, the choice of the areas used will affect the results of any analyses, sometimes producing dramatic differences. Openshaw (1984) and Wong (2009) provided 2 good summaries of the MAUP at different point in time. Two aspects of this problem result from the different ways in which an area can be subdivided into a fixed number of subareas, the zoning problem, and differences in the geographic resolution and number of subareas, the scale problem, following the terminology used by Wong. Studies have shown that increasing the number of subareas will generally result in a decrease in the correlations between variables and a decline in the estimates of parameters in multivariable models (see also Fotheringham and Wong 1991).

The extent of the MAUP for the estimation of the negative exponential model of density decline with distance is examined. Along with the estimates using the census tract data, the model is estimated using data at the block level. Models use the densities of blocks with nonzero population and the densities of blocks having a density exceeding the minimum urban density value. Note that no block-level model is estimated for conventional block densities. Estimation of the transformed model requires taking the logarithm of density. The log of zero is undefined, so blocks with zero density would be missing and excluded, producing the equivalent of the model using blocks with nonzero populations.

Research Context

This study is part of a larger research project looking at patterns of development in large urban areas in the United States from 1950 to 2010. These areas, the urban areas delineated for each census year, are the areas within which the measures of density are calculated and analysis is conducted. This section starts by identifying the urban areas included in the dataset, describes the housing unit data for census tracts that is the core of the data, and discusses how these data have been used to define urban areas from 1950 to 2010. The identification of the locations of the Central Business Districts (CBDs) is discussed. The delineation of the urban core and suburban periphery is the final topic addressed.

This research uses a dataset for the analysis of urban patterns over time 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 better represented the full extent of the metropolitan areas, including those instances in which 2 or more MSAs should more properly be considered to be parts of a single area (Ottensmann 2017). 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 that 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 included as an additional center around which urban development occurred. The decision was made by comparing the populations of census Urbanized Areas (either from the current census or the last census in which the areas were separate) with the largest area. An area 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). Areas with multiple centers have each center included in the name of the urban area.

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 units and 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 and population density measures 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 provides 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 conversions 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 have been defined for the urban patterns research for each census year since 1950 consisting of those tracts contiguous to an urban center meeting a minimum housing unit density threshold. (This is comparable to the way in which the census defines Urbanized Areas using blocks and larger units and Paulsen (2012) defined urban areas using block groups.) 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 an area to be included (U.S. Bureau of the Census 2002, 2011). Using the ratio of population to housing units for the nation as a whole in both 2000 and 2010 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 conventional and not residential density as the total area includes roads, nonresidential uses, and vacant land.

To provide for a set of urban areas that represents the cumulative expansion of the urban areas over time, a further condition was imposed that if a census tract did not exceed the minimum housing unit density and had not been included in the urban area in any given year, it would not be included in urban areas delineated in earlier years even if the density exceeded the minimum. The rule has been imposed in this direction—if rural, then not urban earlier—rather than in the opposite direction—if urban, then urban later—because the more recent data are considered to be generally more accurate.[1]

The location of the CBD must be specified to measure distances for the estimation of the negative exponential models. One of the only efforts by the Census to

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

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 from the centroids of each of the census tracts and blocks in the urban area.

This final part describes the definition of the urban core and suburban periphery used for the initial basic analysis of the distribution of densities within the urban areas. Central cities have been defined for Metropolitan Areas (MSAs) and their predecessors since the middle of the last century. They have frequently been used to identify the older urban core at the center of MSAs. The remainder of an MSA was then seen as encompassing the more recently developed suburban periphery that has generally been referred to as the "suburbs."

The problem with using this central city-suburb delineation to distinguish the older and newer portions of metropolitan areas is that central cities encompass widely varying proportions of metropolitan area populations and older and newer territory. Some of the older central cities have had their boundaries fixed before the end of the nineteenth century. Many of the older central cities had been more-or-less completely developed by the early decades of the twentieth century. Some of the inner suburbs of these metropolitan areas likewise date back to the nineteenth and early twentieth centuries.

Other central cities have been able to continue to expand their boundaries through annexation as their metropolitan areas have grown through the twentieth century. They have continued to see extensive new development within their boundaries, in some instances literally up to the present. As a result, some central cities can have large areas of new development that could be described as being suburban in character, virtually indistinguishable from development in the suburban areas outside of the central cities of other metropolitan areas. This means that the traditional central city-suburb distinction becomes a quite crude and fraught way of distinguishing the older, inner cores of metropolitan areas from the newer suburban areas at the periphery.

As a result of these problems and others affecting comparisons over time, an alternative approach has been taken for the identification of the older and newer portions of urban areas that is consistent across the 59 urban areas. These subareas are denoted as the urban core and suburban periphery. Using the delineations of the urban areas for each census year developed for the urban patterns dataset, the urban core is defined as the urban area in 1950. This is a fixed area, constituting the urban core for each successive census year. It is an area defined in a consistent manner for all of the urban areas and is an area that does not change over time. The suburban periphery is then the portion of the urban area in each succeeding census year that is outside of the urban core. The urban periphery expands over time as the urban area as a whole expands. But like the urban core, the urban periphery is being defined in a consistent manner for all urban areas.

No doubt some will complain that this is an overly generous specification of the urban core. In the Northeast and Midwest, the central cities often encompass only much older portions of the metropolitan area. But this is emphatically not true for many central cities in the South and West. On average, the urban cores actually included a smaller proportion of the urban area populations than the central cities in 2000 (Ottensmann 2019).

Overall Urban Area Density

This section presents results for the measures of the overall density in 2010 for the 59 large urban areas using the different measures of density. It begins with the comparison of population and housing unit density. Next comes the examination of the density measures excluding some of the nonresidential portions of the urban areas. The final part presents the measures of weighted population density.

Population versus Housing Unit Density

The population of an urban area and the number of housing units will naturally be closely related. The ratio of the population to the number of housing units directly affects the densities. Population per housing unit varies across urban areas. Of course differences in household sizes, the numbers of persons occupying housing units, create differences. Occupancy rates are another factor, as vacant units are included in the count of the number of housing units. And in addition to the population residing in housing units, the Census also identifies the population living in group quarters such as college dormitories, nursing homes, and jails. Thus one can have housing units without people and people without housing units.

Table 1 presents summary statistics across the urban areas for conventional population and housing unit density, the total count divided by the total land area of the urban areas. Of course population densities were larger, with a mean of 2,500 persons per square mile, ranging from under 1,300 to a maximum of about 5,600. Mean housing unit density was slightly under 1,100 units per square mile, going from well under 600 to over 1,900. The distributions were skewed in a positive direction, with smaller numbers of areas with the highest densities. For both types of density, the majority of the areas had densities in the lower part of the range, with the distance from

Table 1. Summary Statistics for Conventional Population and Housing Unit Density.

	Conventional Population Density	Conventional Housing Unit Density
Mean	2,543	1,070
Minimum	1,253	548
First Quartile	1,907	844
Median	2,211	1,056
Third Quartile	2,966	1,212
Maximum	5,619	1,931

the minimum to the first quartile being much less than the distance from the third quartile to the maximum.

Differences in the population per housing unit among the urban areas are reflected in the relative differences in densities. The mean population per housing unit for all areas was 2.36, essentially the value for the nation as a whole. However the range in the ratio was surprisingly large, from a minimum of 2.04 persons per housing unit to a maximum of 2.94. Four of the urban areas were clear outliers on the high end, having values of 2.85 and above, with the next highest at 2.64. The area with the highest population per housing unit was the Salt Lake City-Ogden-Provo area. This could be due to larger families among the dominant Mormon population (Lipka 2015). The next 3 outliers were Los Angeles, El Paso, and Fresno. These are areas having large Latino populations including significant numbers of recent immigrants. This would make the overall population skew younger, with more children. Also, lower incomes might increase the practice of families doubling up or otherwise including more persons in the households. The presence of these outliers had a clear effect on the relationships of population and housing unit densities across the urban area. The correlation between population and housing unit density for all 59 areas was 0.96, very high, but increased to 0.98, close to perfect when the four outliers were excluded.

The areas with the lowest populations per housing unit were not as extreme. They were Tampa-St. Petersburg, New Orleans, Pittsburgh, and Dayton. Tampa-St. Petersburg likely had a high proportion of retirees living in 1 or 2-person households. Pittsburgh and Dayton are older areas that have experienced population loss in recent decades, and outmigration of the younger population may have left the population older, living in smaller households as well. New Orleans must be considered to be a special case due to the effects of Hurricane Katrina just 5 years before the 2010 census.

Of course there was major population loss. But in addition, very large numbers of vacant, damaged housing units likely remained with those numbers reducing the population-housing unit ratio.

To provide further context and to serve as a basis for comparison with results that follow, the 6 urban areas (approximately 10 percent) with highest and lowest conventional population densities are shown in Table 2. The list of the areas with the highest population densities may be surprising to some. Los Angeles was the densest urban area with a population density of 5,600 persons per square mile, followed by San Francisco-Oakland-San Jose. New York, often considered to represent the epitome of high density, was third with a considerably lower density of 4,600. This is not an aberration resulting from the definition of the New York urban area used here (including, among other areas, the Connecticut suburbs) being more expansive than the Census definition of the Urbanized Area. (The definitions of the Los Angeles and San Francisco-Oakland-San Jose areas were also more generous, by the way). The Los Angeles Urbanized Area also the highest population density among all Urbanized Areas in 2010, again followed by San-Francisco-Oakland. The New York area actually dropped to fourth place on that list, having been edged out by the Las Vegas Urbanized Area with a slightly higher density.

Not only is the New York urban area not at the top of the list, but the remainder of the list does not include any of the older large urban areas that might be associated with high densities such as Chicago, Philadelphia, or Boston-Providence. In their place are the newer, faster-growing areas of San Diego, Las Vegas, and Miami-Fort Lauderdale-West Palm Beach. The list of the urban areas with the lowest conventional population densities is made up exclusively of somewhat smaller urban areas in the South.

The lists of the areas with the highest and lowest housing unit densities are very similar. New Orleans, with its very low population per housing unit, made the list in fourth place, up from eighth in terms of population density, with San Diego dropping to seventh. Among the areas with the lowest housing unit densities, Harrisburg-York was added in sixth place, up from seventh in population density, with Birmingham dropping several places, out of the top 6.

The general conclusion is that population and housing unit densities present very similar pictures of urban area density. Some variation exists for areas with extreme population-housing unit ratios, but these are clearly exceptional cases. Since population density is most often used in describing the density of urban areas, the results in the rest of this paper will focus largely on population measures. Results involving housing unit density are considered only where the differences are significant.

Table 2. Urban Areas with the Highest and Lowest Conventional Population Density.

Area	Conventional Population Density
Highest Density Areas	
Los Angeles	5,619
San Francisco-Oakland-San Jose	4,824
New York	4,609
San Diego	3,976
Las Vegas	3,920
Miami-Fort Lauderdale-West Palm Beach	3,552
Lowest Density Areas	
Greenville-Spartanburg	1,253
Knoxville	1,387
Raleigh-Durham	1,440
Greensboro--Winston-Salem--High Point	1,473
Charlotte	1,515
Birmingham	1,587

Density Using Different Areas

In addition to the conventional density of the entire urban area, interest has been directed at the density of the residential portions of the area, however defined. No data are available that allow calculation of measures of residential densities. However, block data can be used to exclude parts of the urban area that are completely or mainly nonresidential, yielding density measures that come closer to residential density.

Two alternative measures of density are considered here. The first excludes those blocks having zero population, which are clearly nonresidential. Density is calculated by taking the total population of the urban area (all of which remains) and dividing it by the area of the blocks having nonzero populations. Not all of this area can be considered to be residential, but the resulting, higher measure of density comes closer to being a measure of residential density.

Given that a block having only a small number of residents likely would not be considered to be residential, the second approach excludes some of those blocks as well.

Blocks are excluded having a population density of less than 500 persons per square mile, the minimum for the Census when designating areas as urban. The exclusion of blocks below the density threshold but having nonzero populations will exclude persons living in those blocks. Density is then calculated for blocks having a density exceeding the urban density threshold, the total of the population in these blocks (somewhat less than the total population of the urban area) divided by the total area of the blocks.

Summary statistics for these 2 measures of population density and conventional density, for comparison, are shown in Table 3. Mean density using only the area of blocks with nonzero population was nearly 3,100 persons per square mile, substantially higher than the mean for conventional density of 2,500. But the mean density calculated using the population and area of urban density blocks was much higher, 4,100 persons per square mile.

The mean increase from conventional to nonzero block density was 19 percent. While it ranged from 9 to 52 percent, the middle half of the areas were clustered in the narrow range from 14 to 25 percent. The mean increase from conventional to urban density block density was 63 percent. ranging from 41 to 99 percent. The minimum increase for urban block density was greater than the increase in nonzero block density for all but one of the urban areas, New Orleans, with the maximum increase of 52 percent.

The mean percentage of the total area in the nonzero blocks was 84 percent, ranging from 66 percent to 92 percent. The low value of 66 percent of the area in nonzero blocks was for New Orleans, explaining why it had by far the highest increase

Table 3. Summary Statistics for Conventional Population Density and Density Using Blocks with Nonzero Population and Urban Density Blocks.

	Conventional Population Density Using Total Area	Population Density Using Area of Blocks with Nonzero Population	Population Density Using Population and Area of Urban Density Blocks
Mean	2,543	3,074	4,113
Minimum	1,253	1,379	2,105
First Quartile	1,907	2,248	3,156
Median	2,211	2,640	3,716
Third Quartile	2,966	3,735	5,001
Maximum	5,619	7,167	8,412

in nonzero block density. Excluding more area, the mean percentage of the area in the urban density blocks was much lower at 61 percent, ranging from 48 to 70 percent. So something between one-third and one-half of the total land area of the urban areas was excluded using only the urban density blocks. Despite this dramatic reduction in area considered, using the urban density blocks retained nearly all of the populations of the urban areas. A mean of 98 percent of the total urban area population was included in the urban density blocks. The urban areas with the lowest percentages, going down to the minimum of 94 percent, were also those areas with the lowest conventional density.

While the density levels using these alternative measures were higher than conventional density, relative positioning of the urban areas was not greatly affected. The correlation of nonzero block density with conventional density was a high 0.985. This dropped only slightly to 0.975 for the correlation of urban density block density with conventional density. The lists of the areas with the highest and lowest densities changed only slightly. New Orleans replaced Miami-Fort Lauderdale-West Palm Beach at third on the list of areas with the highest nonzero block density and fourth on the list of areas with the highest urban density block density. Harrisburg-York made the lists of the 6 least dense areas for both of the density measures and Atlanta was also added to the list of the areas with the lowest urban density block density. But even these changes among the 6 most and least dense areas involved urban areas than had been in the top 10 on other lists, so the changes in position were not great.

Weighted Density

Weighted population densities—the means of the densities of small areas weighted by their populations—are the subject of this section. Results for 2 weighted population densities are presented, using the densities of census tracts and the densities of blocks.

Table 4 give the summary statistics for the tract- and block-weighted population densities for the 59 large urban areas along with the conventional population densities for comparison. The first thing that is obvious is that the weighted densities were much higher than the conventional densities. The mean tract-weighted density of 4,900 persons per square mile was nearly twice the mean conventional population density. And the mean block-weighted density of 10,400 was again 2 times larger than the tract-weighted density and 4 times the conventional density. The means that the average resident of the typical urban area experienced population density in his or her census tract that is around double the conventional density of the urban area as a whole. And this was much larger for weighted block density.

The second thing that stands out is the range of the weighted densities across the urban areas. All of the measures of population density considered before this showed substantial variation across the urban areas, with the density for the densest area

Table 4. Summary Statistics for Conventional Population Density and Tract- and Block-weighted Population Density.

	Conventional Population Density	Tract-weighted Population Density	Block-weighted Population Density
Mean	2,543	4,930	10,434
Minimum	1,253	1,608	4,852
First Quartile	1,907	3,081	6,978
Median	2,211	4,154	8,741
Third Quartile	2,966	5,143	11,332
Maximum	5,619	30,847	44,932

around 4 to 5 times the density for the least dense area. But for the weighted densities, the differences were huge, with the maximum weighted densities being nearly 10 to 20 times higher than the minimum densities. The maximum tract-weighted population density of 31,000 is over 6 times the mean density and the block-weighted maximum at 45,000 is over 4 times the mean density. However, as we shall see, these maximum values, for the New York urban area, are extreme outliers.

Table 5 lists the 6 areas with the highest and lowest tract-weighted population densities. Their conventional densities are also shown. The New York urban area had by far the highest tract-weighted density of nearly 31,000 persons per square mile, contrasted with the next highest of 11,300 persons per square mile. However, the 3 areas at the top of the list are the same 3 areas that were the highest on all of the other density measures, New York, San Francisco-Oakland-San Jose, and Los Angeles. But their ordering is reversed.

The contrast with the earlier results for the next 3 places is very different. Chicago, Philadelphia, and Boston-Providence were the next 3 most dense areas as measured by tract-weighted population density. All 3 areas failed to make the list of the top 6 with respect to conventional density. Chicago came close; it was the tenth most dense urban area. Philadelphia was only the twenty-second densest area. And the Boston-Providence area actually had a conventional density below the median for all of the urban areas.

What explains these differences? Chicago, Philadelphia, and Boston-Providence are older large urban areas that experienced high density development in earlier decades (along with New York, which is the extreme example of this). However, this was offset by quite low-density development in the newer, suburban portions of these

Table 5. Urban Areas with the Highest and Lowest Tract-Weighted Population Density

Area	Tract-weighted Population Density	Conventional Population Density
Highest Density Areas		
New York	30,847	4,609
San Francisco-Oakland-San Jose	11,281	4,824
Los Angeles	11,058	5,619
Chicago	9,154	3,350
Philadelphia	8,693	2,763
Boston-Providence	7,534	2,130
Lowest Density Areas		
Greenville-Spartanburg	1,608	1,253
Knoxville	1,960	1,387
Greensboro--Winston-Salem--High Point	2,092	1,473
Charlotte	2,160	1,515
Birmingham	2,238	1,587
Raleigh-Durham	2,354	1,440

urban areas. This produced lower levels of overall conventional density than might be expected from just looking at the urban cores. The amount by which weighted population density exceeds conventional conventional density depends on the variation in the density across the subareas such as the tracts. To be precise, this difference is the variance in the density divided by the conventional density (Ottensmann 2018). And these older large urban area had such high variation. Indeed, of the 6 areas having the highest tract-weighted density, 5 were the top areas with respect to variance in tract density and Boston-Providence was eighth. Other areas with higher conventional densities lacked the very high levels of densities in the center while higher densities in outlying areas resulted in the higher overall conventional densities.

The situation for the areas with the lowest tract-weighted population densities was very different. The 6 areas on this list were the same areas with the lowest conventional density. The list experienced only a modest reordering. This, too, can be explained by the role of the variation in density for population-weighted densities.

These areas had the lowest densities because level of densities were generally low throughout the urban areas. Thus variation in density across the tracts was low. The variance in density for these 6 areas was the lowest of all 59 areas. And therefore the weighted densities remained low.

The large differences in between the tract- and block-weighted population densities makes it clear that no single measure of the weighted density exists. Rather, weighted density depends on the set of subareas used for its calculation. Weighted densities will nearly always be greater with larger numbers of smaller subareas (and must be equal to or greater if the smaller subareas are subdivisions of the larger areas, as with blocks and census tracts). The argument for using weighted density is that it reflects the density experienced by the average resident of the urban area as more persons reside in the denser areas. The choice of the subareas to be used thus depends on assumptions regarding the extent of the area within which residents experience density. Census tracts have long been used as the basic neighborhoods for the analysis of all types of urban phenomena (hopefully with recognition that tracts are imperfect definitions of neighborhoods). So if tracts are not an unreasonable choice in many other contexts, perhaps they might be seen as a reasonable choice for the area within which residents experience density. Of course the density of the block on which one resides is even more immediate, but it is hard to imagine that the density of the block across the street is not of virtually equal importance.

That said, it is important to recognize the limitation of census tracts as units reflecting the area in which residence experience density. For statistical reasons, census tracts are delineated with the goal of including similar populations, with a target of 4,000 and a typical range from 1,200 to 8,000 (U.S. Bureau of the Census 2021b). As a result, census tracts vary widely in land area, depending on the density of the population. Some census tracts in Manhattan have areas under 0.05 square miles (around 30 acres) while many suburban tracts can have areas of several square miles. It is difficult to argue that the experience of density for residents of tracts with such different sizes is limited to the areas of the smallest tracts or encompasses the entire area of the largest tracts.

However, as a practical matter researchers are limited to the sets of areas for which the necessary data are available. In most instances in the United States, this means data from the Census. Potential subareas, from smaller to larger, include blocks, block groups, census tracts, and ZIP codes (actually ZIP code tabulation areas). Many of these have been used for calculating weighted densities. The current results for tract- and block-weighted densities show large differences in values. However, another aspect of the data suggests the issue may not be as problematic as one might expect. The correlation between tract- and block-weighted population densities for the 59 large urban areas is 0.951 while the correlation of those densities with conventional density are 0.861 and 0.759. So the weighted densities are far from perfectly correlated with conventional density, providing a different relative sense of density. But the correlation between the 2 weighted densities is very high, suggesting that the choice of subareas used might not make that much difference for any analysis.[2]

The Distribution of Density within Urban Areas

Differences in population densities between the inner and outer portions of urban area were discussed in relation to weighted densities. This section explicitly addresses the distribution of density within the large urban areas. The first part presents a basic comparison of densities in the urban core and suburban periphery. This is followed by more extensive examination of the negative exponential decline of density with distance from the center of the urban area.

Density in the Urban Core and Suburban Periphery

While innumerable comparisons have been made of various characteristics of central cities and suburbs, few have examined their densities. Two reasons might explain this: The differences in densities may have seemed so obvious—central cities are more dense and the suburbs are less dense—that the measurement of the differences might not have seemed worth pursuing. Or perhaps researchers realized that the suburban portions of Metropolitan Statistical Areas and their predecessors, used in most of the comparisons, included arbitrary amounts of rural territory making measures of density not very meaningful. Of course the suburban areas could have been defined as those portions of Urbanized Areas outside of the central cities, as Schnore (1963) did in his comparison of other socioeconomic characteristics, but that approach never caught on. Comparison of housing unit densities in the urban core and suburban periphery of the large urban areas from 1950 to 2010 were addressed in one of my earlier papers (Ottensmann 2020).

As discussed earlier, the densities of the older portions of the urban areas, those census tracts that constituted the urban areas in 1950, are compared with the densities of the remaining, newer portion of the current 2010 urban area. These parts of the urban area are referred to as the urban core and suburban periphery. This delineation is being used in place of the usual central (or principal) city-suburban distinction as the boundaries of those cities are arbitrary and encompass widely varying portions of their urban areas.

Table 6 presents summary statistics for both conventional population density and the density of blocks with nonzero population for both the urban cores and suburban

[2] I have to make the comment that calculating the tract-weighted densities is less work than the block-weighted densities, making the former choice also the most practical.

Table 6. Summary Statistics for Conventional Population Density and Density Using Blocks with Nonzero Population in the Urban Core and Suburban Periphery.

	Urban Core		Suburban Periphery	
	Conventional Population Density	Population Density Using Area of Blocks with Nonzero Population	Conventional Population Density	Population Density Using Area of Blocks with Nonzero Population
Mean	4,850	6,226	2,097	2,518
Minimum	2,298	2,725	1,115	1,192
First Quartile	3,473	4,423	1,524	1,698
Median	4,471	5,688	1,883	2,247
Third Quartile	5,678	7,001	2,607	3,179
Maximum	12,220	14,804	4,048	5,247

peripheries of the large urban areas. As expected, densities in the core were higher than in the periphery, and nonzero block densities were higher than conventional population densities. The nonzero block densities were on average 29 percent higher than the conventional densities in the urban cores and 18 percent higher in the suburban peripheries. The difference arises from the percentages of the land areas in the nonzero blocks, with the mean of 78 percent in the cores and 85 percent in the suburbs. (The urban cores have more land which is completely nonresidential, which makes sense.) Individual urban areas had percentages in nonzero areas considerably lower and higher, producing greater differences for the relationship between conventional population density and density for blocks with nonzero populations. So as when looking at the comparison of conventional with nonzero block density for the entire urban areas, nonzero block population density makes some difference for a small number of urban areas, but the general pattern is not significantly different. This is confirmed by the correlations of the two types of densities, 0.985 for the densities in both the urban cores and suburban periphery.

While the focus of this paper is on the effects of using different measures of density, it is still interesting to compare suburban densities with core densities, addressing the idea that researchers might have failed to look at the differences because of the assumption that densities were higher in the center and lower in the outer portions of urban areas. The mean densities for the core and periphery in Table 6 indicate that on average this is the case. But looking at the relationship of suburban densities to core densities for the individual urban areas shows a more complex picture.

The mean ratio of conventional population density in the suburban periphery to the density in the urban core was 2.40. It varied widely, however, from a minimum of 1.01 to a maximum of 6.49. In the El Paso urban area, with the minimum value, densities in the core and suburbs are virtually identical. And 2 more urban areas had core densities that were only about 50 percent greater than the suburban densities. At the other extreme, the urban area with the maximum suburban-core ratio of about 6.5 was not surprisingly New York. The areas with the next greatest differences, with ratios just over 4, were Philadelphia and Boston-Providence.

The lists of the urban areas with the highest and lowest conventional densities in the urban core and suburban periphery in Table 7 provide more information related to the differences between the weighted densities and the other measures of density. First, the areas with the highest densities: The 3 areas with the highest densities in the core are those that top the other lists, this time led by New York, followed by Los Angeles and San Francisco-Oakland-San Jose. It may seem surprising that the core density for New York is not that much higher than that of Los Angeles, but remember that the urban core is not Manhattan or New York City, it is the entire New York urban area as of 1950, a much larger area with other areas having lower densities. San Diego stays on the list of the areas with highest core densities, but then come Chicago and Philadelphia, older areas with higher densities in their cores that also had among the highest weighted densities.

The older urban areas including New York drop off the list of urban areas with the highest densities in the suburban periphery. All areas here are relatively newer, more rapidly growing urban areas. Those from the list of areas with the highest core densities are joined by Las Vegas, Miami-Fort Lauderdale-West Palm Beach, and El Paso. Obviously those areas with high densities in both the core and suburbs, led by Los Angeles, will also have high overall densities. The high suburban densities of Las Vegas and Miami-Fort Lauderdale-West Palm Beach can also place them on the list of areas with the highest overall conventional densities. The high core densities of areas such as Chicago and Philadelphia, and of course New York, combined with far lower suburban densities, place these areas among those with the highest weighted densities. (New York, for example, had the twenty-ninth highest suburban density, below the median for all areas.)

The list of areas with the lowest densities in their urban cores is not dissimilar to the lists of areas with the lowest overall densities, with 4 also being on that list. All are again somewhat smaller urban areas in the South. And 3 of those areas also are among the areas with the lowest densities in their suburban peripheries, so they have lower overall densities. But the list of the areas having the lowest densities in the suburbs includes some areas that have not appeared on prior lists—Albany-Schenectady-Troy, Pittsburgh, and Harrisburg-York. These are areas in the Northeast, not among the

Table 7. Urban Areas with the Highest and Lowest Conventional Population Density in the Urban Core and Suburban Periphery.

Urban Core		Suburban Periphery	
Area	Conventional Population Density	Area	Conventional Population Density
Highest Density Areas			
New York	12,220	Los Angeles	4,048
Los Angeles	10,310	Las Vegas	3,860
San Francisco-Oakland-San Jose	9,544	San Francisco-Oakland-San Jose	3,715
San Diego	8,212	San Diego	3,361
Chicago	7,927	Miami-Fort Lauderdale-West Palm Beach	3,178
Philadelphia	7,179	El Paso	3,120
Lowest Density Areas			
Greenville-Spartanburg	2,298	Albany-Schenectady-Troy	1,115
Birmingham	2,388	Greenville-Spartanburg	1,134
Knoxville	2,388	Pittsburgh	1,187
Oklahoma City	2,896	Knoxville	1,222
Greensboro--Winston-Salem--High Point	2,915	Harrisburg-York	1,273
Nashville	2,927	Greensboro--Winston-Salem--High Point	1,300

largest areas, that have extremely low densities in the suburbs more like low-density urban areas otherwise found in the South.

Negative Exponential Decline of Density

Most efforts devoted to examining the variation of population density within urban areas have assumed that densities decline as a negative exponential function of distance from the center and have estimated the parameters of that model. Here the model is estimated using different measures of population density, with both census tracts and blocks being used as the units of analysis.

The population densities of census tracts have been used most often when the negative exponential model has been estimated using data for small subareas within the urban area.[3] The initial models presented here estimate the model parameters using conventional population and housing unit densities for the census tracts in the urban areas. Next are models using different measures of tract population density. Comparison of results for models using densities for tracts and blocks follow.

Table 8 present summary statistics for the estimates of models using tract population and housing unit densities. The natural logarithm of density is regressed on distance from the central business district (CBD) in miles. This table and the following tables with exponential model results give the values for the estimated density gradient, β in the equations presented earlier, the central density D_0 , and R^2, the goodness-of-fit, the measure of the proportion of the variance in the log of density accounted for by the model.

The mean estimated density gradient using housing unit density, 0.067, was higher than the gradient for population density, 0.058. The difference was statistically significant. Similar increases can be seen for the minima and maxima. One possible explanation for the difference could be that lower numbers of persons per housing unit in tracts closer to the center would depress population density relative to housing unit density, resulting in a lower difference in density between the center and more outlying areas, therefore producing a lower density gradient.

The fit of the model was also better using housing unit density, with R^2 of 0.193 for housing unit density versus 0.173 for population density, also a statistically significant difference. The standard monocentric model in urban economics, while describing the decline of population density, is actually referring to the tradeoffs made by individuals in locating residences. These decisions about housing units are seen as equivalent to population densities because of the assumption in the model that all households are identical. And since they are not, housing units and housing unit densities might be more appropriate than population. (Note that estimated central densities are naturally far lower the model using housing unit densities because housing unit densities are a fraction of population densities.)

While housing unit densities may actually be a better choice than population densities for estimating the negative exponential model, the following models will all use population densities. This is to keep the results more comparable with other research on the negative exponential decline of density because population densities have been almost universally used in such studies. The one exception using housing unit densities of which I am aware is my work estimating the model for the large urban

[3] A frequently used alternative for estimating the parameters of the exponential model described by Mills (1972) *assumes* density declines with the negative exponential of distance and uses the populations of the central city and the entire metropolitan area and the radius of the central city to make a 2-point estimate of the model parameters. Such a procedure does not allow assessment of the goodness-of-fit of the model.

Table 8. Summary Statistics for Density Gradient, Central Density, and R^2 for the Negative Exponential Model Estimated Using Tract Conventional Population Density and Conventional Housing Unit Density.

	Tract Conventional Population Density	Tract Conventional Housing Unit Density
Density Gradient		
Mean	0.058	0.067
Minimum	0.014	0.022
Maximum	0.226	0.240
Central Density		
Mean	6,467	2,906
Minimum	2,572	1,257
Maximum	33,187	12,856
R^2		
Mean	0.173	0.193
Minimum	0.007	0.023
Maximum	0.500	0.510

areas for years from 1950 to 2010. In that case, housing unit densities were used not because they were necessarily seen as a superior measure but because they were the only measures of density for the earlier years being considered (Ottensmann 2016).

Next are the estimates of the negative exponential decline of density using alternative measures of tract population density—conventional density, density using the area of blocks with nonzero population, and density using the population and area of block having urban densities. It would have been possible to calculate block-weighted densities for individual census tracts. However questions about the confounding of the block and tract densities and the meaning that would be associated with any models estimated using such densities led to the decision not to do so.

Summary results for the negative exponential models estimated using the 3 alternative measures of tract population density are presented in Table 9. The density gradient increased from 0.58 for the estimates use tract conventional density to 0.69 for the nonzero block density. But for urban density block density, the gradient dropped back down to 0.56, nearly the same as for conventional density. Differences from the

Table 9. Summary Statistics for Density Gradient, Central Density, and R^2 for the Negative Exponential Model Estimated Using Tract Conventional Population Density, Tract Population Density Using the Area of Blocks with Nonzero Population, and Tract Population Density Using the Population and Area of Urban Density Blocks.

	Tract Conventional Population Density Using Total Area	Tract Population Density Using Area of Blocks with Nonzero Population	Tract Population Density Using Population and Area of Urban Density Blocks
Density Gradient			
Mean	0.058	0.069	0.056
Minimum	0.014	0.025	0.024
Maximum	0.226	0.237	0.193
Central Density			
Mean	6,467	8,509	8,864
Minimum	2,572	3,337	3,890
Maximum	33,187	40,164	41,454
R^2			
Mean	0.173	0.257	0.257
Minimum	0.007	0.069	0.092
Maximum	0.500	0.632	0.625

nonzero block gradient were statistically significant. It may be that by excluding higher proportion of blocks in the inner portions of the urban areas, as observed in the core-periphery comparisons, nonzero block densities were increased more compared to conventional density in the core than at the periphery. This would produce a higher density gradient. And it is further possible possible that the exclusion of blocks having less than urban density is more uniform across the urban areas, reversing this effect.

This differential effect of excluding varying proportions of blocks in different parts of the urban areas may be reflected in the estimates of central densities as well. Densities necessarily increase as more blocks are excluded, and this was seen in the results for the densities for the entire urban areas presented in Table 3. The same holds for the estimates of the central densities. However for those urban area densities, the mean density using nonzero blocks of about 3,100 was closer to the mean conventional density of 2,500 than to the mean density using urban density blocks, which was 4,100. But for the estimate of mean central density using nonzero blocks of 8,500 was much closer to the mean for urban density blocks of 8,900 than to the mean using conventional density, which was 6,500. This may have been caused by the relatively higher densities near the center resulting from the exclusion of higher proportions of zero density blocks.

The two models estimated using densities with blocks excluded have a better fit and explained more of the variation in tract population density than the model using conventional tract density. The R^2 values for both of those models were identical, 0.257, while the value for the conventional density model was 0.173. The monocentric model addresses the choices made with respect to lot size and density, in other words, residential density. It is reasonable that models estimated using density measures that exclude at least some nonresidential area and come closer to the residential density assumed in the model would perform better. Take the example of a census tract in which half of the blocks and tract area were nonresidential, with zero population, and half of the blocks were completely residential excluding the blocks with zero population would double the density, reflecting the density choice of residents in that tract at that distance from the center.

The modifiable areal unit problem (MAUP) suggests that the choice of areas for conducting an analysis can affect the results. As the densities for blocks are available along with tracts, the opportunity is presented to compare the effect of using blocks for the estimation of the negative exponential model as compared with census tracts. As mentioned earlier, comparisons using conventional density are not possible given the manner in which the models are being estimated. Estimation using the linear-transformed model requires the use of the logarithm of density. But the logarithm is not defined for zero, so attempting to estimate the model for all blocks results in the zero density blocks being excluded as missing values, resulting in estimates identical to those obtained by explicitly excluding those blocks. Therefore comparison will be made of estimates of the model using tract density with nonzero blocks and block densities using nonzero blocks. Estimates for densities using urban density blocks are not presented as they show essentially the same picture as the nonzero block densities.

Table 10 presents the usual results for a model estimated using census tracts and tract densities with nonzero blocks and a model estimated using nonzero blocks and their densities. The mean density gradient dropped significantly from 0.069 using tracts to 0.044 using blocks. A comparable decline occurred for the estimates of central density, from 8,500 to 7,000. These differences are what would generally be expected in moving to the estimation of a model using larger numbers of more disaggregated areal units.

The most dramatic differences comes with the fit of the model. The mean R^2 plummeted from 0.257 using tracts to 0.058 using blocks. A decline again is exactly what was expected though the magnitude of the drop was not. Aggregating smaller areas

Table 10. Summary Statistics for Density Gradient, Central Density, and R^2 for the Negative Exponential Model Estimated Using Tract Population Density Using the Area of Blocks with Nonzero Population and Block Population Density Using Blocks with Nonzero Population.

	Tract Population Density Using Area of Blocks with Nonzero Population	Block Population Density Using Blocks with Nonzero Population
Density Gradient		
Mean	0.069	0.044
Minimum	0.025	0.014
Maximum	0.237	0.168
Central Density		
Mean	8,509	6,959
Minimum	3,337	2,899
Maximum	40,164	17,811
R^2		
Mean	0.257	0.058
Minimum	0.069	0.003
Maximum	0.632	0.245

into fewer, larger areas will nearly always increase the fit of models and vice versa. In this case, the effect was very large.

Despite the very large differences in the parameter estimates in moving from tracts to blocks and the extremely poor fit of the models using the block data, the relative magnitudes of the tract and block results were reasonably consistent. The correlation of the tract and block density gradient estimates across the urban areas was 0.920 and for the central density it was 0.901. The correlation was somewhat lower but still quite high, 0.794, for the R^2 values. In other words, the estimates using block densities are still legitimate estimates. The effect of the modifiable areal unit problem is this case was mixed: Changing the areal units used had a major effect on the magnitude of the estimates but a more limited effect on the relative patterns of those estimates across the urban areas.

Conclusions

As expected, conventional housing unit density for urban areas closely tracked population density. Relative positions of urban areas were largely the same. The few variations were caused by those urban areas with extreme population to housing unit ratios that might be of interest in certain cases.

The difference between housing unit and population density was more significant for the estimation of the negative exponential model. When estimated using conventional densities, the gradient was higher for housing unit density and the fit of the model as measured by R^2 is also greater. The argument can certainly be made that housing unit density is a better measure of density for looking at the decline with distance from the center.

The measures of population density calculated using the areas excluding blocks that were completely or largely nonresidential showed very similar patterns across the urban areas to conventional density. The values were obviously higher, with the mean density using only urban density blocks 60 percent greater than conventional density, 4,100 versus 2,500 persons per square mile. These densities, while not limited to completely residential areas, give a much better idea of the levels of residential densities.

The standard monocentric model predicts the negative exponential decline of *residential* densities. But estimates of model parameters have generally been made using conventional population densities, which reflect the inclusion of large areas of nonresidential land. Using the measures of density that excluded the areas of blocks that were completely or largely nonresidential yielded models that accounted for significantly more of the variation in density, increasing the mean R^2 values by nearly 50 percent from 0.173 to 0.257 using both both alternative measures of density. This provides confirmation that while these are not measures of pure residential density, they seem to go along way in moving towards that from conventional density.

Weighted population density is a completely different form of density measure as shown by the results. All weighted densities are larger than conventional densities, but the extent of the increase varies widely depending on the variation in densities within the urban areas. Areas having some very dense portions, near their centers, offset by low densities in the areas farther out saw the greatest increases for weighted density. These include especially the large, older urban areas in the Northeast and Midwest, which had some of the highest weighted densities despite some having relatively modest conventional densities. And this makes the New York urban area an extreme high outlier in terms of weighted density. Differences in densities between the urban core and suburban periphery are consistent with the observations for weighted densities.

The comparison of the estimates of the negative exponential model using tracts and blocks as the units analysis provides some helpful understanding of the impact of the modifiable areal unit problem in this context. Estimates of the density gradient and central density declined significantly when changing from tracts to blocks as the units of analysis as expected with the shift to larger numbers of smaller areas. This occurred as R^2 fell to near zero. However it seems somehow encouraging that even with this drastic drop in the goodness-of-fit, estimates of the model parameters were consistent with, if lower than, those obtained using the tract densities, with high correlations across the urban areas. This suggests that while the choice of the areal units can affect the magnitudes of estimates, patterns of variation across the urban areas may not be that greatly affected.

No single measure of density can be the most suitable for all analyses of urban areas. Understanding the range of alternatives and how they differ can play a role in making the most appropriate selection.

References

Alexander, Ernest R. 1993. Density measures: a review and analysis. *Journal of Architectural and Planning Research* 10, 3 (Autumn): 181-202.

Baser, Onur. 2000. Population density index and its use for distribution of Covid-19: a case study using Turkish data. *Health Policy* 125, 2 (February): 148-154.

Boeing, Geoff. 2018. Estimating local daytime population density from census and payroll data. *Regional Studies, Regional Science* 5, 1: 179-182.

Borruso, Giuseppe. 2003. Network density and the delimitation of urban areas. *Transactions in GIS* 7, 2 (March): 177-191.

Boyko, Christopher T., and Rachel Cooper. 2011. Clarifying and re-conceptualizing density. *Progress in Planning* 76, 1 (July): 1-61.

Brownstone, David, and Thomas F. Golob. 2009. The impact of residential density on vehicle usage and energy consumption. *Journal of Urban Economics* 65, 1 (January): 91-98.

Burchell, Robert W. 1997. Economic and fiscal costs (and benefits) of sprawl. *The Urban Lawyer* 29, 2 (Spring): 159-181.

Burchell, Robert W., and 9 others. 1998. *The Costs of Sprawl—Revisited*. Transit Cooperative Research Program Report 39. Transportation Research Board, National Research Council. Washington DC: National Academy Press.

Burchell, Robert W., and 7 others. 2002. *Costs of Sprawl—2000*. Transit Cooperative Research Program Report 74. Transportation Research Board, National Research Council. Washington DC: National Academy Press.

Carlino, Gerald, Satyajit Chatterjee, and Robert Hunt. 2006. Urban density and the rate of invention. Working Paper NO. 06-14. Federal Reserve Bank of Philadelphia. Downloaded from *https://www.researchgate.net/publication/314904310_Urban_Density_and_the_Rate_of_Invention* on June 6, 2021.

Carozzi, Felipe, Sandro Provenzano, and Sefi Roth. 2020. Urban density and Covid-19. CEP Discussion Paper No 1711. London School of Economics and Political Science. Downloaded from *https://cep.lse.ac.uk/pubs/download/dp1711.pdf* on July 19, 2021

Churchman, Arza. 1999. Disentangling the concept of density. *Journal of Planning Literature* 13, 4 (May): 389-411.

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

Craig, John. 1984. Averaging population density. *Demography* 21, 3 (August): 405-412.

Craig, John. 1985. Better measures of population density. *Population Trends* No. 39: 16-21.

Cramer, Victoria, Svenn Torgersen, and Einar Kringlen. 2004. Quality of life in a city: the effect of population density. *Social Indicators Research* 69, 1 (October): 103-116.

Ellis, John G. 2004. Explaining residential density. *Places* 16, 2 (Summer): 34-43.

Ewing, Reid. 1997. Is Los Angeles-style sprawl desirable? *Journal of the American Planning Association* 63, 1 (Winter): 107-126.

Fotheringham, A. Stewart, and David Wong. 1991. The modifiable areal unit problem in multivariable statistical analysis. *Environment and Planning A* 23, 7 (July): 1025-1044.

Galster, George, and 5 others. 2001. Wrestling sprawl to the ground: defining and measuring an elusive concept. *Housing Policy Debate* 12, 4: 681-717.

Garland, Patrick, and 5 others. 2020. The COVID-19 pandemic as experienced by the individual. Downloaded from *https://arxiv.org/abs/2005.01167* on June 2, 2021.

Glaeser, Edward L., and Matthew E. Kahn. 2004. Sprawl and urban growth. In *Handbook of Regional and Urban Economics*, vol. 4, J. V. Henderson and J. F. Thisse, eds. Amsterdam: Elsevier, pp. 2481-2527.

Gordon, Peter, and Harry W. Richardson. 1997. Are compact cities a desirable planning goal? *Journal of the American Planning Association* 63, 1 (Winter): 95-106.

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 (August): 183-199.

Hawley, Amos H. 1972. Population density and the city. *Demography* 9, 4 (November): 521-529.

Kaddoura, Ihab, and Tilmann Schlechter. 2021. the impact of trip density on the fleet size and pooling rate of ride-hailing services: a simulation study. *Procedia Computer Science* 184: 674-679.

Knudsen, Brian, Richard Florida, Kevin Stolarick, and Gary Gates. 2008. Density and creativity in U.S. regions. *Annals of the Association of American Geographers* 98, 2 (June): 461-478.

Ladd, Helen F. 1992. Population growth, density and the cost of providing public services. *Urban Studies* 29, 2 (April): 273-295.

Lee, Sungwon, and Bumsoo Lee. 2014. The influence of urban form on GHG emissions in the U.S. household sector. *Energy Policy* 68 (May): 534-549.

Lipka, Michael. 2015. Mormons more likely to marry, have more children than other U.S. religious groups. Pew Research Center. Downloaded from *https://www.pewresearch.org/fact-tank/2015/05/22/mormons-more-likely-to-marry-have-more-children-than-other-u-s-religious-groups/* on July 25, 2021.

McDonald, John F. 1989. Econometric studies of urban population density: a survey. *Journal of Urban Economics* 26, 3 (November): 361-385.

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

Morton, Anthony B. 2015. Population-weighted density, density-weighted population, granularity, paradoxes: A recapitulation. *arXiv*:1412.4332.v2. (website). Accessed from *https://arxiv.org/pdf/1412.4332.pdf* on February 8, 2017.

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

Openshaw, Stan. 1984. *The Modifiable Areal Unit Problem*. Concepts and Techniques in Modern Geography No. 38. Norwich: Geo Books.

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. 2016. The negative exponential decline of density in large urban areas in the U.S., 1950-2010. Available at *https://urbanpatternsblog.files.wordpress.com/2016/12/negative-exponential-density-decline.pdf*.

Ottensmann, John R. 2017. On the choice of Combined Statistical Areas. Available at *https://urbanpatternsblog.files.wordpress.com/2017/04/combined-statistical-areas.pdf*

Ottensmann, John R. 2018. On population-weighted density. Available at *https://urbanpatternsblog.files.wordpress.com/2018/02/on-population-weighted-density.pdf*.

Ottensmann, John R. 2019. Central city and suburb versus urban core and suburban periphery. Available at *https://urbanpatternsblog.files.wordpress.com/2019/09/central-city-vs-urban-core.pdf*.

Ottensmann, John R. 2020. Population density in the urban core and suburban periphery of large urban areas in the U.S., 1980-2010. Available at *https://urbanpatternsblog.files.wordpress.com/2020/04/core-periphery-density.pdf*.

Paulsen, Kurt. 2012. Yet even more evidence on the spatial size of cities: urban spatial expansion in the US, 1980-2000. *Regional Science and Urban Economics* 42, 4 (July): 561-568.

Paulsen, Kurt. 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., and 5 others. 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 I. 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.

Rae, Alasdair. 2018. Think your country is crowded? These maps reveal the truth about population density across Europe. Downloaded from *https://theconversation.com/think-your-country-is-crowded-these-maps-reveal-the-truth-about-population-density-across-europe-90345* on June 6, 2021.

Rapoport, Amos. 1975. Toward a redefinition of density. *Environment and Behavior* 7, 2 (June): 133–158.

Real Estate Research Corporation. 1974. *The Costs of Sprawl: Environmental and Economic Costs of Alternative Residential Development Patterns at the Urban Fringe*. Vol. 2, *Detailed Cost Analysis*. Prepared for the Council on Environmental Quality, the Office of Policy Development and Research, Department of Housing and Urban Development, and the

Office of Planning and Management, Environmental Protection Agency. Washington DC: U.S. Government Printing Office.
Schnore, Leo F. 1963. The socio-economic status of cities and suburbs. *American Sociological Review* 28, 1 (February): 76-85.
Sifakis, N.I., 1998. Quantitative mapping of air pollution density using earth observations: a new processing method and application to an urban area. *International Journal of Remote Sensing* 19, 17: 3289-3300.
Smith, Betty E. 1997. A review of monocentric urban density analysis. *Journal of Planning Literature* 12, 2 (November): 115-135.
Theobald, David M. 2001. Land-use dynamics beyond the urban fringe. *Geographical Review* 91, 3 (July): 544-564.
Thrall, Grant Ian. 1988. Statistical and theoretical issues in verifying the population density function. *Urban Geography* 9, 5: 518-537.
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.
U.S. Bureau of the Census. 2021a. 2010 census tallies. Accessed at *https://www.census.gov/geographies/reference-files/time-series/geo/tallies.html* on June 30, 2021.
U.S. Bureau of the Census. 2021b. Geography program glossary. Accessed at *https://www.census.gov/programs-surveys/geography/about/glossary.html#par_textimage_13* on July 5, 2021.
U.S. Geological Survey. 2021. National Land Cover Database. Accessed at *https://www.usgs.gov/centers/eros/science/national-land-cover-database?qt-science_center_objects=0#qt-science_center_objects* on June 30, 2021.
Urban Institute and Geolytics. 2003. *Census CD Neighborhood Change Database (NCDB): 1970-2000 Tract Data*. Somerville, NJ: Geolytics.
Walton, D., S. J. Murray, and J. A. Thomas. 2008. Relationships between population density and the perceived quality of neighborhood. *Social Indicators Research* 89, 3 (December): 405-420.
Windsor, Duane. 1979. A critique of *The Costs of Sprawl*. *Journal of the American Planning Association* 45, 3 (July): 279-292.
Wong, David. 2009. The modifiable areal unit problem (MAUP). in *The Sage Handbook of Spatial Analysis*, A. Stewart Fotheringham and Peter A. Rogerson, eds. Thousand Oaks, CA: Sage Publications, pp. 105-123.
Wirth, Louis. 1938. Urbanism as a way of life. *American Journal of Sociology* 44, 1 (July): 1-24.