On Population-Weighted Density

John R. Ottensmann
Indiana University-Purdue University Indianapolis
john.ottensmann@gmail.com
urbanpatternsblog.wordpress.com
February 2018

Abstract

Population-weighted density is the mean of the densities of subareas of a larger area weighted by the populations of those subareas. It is an alternative to the conventional density measure, total population divided by total area. This paper shows that population-weighted density is equal to conventional density plus the variance in density across the subareas divided by the conventional density. This density alternative is very dependent on the size and configuration of the subareas, an issue that has not been adequately addressed by most users. Urban sprawl is associated with lower densities, and the choice of the appropriate density measure is dependent upon the negative consequences of sprawl being considered. Comparison of conventional and population-weighted densities show the latter to be more highly skewed and to sometimes rank urban area densities very differently. Population-weighted density is more strongly related to the size of the urban area, especially size in earlier years, demonstrating the effect of the timing of urban growth on density.

Introduction

Density is one of the most fundamental characteristics of an urban area. The basic idea of density is simple—population or some other count divided by area. Some have proposed an alternative measure of density generally called population-weighted density, which consists of the mean of the densities of small subareas weighted by the populations of those smaller areas. Because the densities are weighted by the subarea populations, this provides a measure of density that reflects the average density as experienced by the residents of the urban area within their subareas.

Population-weighted density is a valid alternative to conventional density that captures aspects that the other measure does not. But some proponents of this measures have gone farther. It is one thing to maintain that population-weighted density might be a "more meaningful" or "informative" measure of density than the conventional measure (Avent 2011, Bradford 2008a). But others claim that population-weighted density is "a better way to measure density" (Florida 2012) and "the best way to measure density (Davies 20012b).” Indeed, the conventional measure of density is dismissed as “a meaningless number” (Bradford 2008a) that is a “counterintuitive” measure of density (Bradford 2008c, Davies 2012a) and “leads to some funny results (Yglesias 2016).” Duke (2012) went so far as to say that population-weighted density eliminated the “boundary games” associated with conventional density, essentially implying that those using that measure were intentionally attempting to mislead.

A more nuanced approach is provided by Craig (1984), who says that “the conventional crude population density is not a good measure of the density *at which the population lives*.” The emphasis is his, highlighting that population-weighted density is a better measure for addressing that particular question. He further states (Craig 1985) that ”which measure is appropriate depends on whether population density is being used as a land-use measure or as a social indicator of the population densities experienced by the population.” This paper adopts Craig’s view that population-weighted density is an alternative measure that has advantages for certain purposes but cannot be considered to be a generally superior measure of density.

A simple analogy can illustrate how an alternative measure can be better for some purposes while not being a better overall measure. I grew up in the upper Midwest where temperatures in winter could get very cold, sometimes well below zero. At those times, we would pay close attention to the wind chill, a composite of the actual temperature and the wind speed. This alternative temperature measure was important not just as an indicator of how cold one would feel outside but as an indicator of the risk of frostbite and hypothermia. But certainly wind chill cannot be considered a better measure of temperature that could supplant the conventional, actual temperature. It would not be appropriate to substitute high and low wind chills in reporting daily weather, nor would wind chill provide an indication of whether certain substances would freeze.

This paper comprehensively examines population-weighted density and its use as an alternative to the conventional density measure. It begins with the definition of the densities and the demonstration that population-weighted density is a simple function of conventional density and the variance in density across the subareas. Issues that arise with both conventional and population-weighted density are discussed next. Population-weighted density has been used in a variety of contexts and some examples are presented. Urban sprawl is associated at least to some extent with low-density development, and both conventional density and population-weighted density can be relevant to the examination of sprawl in different contexts. Finally, conventional and population-weighted densities for 59 large urban areas in the United States in 1950 and 2010 are compared.

The Relationship between Population-Weighted Density and Conventional Density

This section presents the derivation of the relationship between population-weighted density and conventional density. It starts with the definitions of the two measures of density. The density D of an urban area (which will be referred to as "conventional density" to distinguish it from population-weighted density) is the total population or other measure of interest such as housing units P divided by the total area A:

$$D = \frac{P}{A}$$

By definition, the total population and area of the urban area are simply the sums over census tract (or other subarea) populations and areas:

$$P = \sum p_i$$

$$A = \sum a_i$$

where p_i is the population and a_i is the area of tract i. The density d_i of tract i is its population divided by its area:

$$d_i = \frac{p_i}{a_i}$$

Also relevant for comparisons with population-weighted density, conventional density is the mean of the densities of census tracts (or other subareas) weighted by their areas:

$$D = \frac{1}{A}\sum a_i d_i = \frac{1}{A}\sum p_i = \frac{P}{A}$$

Thus, conventional density is area-weighted density as contrasted with population-weighted density.

Population-weighted density D_P is the mean of the census tract densities weighted by the populations of the tracts:

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

Craig (1984, 1985) suggests that the extent to which population-weighted density exceeds conventional density will depend on the variation in density across the subareas. However, in comparing population-weighted density with conventional density, Craig (1984, 1985) and Dorling and Atkins (1995) have used the absolute differences and the ratios of the two types of densities.

We now show that the difference between population-weighted density and conventional density is a simple function of the variance in density across the census tracts and conventional density.[1]

Calculating the variance requires the mean density across the tracts. For the calculation of both the mean density and the density variance, it is necessary to weight the calculations by the areas of the tracts, as these differ. So the mean density is therefore the same as the overall density:

$$\bar{D} = \frac{1}{A}\sum a_i d_i = \frac{1}{A}\sum p_i = D$$

The density variance s_D^2 is then the area weighted sum of the squared differences between the tract densities and the mean density divided by the total area:

$$s_D{}^2 = \frac{1}{A}\sum a_i \left(d_i - \bar{D}\right)^2$$

Note that since we have information on all of the tracts, not just a sample, we do not need to do the equivalent of the *(n - 1)* correction for the weighted variance.

The manipulation of the variance starts with the squaring of the binomial term in parentheses, splitting the sum, factoring out constants from the sums, and substituting conventional density for mean density:

$$s_D{}^2 = \frac{1}{A}\left(\sum a_i d_i{}^2 - 2D\sum a_i d_i + D^2 \sum a_i\right)$$

The sum in the second term, tract area times density, is the sum of the tract populations, the total population *P*, which is equal to the area times the density, *AD*. Substituting that along with the total area *A* for the sum of the tract areas in the third term gives

[1] Lewontin and Levins (1989) attempted to derive the relationship between conventional and population-weighted density, taking a more indirect approach that required making approximations and producing a result they described as a "poor approximation." Their result includes both variances and covariances. Then by assuming equal areas for the subdivisions they get a result similar to that reached here.

$$s_D{}^2 = \frac{1}{A}\left(\sum a_i d_i{}^2 - 2AD^2 + AD^2\right) = \frac{1}{A}\left(\sum a_i d_i{}^2 - AD^2\right) = \frac{1}{A}\sum a_i d_i{}^2 - D^2$$

To show this to be related to the difference between population-weighted density and density, begin by substituting $p_i / a_i = d_i$ for one d_i in the first term:

$$s_D{}^2 = \frac{1}{A}\sum a_i d_i \left(\frac{p_i}{a_i}\right) - D^2 = \frac{1}{A}\sum d_i p_i - D^2$$

Multiply the first term by D/D and substitute $P = AD$,

$$s_D{}^2 = \frac{D}{AD}\sum d_i p_i - D^2 = \frac{D}{P}\sum d_i p_i - D^2$$

Factor D out from both terms, giving

$$s_D{}^2 = D\left(\frac{1}{P}\sum d_i p_i - D\right)$$

The first term inside the parentheses is the formula for population-weighted density, so

$$s_D{}^2 = D\left(D_P - D\right)$$

The variance in density is the difference between population-weighted density and conventional density multiplied by the conventional density. Therefore the difference between population-weighted density and conventional density is

$$D_P - D = \frac{s_D{}^2}{D}$$

which is the density variance divided by the conventional density. Alternatively, population-weighted density is

$$D_P = D + \frac{s_D{}^2}{D}$$

so population-weighted density is simply conventional density plus the variance of density divided by conventional density.

Issues with Conventional and Population-Weighted Density

Issues arise with respect to both conventional and population-weighted density suggesting that their use as measures of density is not as straightforward as it may seem. The major issue with conventional density is that this measure is very sensitive to the extent of the territory included in the urban area or other area for which density is being calculated. This has been noted by many including Barnes (2001), Davis (2015), and Bradford (2008c).

This problem is exemplified in considering the densities of Metropolitan Statistical Areas (MSAs). These areas are defined using counties as the building blocks. Counties included in MSAs can include substantial areas that are very lightly populated and not clearly related to the urban area. This additional area reduces the density calculated for the area. Extreme examples of this are in the West, where counties can be very large. San Bernardino County is one of 2 counties making up the Riverside-San Bernardino-Ontario MSA. This is the largest county in the lower-48 states. It extends from Los Angeles County about 200 miles east to the Nevada and Arizona borders and has a land area of over 20,000 square miles, larger than 9 states. Nearly all of the population is concentrated in the southwestern corner of the county. A major portion of the county is the Mohave Desert. This area is not just rural, it is essentially empty. Driving across the county on Interstate 40, you encounter a warning sign before an interchange that informs you that it will be over 80 miles until the next gas station. Obviously including all of this empty land in any calculation of density gives an extremely artificially lowered value that does not reflect the character of the actual urban development in the MSA adjacent to Los Angeles.

Morton (2015) states that therefore conventional density is “a contentious measure, owing to the question of where to place the boundary.” He explicitly notes the arbitrary nature of administrative boundaries. But he goes on to say that this problem can be addressed by defining the urban area by aggregating small areas using a minimum-density cutoff. This is exactly the approach taken by the census in defining Urbanized Areas and in the current research reported here.

The conclusion with respect to this issue should be that conventional density should be used as a measure of density only for areas that have been defined in such a ways that they do not include sparsely settled territory or other area that biases the measure, as is the case with MSAs. Note that population-weighted density, by giving low weights to the low-density subareas, minimizes this problem for such areas. (It does not, however, eliminate the problem, as noted by Coombes and Raybould 2001).

This issue has led some of the more ardent proponents of population-weighted density to strongly dismiss conventional density. Davis (2015) makes the point about problem of arbitrary boundaries for conventional density. Loader (2013) says conventional density is not very meaningful “because of the inclusion of vast swaths of unpopulated land within "metropolitan areas." Both fail to acknowledging the possibility that conventional density could be appropriate in other contexts where the boundaries are not so arbitrary. Duke (2012) says population-weighted density "is less impacted by large unpopulated areas, largely eliminating boundary games," essentially implying that those using conventional density are intentionally trying to mislead by manipulating the areas for which density is being computed. (This is, of course, ridiculous in choosing to compute densities for MSAs, the boundaries of which are not defined by the those computing densities for these areas.)

Population-weighted density requires densities for subareas of the larger area for which the density measure is to be computed. The selection and nature of the set of subareas to be used is the source of numbers of issues related to this measure of density. Perhaps the most important question for computing population-weighted density is the number and sizes of the subareas. Craig (1984, 1985) and Morton (2015) make the point that the use of smaller subareas will increase the population-weighted density. The previous section demonstrated that population-weighted density increases with the variance in density across the subareas. The smaller the subareas, the greater the variation, since larger subareas incorporate more of the density variation within those areas, resulting in less between-subarea variation. Craig (1985) says that the greater the number of subareas, the greater variation in density. Lewontin and Levins (1989) argue that that smaller areas will have greater within-area homogeneity and higher levels of between-area heterogeneity, making for greater population-weighted variance.

Many different subarea divisions have been used for the calculation of population-weighted densities for urban and metropolitan areas. In most instances, the authors have simply declared which subareas are being used without providing any explanation or justification for their choice. Consider just the units used for calculating population-weighted density for urban areas in the United States: Starting with the smaller subareas, Barnes (2001) used traffic analysis zones and Lee and Lee (2004) used block group data from the census. Census tract densities were the source for the densities calculated by Bradford (2008b) and by the Census for their report on change in metropolitan areas (U.S. Bureau of the Census 2012b). Glaeser and Kahn (2004) used ZIP codes. Rappaport (2006, 2008) calculated population-weighted density using as subareas the intersections of county subdivisions and municipalities within metropolitan areas.

The effect of the choice of subareas, especially across such a wide spectrum, on the values for population-weighted density is huge. Direct comparisons among many of the studies cited above is difficult because they calculate densities for different areas and use data for different years. The population-weighted densities calculated by two of the studies can be directly compared as they use census data for 2000 and calculate the densities for MSAs using post-2000 definitions. The census report (U.S. Bureau of the Census 2012b) used census tracts as the subareas. Rappaport (2006) used the intersections of county subdivisions and municipalities. Comparisons are made for the densities of the Los Angeles and New York areas, perhaps the most common comparison for works presenting population-weighted densities. The choice of subareas does not affect the conventional densities. The census reported a density for Los Angeles of 2,646 persons per square mile and for New York, 2,826. For population-weighted density, the census reported the value of 12,100 persons per square mile for Los Angeles using census tracts as the subareas, while Rappaport's density using the much larger subareas was 7,800. And the difference for the New York area was even greater, 31,300 versus 18,900. (Rappaport rounded densities to the nearest hundred, so the same has been done for the more detailed population-weighted densities from the census report.) So this leaves hanging the question of what is the population-weighted density? The only answer can be that "it depends" on the set of subareas used for the calculation.

The logic underlying population-weighted density is that it represents the average density experienced by the residents of the urban area. But this raises the question as to what might be the size of the area around their dwellings within which people experience density. This is a question that has generally not been addressed but is fundamental to the selection of the appropriately sized subareas for computing population-weighted density.

Barnes (2001) does address this issue, saying that "the density of a traffic zone is thus a good measure of the immediate few blocks around a person's home." However, he then proceeds to raise another issue relating to subareas and population-weighted density when he describes the sizes of traffic zones as ranging from as small as a single block to more than a square mile in outlying, less dense areas. This large variation in size affects not only traffic analysis zones but other units frequently used as subareas for population-weighted density, including block groups, census tracts, and ZIP codes. It is difficult to believe that someone's experience of the density within a few blocks of their residence is the same as their experience of the overall density within an area that could be measured in miles.

Barnes (2001) minimizes the problem by suggesting that nearby areas will tend to have similar densities, which is true. Except when it isn't. Remember that population-weighted density is based on the premise that densities (and the experiences of residents of density) varies across the urban area. As has been shown, population-weighted density varies with the variance of density across the subareas. So maintaining that the variation in the sizes of the subareas is not a problem because densities do not vary that much is necessarily a rather weak argument.

A fairly small, minor point can also be added with respect to the experience of density by the residents of subareas. The density of the entire subarea is weighted by the total population living within that area. So the assumption must be that every resident of the subarea shares that same experience of density. But what about a person living near the boundary of that subarea? What he or she experiences in terms of density may depend as much on the densities of adjacent subareas as on the density of the subarea of residence. Once again, a tendency of densities to frequently be similar in adjacent subareas would reduce the significance of this issue.

One other characteristic of urban areas and the distributions of both residences and other uses across the subareas will exacerbate the differences in densities between nearby areas. The densities generally being considered, both for the entire urban area for conventional density and for the subareas used for population-weighted density, are gross densities. That is, they are the total population of the area or subarea divided by the total land area. This land area includes both residential areas and other land uses. The latter could include, for example, commercial and industrial uses, parks, and even vacant land. Working with census data does not allow the computation of net residential densities, which is population divided by the amount of residential land (however defined).

This, of course, affects both conventional density and population-weighted density. For conventional density, however, if the proportion of the land that is residential in the entire urban area remains relatively constant over time or between different urban areas, the inclusion of the nonresidential areas will have a limited effect when making comparisons. And this is probably not an unreasonable assumption.

For population-weighted density, on the other hand, the effect of the non-residential land will depend upon its distribution within the urban area relative to residential areas and the the locations of the boundaries of the subareas in relation to this distribution. Morton (2015) implicitly raises this issue in discussing the effect of larger subareas and the locations of the subarea boundaries on densities. A simple example will show what can happen. Suppose an urban area includes a residential area with a uniform density that extends across multiple subareas such as census tracts. One census tract lies totally within this residential area—all of the land is residential. A second census tract having the same total land area falls on the edge of this larger residential area such that half of its area is residential and the other half includes other, nonresidential uses. Residents of both tracts live in a residential area having the same density (by definition). It is therefore reasonable to assume that they might have the same experience of density. But the second tract on the edge of the area would have half the population of the first tract and therefore half the gross population density that would enter into the calculation of population-weighted density.

The effect of such density variation resulting from nonresidential uses will depend on the densities involved. If these are high-density areas, the lower density of the subarea with the nonresidential use could have the effect of reducing the total variance in density across the urban area. This would, of course, lower the population-weighted density. On the other hand, if the area involved had a density lower than the average for the urban area, the artificially lowered density in the subarea with the nonresidential use would increase the overall variance and therefore the population-weighted density. This means, of course, that at least some of these effects will tend to cancel out. But it would take a rather heroic assumption to believe that the relative effects of the decreases and increases in the variance would be similar across areas with distributions of both residential and nonresidential uses and different as, for example, New York and Los Angeles. And the greater land areas of outlying subareas would increases the likelihood of areas of nonresidential use in those areas.

One final question with respect to population-weighted density relates to how one calculates the average density. The first reaction is probably that this is obvious: It is the mean, the sum of the weighted densities divided by the total population. But this is the arithmetic mean. This is not the only measure of central tendency and not even the only method of calculating a mean. Craig (1984, 1985) makes the argument that in comparing densities, relative or proportional differences are more important than absolute differences. A density difference of 100 persons per square mile is more important for an area with a density of 1,000 persons per square mile than for an area with a density of 5,000 or more. For this reason, he proposes that the geometric mean be used for calculating population-weighted densities rather than the arithmetic mean. The population-weighted geometric mean of density is:

$$D_{PG} = \prod d_i^{p_i/P}$$

This becomes easier to understand and compute after taking the logarithms of both sides:

$$\log\left(D_{PG}\right) = \frac{1}{P}\sum p_i \log\left(d_i\right)$$

So the log of the population-weighted geometric mean of density is the population-weighted arithmetic mean of the log of the subarea densities. In other words, this is the population-weighted mean of the magnitudes of density.

The mean is one measure of the central tendency of a distribution. The distribution of densities across subareas in an urban area can often be quite skewed. For skewed distributions, the suggestion is often made that the median may be a more informative measure of central tendency than the mean, which is affected by extreme values. So one might also considered the alternative of the population-weighted median density. This would be the density level of the subarea for which half the population of the urban area resides in subareas with lower densities or in that subarea and half live in areas having higher densities or in that subarea.

The conclusion must be that the population-weighted density for any urban area can take on a wide range of values depending upon the subareas employed and the method of calculation. No single correct value exists for the population-weighted density.

Uses of Population-Weighted Density

Population-weighted density has been employed in a variety of ways. The most basic is as a descriptive measure of density, often in comparison with conventional density. Craig (1984) looked at conventional density and population-weighted density (including the geometric mean) for all of Great Britain using subareas of different sizes. In a subsequent article Craig (1985) compared conventional and population-weighted densities in regions and counties of England and Wales. Dorling and Atkins (1995) provided more detailed and updated results. Zhao and Kaestner (2010) examined at the effect of urban sprawl as measured by density on obesity. They presented population-weighted density in a descriptive table but then used only measures of percentages of conventional density less than various thresholds in their analysis.

Considering studies related to transportation and the environment, Barnes (2001) used both population- and employment-weighted densities as measures of land use in examining the effect on multiple aspects of travel behavior in the U.S. He described his measure of population-weighted density as a “new” measure to which he gave the name “perceived density.” Lee and Lee (2004) showed higher population-weighted density to be associated with major reductions in carbon dioxide greenhouse gas emissions from household travel and from residential energy use in the 125 largest Urbanized Areas in United States. Population-weighted density was calculated using block group data. They stated that “we use this alternative measure because it better captures the population density that typical residents of an urban area experience in their daily lives than do conventional density measures.”

Considering the regional convergence of income in Europe, Castro 2003 considered the relationships to both conventional and population-weighted density. Looking at how economic factors affect crowding, Rapport examined the effects of both total urban factor productivity (2006) and urban consumption amenities (2008) on population-weighted density.

Not surprisingly, population-weighted density has been employed in studies of urban sprawl. For example, Glaeser and Kahn (2004) calculated both population-weighted density and employment-weighted density as measures of sprawl using ZIP code data. Eidlin (2010) argued that population-weighted density was a better measure of urban sprawl and also suggested the use of the ratio of population-weighted density to conventional density as an additional measure. In addressing sprawl, Yglesias (2016) used tract-weighted density change to identify those urban area becoming more dense.

In a policy-making context, Coombes and Raybould (2001) evaluated measures of settlement patterns used for the allocation of funding to local authorities in Britain. They then considered additional measures including population-weighted density, looking at the relationships to improved indices they were proposing. Hanlon, *et al.* (2005) hypothesized that lower density makes the provision of health services more difficult. They used a variant of population-weighted density to look at the relationship between density and maternal health coverage, arguing that uninhabited portions of countries should not factor in to density estimates because these areas were not relevant to the provision of health care.

Some studies have explicitly chosen to use population-weighted density as a measure that would be appropriate for the areas being considered, areas for which the calculation of conventional density would be inappropriate because of the nature of the boundaries. One example is the study by the census (U.S. Bureau of the Census 2012b) of Metropolitan Statistical Areas (MSAs) and micropolitan areas for 2000 and 2010. One chapter looks at the densities of these areas, beginning with the presentation and discussion of conventional densities. As discussed above, this is rather questionable for areas such as MSAs where the boundaries and extent of the areas included can vary so widely and arbitrarily. It is only after looking at the areas with the highest and lowest densities that the report acknowledges this problem associated with the sizes of counties.

To address the problem of using conventional densities with MSAs, they turn to population-weighted densities. It is certainly appropriate to calculate and present population-weighted densities for such areas. But they present population-weighted density more as a means of addressing the problem with conventional densities rather than as a fundamentally different measure of density. They do make the brief statement that "to gain perspective on the densities at which people live, in addition to overall density, the population-weighted density for 2010 was computed," and a brief discussion of the differences is offered. But the extent to which they fail to make completely clear that population-weighted density is a fundamentally different measure comes only a couple of paragraphs later when they make the statement about population-weighted density showing "that the areas with people living at the highest density levels" are largely in California and the Northeast corridor. Here the "where people live" is referring to entire metropolitan areas, as opposed to the use of virtually identical language in presenting the idea of population-weighted density. There they referred to the "densities at which people live" which referred to their census tracts of residence. These not the clearest statements about what population-weighted density is and how it differs from conventional density and is a distinct measure of density.

Loader (2013) was comparing metropolitan areas in Australia and used population-weighted density, saying that it "goes a long way to overcoming having to worry about the boundaries of the 'urban area' of a city." He does briefly describe this population-weighted density as "the residential density in which the 'average resident' lives." But 2 paragraphs later, he describes population-weighted density simply as providing information about the "average density" of a city. Population-weighted density is once again promoted more as a means of overcoming boundary problems without fully clarifying that it is a fundamentally different measure of density.

Population-Weighted Density and Urban Sprawl

Some critics of urban sprawl have turned to population-weighted density as a measure of sprawl. However defined, urban sprawl is generally associated with low-density development. And population-weighted density overcomes one of the paradoxes faced by some sprawl opponents. New York is obviously a very dense area, while the Los Angeles area is often taken as an exemplar of urban sprawl. (See, for example, the article "Is Los Angeles-Style Sprawl Desirable?" Ewing 1997). The difficulty arises because the Los Angeles Urbanized Area has the highest (conventional) density of any area in the United States. The New York area, while very dense, is several places behind. New York City has very high densities, but the variation in density within the broader area is far higher than for Los Angeles. And as demonstrated above, this high density variance results in a much higher population-weighted density for New York as compared to Los Angeles.

Eidlin (2010) argued that population-weighted density is a better measure of urban sprawl, citing the New York-Los Angeles comparison (and describing Los Angeles as an example of high-density sprawl, which some might consider an oxymoron). He suggests that the problem is that suburban areas in Los Angeles "are not dense enough to support traditional urban amenities like frequent and high quality public transit and bustling commercial districts with sidewalk cafes and pedestrian-oriented retail." This is presumably in contrast to the dense core of the New York area, which has these attributes (and evidently those living in the less-dense suburbs in the New York area can be ignored).

Barnes (2001), using population-weighted density in a study of transportation outcomes, also makes the New York-Los Angeles comparison. The argument for the use of population-weighted density and the comparison of the population-weighted densities for the two urban areas have also been made by numerous online commentators (e.g., Bradford 2008a, Florida 2012). Davies (2012a) goes farther, describing the fact that Los Angeles is more dense (in terms of conventional density) as being "counter-intuitive." And Yglesias (2016) said that conventional density "leads to some funny results" while tract-weighted population density "tends to be more accurate."

In a post on his New York Times blog simply titled "Density," Krugman (2013) displayed enthusiasm for population-weighted density. He begins by raising the economic question of whether population growth leads to increases in land prices, relating this to density and changes in density. He then goes on to exclaim, "And we have data!" referring to Florida's article (2012) about the census report providing population-weighted densities for MSAs (discussed earlier). He fails to note that the census has provided conventional densities for Urbanized Areas (for which such measures are appropriate) since 1950. Again citing Florida, Krugman says "the new measure conveys a much better sense of how metros differ" and gives the New York-Los Angeles comparison. He finally uses the trends in population-weighted densities from 2000 to 2010 to address the question posed at the beginning about land prices and population growth. Note that in doing so, he does not address whether population-weighted or conventional density would be more appropriate for addressing the question, apparently assuming that that the population-weighted version is simply a better measure of density. I am not going to question a Nobel Prize-winning economist. But the standard monocentric model that has been the foundation of urban economics includes density, and it is conventional density.

This discussion of population-weighted density and urban sprawl raises the question of whether it might be a better measure of sprawl than conventional density. The introduction to this paper made the point that the two different densities measure of different qualities, and neither should be considered superior. As to which might be a better measure of sprawl, it is depends upon the issue being addressed.

Urban sprawl is discussed in the context of the negative consequences that such development patterns are believed to bring about. It is therefore reasonable to focus the consideration of a choice of a "better" measure of density on the relationships of the measures to those outcomes. For the effects of sprawl on individuals within their immediate residential environments, population-weighted density is obviously the more appropriate measure. The higher the population-weighted density, the greater the proportion of individuals living in denser, more walkable neighborhoods, for example. The effects of sprawl on obesity (Zhao and Kaestner 2010) would likewise be related to the local experience of density.

Other negative outcomes of sprawl, on the other hand, depend upon the total amount of land devoted to urban development. The loss of prime agricultural land and sensitive ecological areas are affected by the conversion of land to urban use and the quantity of such land that is converted. For any given population, the amount of land urban is directly related to the conventional, area-weighted population density. The distribution of the population within the urban area is irrelevant. Likewise, sprawl has been criticized for the higher costs associated with providing urban services, particularly linear services such as roads and utilities, to lower-density development. Again, those costs will depend to a considerable extend on the size of area served, not on the distribution of the population within the larger urban area. So conventional, area-weighted density is a more appropriate measure of sprawl with respect to such negative outcomes.

For other consequences of sprawl, relationships to the two measures of density are more complex. Transportation provides one good example. Even for something as simple as access to public transportation, both measures of density are relevant. For the question of the proportion of the population that can be served within walking distance by a public transportation system limited to a portion of the urban area, population-weighted density is most relevant. The concentration of the population in certain areas makes it easier to provide convenient access to more people. On the other hand, if the question is changed to one of what would be required to provide access to public transit within walking distance for everyone in the urban area, then the size of the area is the only relevant concern and conventional density becomes relevant. In general, the relationship of the distribution of the population (along with the distribution of other destinations) to transportation outcomes is more complex and is not limited to a relationship to one type of density or the other.

Data and Definition of Urban Areas

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 populations 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).

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 2012a).

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 included 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 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.[2]

Analysis of Conventional and Housing-Unit-Weighted Density

Conventional housing-unit density and housing-unit-weighted density have been calculated for the 59 large urban areas for 1950 and 2010. (Housing units are being used in the dataset rather than population, hence the housing-unit densities. The values are in housing units per square mile.) First is the comparison of the two measures of density, followed by the examination of the changes from 1950 to 2010. Next is the consideration of the relationships of the density measures to the variance in density and the sizes of the urban areas at various times. Densities are shown to vary by region. A listing of the conventional and housing-unit-weighted densities and the density variance in 2010 for each of the urban areas is given in the Appendix.

Comparison of the Density Measures

First is the comparison of conventional and housing-unit-weighted densities for 1950 and 2010. Table 1 presents summary statistics for the measures for the two years. The housing-unit-weighted densities are of course higher, which they necessarily must be. Mean housing-unit-weighted density is nearly 3 times larger than conventional

Table 1. Summary Statistics for Conventional and Housing-Unit-Weighted Densities for 59 Large Urban Areas, 1950 and 2010.

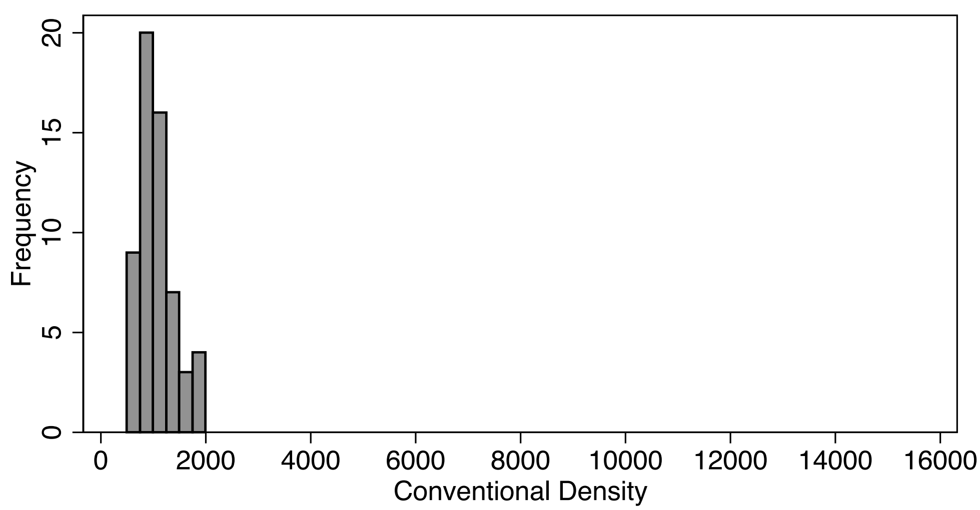
Density Measure	Mean	Standard Deviation	Minimum	Maximum
Densities in 1950				
Conventional	1,268	518	636	2,975
Housing-Unit-Weighted	3,042	2,639	905	18,341
Densities in 2010				
Conventional	1,080	330	546	1,928
Housing-Unit-Weighted	2,329	1,924	732	14,889

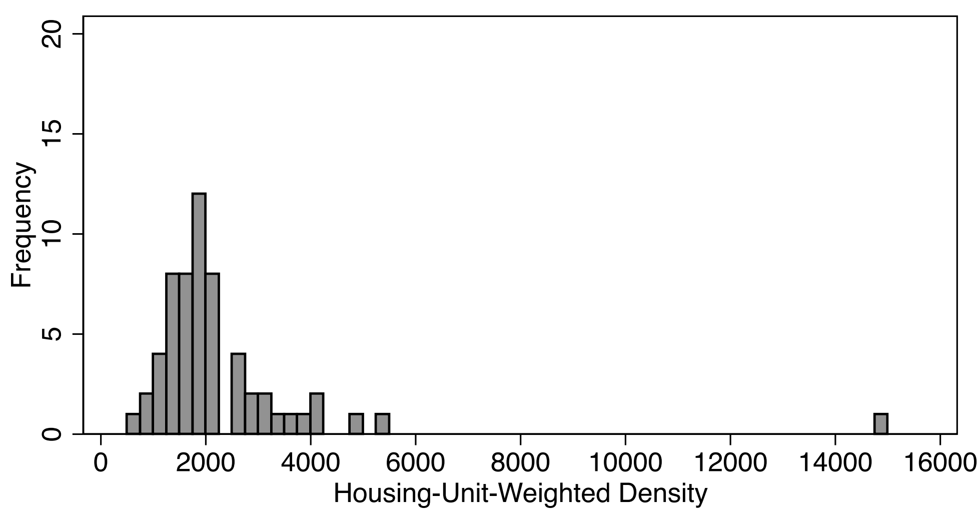
[2] More detail on the construction of the dataset and the delineation of the urban areas is provided in Ottensmann (2014).

density in 1950 and over twice as large in 2010. More interesting are the minimum and maximum values. The smallest weighted densities in both 1950 and 2010 are not that much greater than the minimum conventional densities. The maximum values, on the other hand, are much, much higher than the conventional densities, about 18,000 versus 3,000 in 1950 and 15,000 versus 2,000 in 2010. Not surprisingly, the highest weighted density is for the New York area in both years.

The very wide range of the weighted densities as compared with the conventional densities suggests making comparisons of the distributions of the two measures. Figure 1 shows the histograms of the two distributions for 2010 using the

Figure 1. Histograms of Conventional Density and Housing-Unit-Weighted Density, 2010.



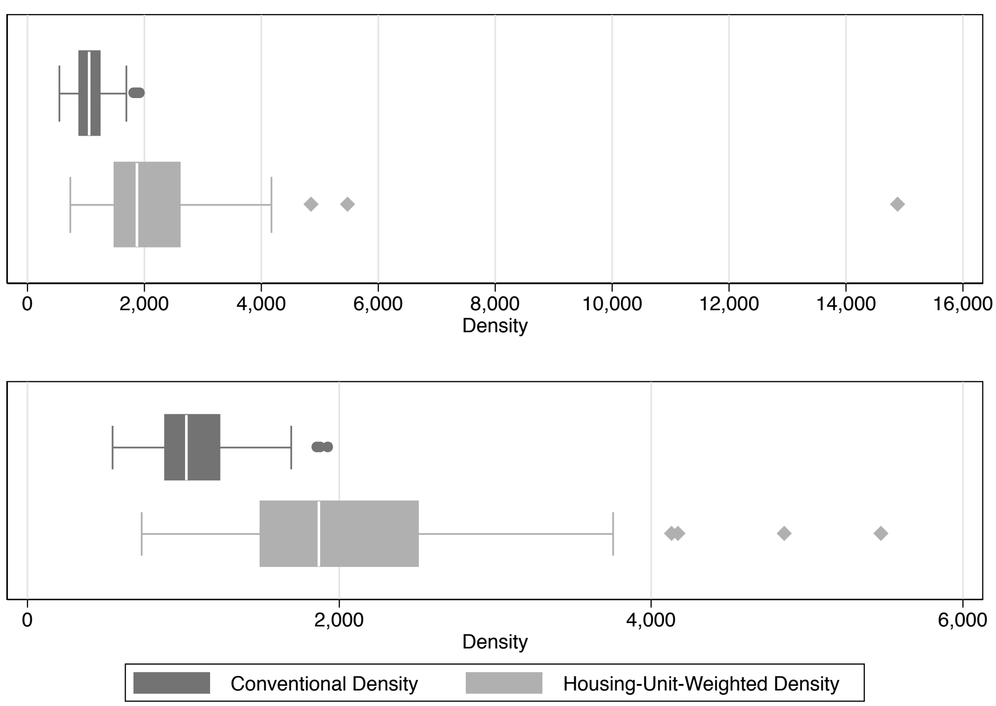


same scale for each. For conventional density, all of the areas are tightly clustered below 2,000 units per square mile. Weighted densities have a much more extended tail, going up to over 5,000 units per square mile for the second largest, with the maximum value for New York at nearly 15,000 units. This shows clearly the extent to which New York is a major outlier in terms of weighted density while at the same time other urban areas also have high values relative to conventional density.

The histograms clearly showed the presence of the outliers, especially New York. They were less effective in allowing the comparison of the distribution of the majority of the areas with respect to the 2 measures. For this, box plots of the two density measures are shown in Figure 2. Two sets of the conventional- and weighted-density box plots are presented. The top set shows the full range, including the maximum value for New York. The lower-set limits \the range to 6,000 units per square mile, excluding the very high value for New York and expanding the rest of the plots to better allow comparisons of the remainder of the distributions.

The middle of the distribution of the housing-unit-weighted densities—the range between the first and third quartiles—falls above all but the highest areas in terms of conventional density. And the top quartile extends yet higher. Conversely, most of the

Figure 2. Box Plots of Conventional Density and Housing-Unit-Weighted Density, 2010.



conventional densities fall into the first quartile of weighted densities. So the overlap of the two distributions is quite limited.

Comparison of the areas with the highest densities on the two measures provides additional insight into both the differences and similarities. Table 2 lists the 10 areas with the highest densities in 2010, the top portion in terms of conventional densities and the lower part with respect to housing-unit-weighted densities. In addition to the rank and density for the measure used to order each table, the listings also give the density and ranking for the other density measure.

Looking first at the areas with the highest conventional densities in Table 2. Los Angeles has the highest housing-unit density among all of the areas. New York is somewhat lower, in fourth place. Note that this is the same as for population densities for all census Urbanized Areas in 2010—Los Angeles is the most dense and New York is fourth. Some of the areas on this list of the most dense areas are places many would not have thought of as being especially dense, including Las Vegas, San Diego, Portland, and Denver. Indeed, 3 of these 4 areas—all but San Diego—are not ranked in the top 10 with respect fo housing-unit-weighted density—though none falls to lower than 16th position.

Now consider the densest areas with respect to housing-unit-weighted density shown the lower portion of Table 2. New York has by far the highest weighted density. Los Angeles' ranking has dropped, but only to fourth position. It still has one of the higher housing-unit-weighted densities. Additions to this top-10 ranking that were not on the list as having the highest conventional densities are Chicago, Philadelphia, and Boston. These are areas that many might have considered among the more dense urban areas in the country. And while they are dense in terms of weighted density, this is not so much the case with respect to their conventional densities. Chicago was ranked 11th for conventional density, so it almost made the top-10 list. But of the 59 large urban areas in the dataset, Philadelphia was only 22nd in terms of conventional density and Boston was way down that list in the 41st position. This clearly shows how in some cases conventional density and weighted density can give a very different senses of urban area densities.

The lists of areas with the lowest densities showed less interesting variation between conventional and housing-unit-weighted densities. For the areas with the lowest densities, all but one of the 10 least dense in terms of housing-unit-weighted density in 2010 were in the South, with Indianapolis in tenth place. And all but one were also either in the 10 least dense areas in terms of conventional density or very close. The 5 least dense areas with respect to both types of density were Greenville-Spartanburg, Knoxville, Greensboro—Winston-Salem—High Point, Charlotte, and Raleigh-Durham.

Table 2. The Ten Most Dense Urban Areas Ranked by Conventional Density and by Housing-Unit-Weighted Density, 2010.

Area	Conventional Density		Housing-Unit-Weighted Density	
	Value	Rank	Value	Rank
Los Angeles	1,928	1	4,173	4
San Francisco-Oakland-San Jose	1,880	2	5,476	2
Las Vegas	1,859	3	2,888	11
New York	1,822	4	14,889	1
New Orleans	1,692	5	3,209	9
Miami-Fort Lauderdale-West Palm Beach	1,639	6	3,756	6
San Diego	1,507	7	3,095	10
Portland	1,435	8	2,508	16
Washington-Baltimore	1,382	9	3,731	7
Denver	1,357	10	2,606	15

Area	Housing-Unit-Weighted Density		Conventional Density	
	Value	Rank	Value	Rank
New York	14,889	1	1,822	4
San Francisco-Oakland-San Jose	5,476	2	1,880	2
Chicago	4,856	3	1,342	11
Los Angeles	4,173	4	1,928	1
Philadelphia	4,133	5	1,134	22
Miami-Ft Lauderdale-W Palm Beach	3,756	6	1,639	6
Washington-Baltimore	3,731	7	1,382	9
Boston	3,445	8	909	41
New Orleans	3,209	9	1,692	5
San Diego	3,095	10	1,507	7

Changes in Densities

Next is the consideration of the changes in the two density measures from 1950 to 2010. Two measures of change have been calculated, the difference between the 1950 and 2010 densities and the percentage changes over the period. Table 3 presents summary statistics for these measures of change. On average, both types of densities declined over the period, with weighted density showing a somewhat greater drop of 12 percent versus 5 percent for conventional density. But more significant were the tremendous ranges in the changes in both types of density. The changes in conventional density ranged from a decline of over 1,000 units per square mile to an increase that was just as large. The largest drop in weighted density was over 3,700 while another area saw the weighted density increase by nearly 1,700 units per square mile. Obviously no consistent pattern of change occurred for either type of density. The smallest and largest percentage changes were more extreme because some of these changes started from bases of very low densities in 1950.

Relationships Among Density, Variance, and Size of Area

The measures of density are related to one another, to the variance of density in the urban area, and to the size of the area. These relationships are examined here. Starting with the relationship of housing-unit-weighted density to conventional density, Figure 3 is the scatter plot of weighted versus conventional density. Weighted density is reasonably closely related to conventional density, increasing in a generally linear

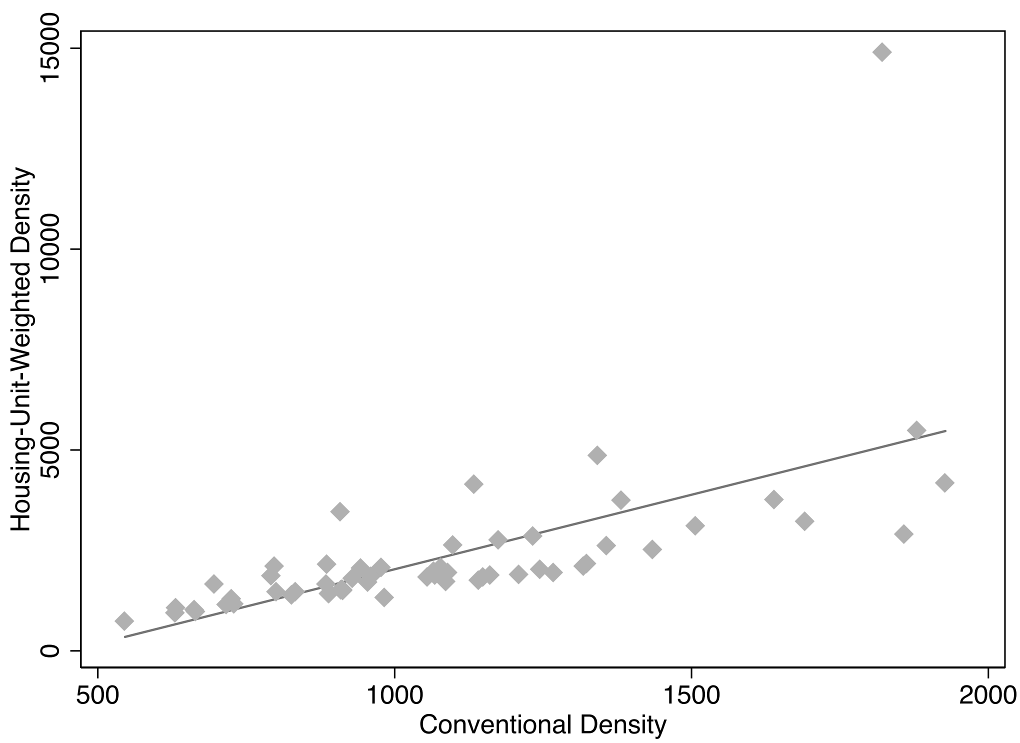
Table 3. Summary Statistics for Changes in Conventional and Housing-Unit-Weighted Densities, 1950 to 2010.

Density Change Measure	Mean	Standard Deviation	Minimum	Maximum
Change in Density, 1950-2010				
Conventional	-187	479	-1,153	1,154
Housing-Unit-Weighted	-713	1,166	-3,729	1,770
Percentage Change in Density, 1950-2010				
Conventional	-5.3	40.8	-55.2	163.8
Housing-Unit-Weighted	-12.1	39.1	-61.8	158.2

fashion, with one major exception. New York's position as the outlier with the extremely high housing-unit-weighted density in comparison to its conventional density is very obvious.

Table 4 presents the correlations among both measures of density, the variance of density across the census tracts, and various measures of the size of the urban area. As expected from the scatter plot, the correlation between conventional and weighted density is high, 0.85, and clearly highly statistically significant. Both measures of density were likewise highly correlated with the variance in density. This is not surprising. The threshold for the inclusion of tracts in an urban areas is the same for all areas, so all areas will have similar minimum tract densities. The higher the overall densities, the greater the range and therefore the larger the variance. While the correlation of the variance with conventional density was large, 0.75, the correlation with weighted density was extremely high at 0.97. This relationship to the variance is not surprising given that, as shown above, weighted density is equal to conventional density plus the variance divided by the conventional density.

Figure 3. Scatter Plot of Housing-Unit-Weighted Density versus Conventional Density, 2010.



Now to the relationships of the densities to the size of the urban area. First is the consideration of the correlations to the size of the area contemporaneously with the density measurement, 1950 sizes with 1950 densities and 2010 sizes with 2010 densities. In keeping with the use of housing units, size is measured in terms of the number of housing units in the urban areas. Once again there are strong positive relationships. Densities clearly increase with the size of the urban area. The relationships are especially strong for weighted densities, with correlations of 0.90 and 0.87 in 1950 and 2010 as opposed to somewhat smaller correlations of 0,67 and 0.58 for conventional density.

This relationship of density to size raises the question of size at what point in time. Urban areas develop and their patterns, including densities, are established over long periods of time. Cities were far more dense in the 19th century when they were dependent upon walking and public transit as the primary modes for movement. The development and widespread use of the automobile provided more flexibility and enabled settlement further from the centers at lower densities. This suggests that urban areas that grew to a larger size before the automobile came into general use might be more dense than areas that have seen more of their development in the automobile era.

The sizes of the urban areas in 1910 are considered to examine this hypothesis. That year was chosen for 2 reasons: This is prior to the period when automobile use became common enough to substantially contribute to urban mobility. In addition, 1910 was the first year in which the census delineated and reported populations for metropolitan areas that included the suburbs along with the major cities.

The Census Bureau first defined Metropolitan Districts around cities of 100,000 or more in 1910 and reported populations and other data for these areas. (The Metropolitan Districts were the forerunners of the current Metropolitan Statistical Areas but were comprised of incorporated municipalities and minor civil divisions, not counties.) The Metropolitan District populations in 1910 are being used as the measure of urban area size in the early twentieth century for those areas for which they were delineated. For cities with populations less than 100,000, the city populations are used. This seems reasonable because those areas were likely to have limited suburban population outside the city. And Las Vegas is not included as it was not incorporated as a city until 1911 so there was no census report of population for 1910 (U.S. Bureau of the Census 1913).

For conventional density in 1950, the correlation with 1910 population was just as high as the correlation with 1950 housing units. 0.69 versus 0.67. This supports the hypothesis that the development of the urban area in the pre-automobile period had a significant effect of later density levels. But by 2010, the correlation with 1910 population had dropped to 0.36 as compared with the correlation of 0.69 with housing units in 2010. This could reasonably be attributed to the longer time period and the greater amounts of subsequent development affecting urban patterns and densities.

Table 4. Correlations Among Densities, Density Variance, and Sizes of Urban Areas, 1950 and 2010.

	Conventional Density	Housing-Unit-Weighted Density	Density Variance	Housing Units in Current Year	Population in 1910
Correlations for 1950					
Conventional Density	1.000				
Housing-Unit-Weighted Density	0.847 ***	1.000			
Density Variance	0.739 ***	0.971 ***	1.000		
Housing Units in 1950	0.672 ***	0.895 ***	0.930 ***		
Population in 1910	0.685 ***	0.913 ***	0.942 ***	0.961 ***	1.000
Correlations for 2010					
Conventional Density	1.000				
Housing-Unit-Weighted Density	0.635 ***	1.000			
Density Variance	0.556 ***	0.987 ***	1.000		
Housing Units in 2010	0.581 ***	0.866 ***	0.847 ***	1.000	
Population in 1910	0.359 **	0.906 ***	0.903 ***	0.799 ***	1.000

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

A quite different picture is seen when looking at housing-unit-weighted densities. For the 1950 densities, the correlations with 1910 populations are about the same (actually a bit higher) than the correlations with contemporaneous size. The difference was 0.90 versus 0.91. But the more interesting result is for 2010 weighted density, with a correlation of 0.91 with 1910 population and 0.87 for 2010 housing units. The correlation of housing-unit-weighted density with the size of the urban area has not declined over the 100-year period since 1910. Weighted density in 2010 is as strongly related to the size of the area a century earlier as it was to the current size in 2010. This would seem to suggest that the some of densest portions of urban areas which are contributing to higher weighted densities were established by 1910 before the widespread use of the automobile, and their effects on weighted densities are continuing up through the present.

Differences in Density by Region

The final question relates to how the 2 types of densities vary by region of the country. Table 5 gives the mean values for the 2 measures of density in 1950 and 2010 for the urban areas in each of the census regions. In what may seem surprising to some, areas in the West had by far the highest mean conventional density in 2010. Areas in the South had the average lowest density, but this was not that much below the means for the Northeast and Midwest. This is in contrast to 1950 in which the urban areas in the Northeast had by far the highest conventional density.[3]

Perhaps not surprisingly, the patterns for average housing-unit-weighted densities was very different. Urban areas in the Northeast had by far the highest average weighted densities of over 6,000 units per square mile in 1950 and nearly 3,900 in 2010. This is the location of many of the larger and older urban areas that experienced rapid growth in the nineteen and early twentieth century. However, the mean value in

Table 5. Mean Conventional Density and Housing-Unit-Weighted Density by Region, 1950 and 2010.

Region	Mean Density 1950		Mean Density 2010	
	Conventional Density	Housing-Unit-Weighted Density	Conventional Density	Housing-Unit-Weighted Density
Northeast	1,814	6,033	1,009	3,875
Midwest	1,536	3,500	1,020	2,087
South	1,023	1,968	955	1,687
West	1,112	2,612	1,425	2,715
Significance	$p < 0.001$	$p < 0.001$	$p < 0.001$	$p < 0.05$

[3] For more detailed analysis of urban area densities and changes over time, see Ottensmann (2016).

2010 is strongly affected by the extremely high value for New York. Excluding New York reduces the mean for the remaining urban areas in the Northeast to just under 2,500. The urban areas in the West, much newer and younger, went from nearly the lowest mean density in 1950 to an average of 2,700 units per square mile in 2010, second to those in the Northeast with New York included, but higher than the Northeastern areas without New York.

Once again, areas in the South were at the bottom at under 1,700. One note on the mean densities for the South: The Washington-Baltimore urban area is included in the South census region though it is arguably more similar to the other large urban areas in the Northeast corridor. It had, along with Miami-Fort Lauderdale-West Palm Beach and New Orleans by far the highest densities, both conventional and weighted, in the South. Excluding Washington-Baltimore would have resulted in somewhat lower mean densities for the South. Excluding all 3 of those outliers would have resulted in a substantially lower mean.

Conclusions

Population-weighted density is a distinct, alternative measure of density. As demonstrated, population-weighted density is equal to conventional density plus the variance of density across the subareas used for its calculation divided by the conventional density, which is simply total population divided by total area. Conventional density is also equivalent to area-weighted density calculated across any set of subareas.

Population-weighted density is a measure of density that reflects the average density as experienced by the residents of the urban area within their subareas. It captures an aspect of density that differs from conventional density. But neither density can be seen as being inherently superior. As measures of urban sprawl, population-weighted density may be more appropriate in considering negative effects of low densities and sprawl on people within their immediate neighborhoods. However, conventional, area-weighted density is more relevant to consideration of the effects of sprawl in causing the loss of land and increased costs of providing services associated with the greater land areas to be served.

The size and configuration of the subareas used for the calculation of population-weighted density has major implications for the values obtained and for their interpretation. Many of those using population-weighted density have devoted little or no attention to these critical issues.

Population-weighted densities must be greater than conventional densities with the increase associated with the level of variation in densities across the subareas. Because of this effect of density variation, population-weighted densities are also more highly skewed towards larger values, with the population-weighted density of the New

York area being an extreme example. This effect of the density variance makes it possible for areas having very modest conventional densities to have relatively high population-weighted densities. An example is the Boston urban area, which is was forty-first in conventional density among the 59 urban areas consider but has the eighth highest housing-unit-weighted density.

While both types of density are positively related to the sizes of urban areas, housing-unit-weighted density is more highly correlated with size than conventional density. This holds not only for the correlations of weighted density with the current size of an urban area but with the size in earlier years, going back to 1910. How large the urban area had grown by that year continues to be very highly correlated with population-weighted densities all the way up to the present.

The effect of the timing of urban growth on the two types of densities can be seen in regional differences as well. In 1950, mean population-weighted density was highest for the older urban areas in the Northeast which had experienced considerable growth in the nineteenth and early twentieth centuries. But by 2010, mean conventional density was greatest for the urban areas in the West which have experienced more of their growth in recent decades.

To reemphasize the point made in the introduction: Population-weighted density captures a different aspect of density compared with conventional density. It is not a better measure of density. It has advantages and disadvantages compared to conventional density. As such, it can be a useful alternative in some contexts.

References

Avent, Ryan. 2011. *The Gated City*. Kindle Single. Amazon.

Barnes, Gary. 2001. Population and employment density and travel behavior in large U.S. cities. Final Report 2001-24. St. Paul: Minnesota Department of Transportation. Downloaded from https://www.lrrb.org/pdf/200124.pdf on January 1, 2018.

Bradford, Chris. 2008a. Perceived density. *Austin Contrarian* (website). Accessed from http://austinzoning.typepad.com/austincontrarian/2008/03/perceived-densi.html on February 7, 2017.

Bradford, Chris. 2008b. Density calculations for U.S. urbanized areas, weighted by census tract. *Austin Contrarian* (website). Accessed from http://austinzoning.typepad.com/austincontrarian/2008/03/weighted-densit.html on February 7, 2017.

Bradford, Chris. 2008c. More on weighted density for the crackpot blogging stats geeks. *Austin Contrarian* (website). Accessed from http://www.austincontrarian.com/austincontrarian/2008/03/more-for-the-cr.html on February 7, 2017.

Castro, José Villaverde. 2003. Regional convergence, polarisation and mobility in the European Union, 1980-1996. *Journal of European Integration* 25, 1: 73-86.

Coombes, Mike, and Simon Raybould. 2001. Public policy and population distribution: Developing appropriate indicators of settlement patterns. *Environment and Planning C: Politics and Space* 19, 2 (April): 223-248.

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.

Davies, Alan. 2012a. Is Los Angeles really the densest city in the US? *The Urbanist* (website). Accessed from https://blogs.crikey.com.au/theurbanist/2012/10/17/is-los-angeles-really-the-densest-city-in-the-us/ on February 8, 2017.

Davies, Alan. 2012b. Are Australian cities sprawling at ever-lower densities? *The Urbanist* (website). Accessed from https://blogs.crikey.com.au/theurbanist/2013/11/06/are-australian-cities-sprawling-at-ever-lower-densities/ on February 8, 2017.

Davis, Kenneth. 2015. Weighted density and city size: Who knew Milwaukee is so dense. *Geovisualist* (website). Access from https://geovisualist.com/2015/05/18/who-knew-milwaukee-is-so-dense/ on February 8, 2017.

Dorling, Daniel, and David Atkins. 1995. *Population density, change and concentration in Great Britain 1971, 1981 and 1991*. Studies on Medical and Population Subjects, No. 58. London: HMSO. Downloaded from http://www.dannydorling.org/wp-content/files/dannydorling_publication_id1449.pdf on February 6, 2018.

Duke, Martin H. 2012. Population-weighted density: How Seattle stacks up. *Seattle Transit Blog* (website). Accessed from https://www.seattletransitblog.com/2012/10/27/population-weighted-density-how-seattle-stacks-up/ on February 8, 2017.

Eidlin, Eric. 2010. What density doesn't tell us about sprawl. *Access* 37 (Fall): 2-9.

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

Florida, Richard. 2012. America's truly densest metros. *CityLab* (website). Accessed from http://www.citylab.com/housing/2012/10/americas-truly-densest-metros/3450/ on February 8, 2017.

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.

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

Hanlon, Michael, Roy Burstein, Samuel H. Masters, and Raymond Zhang. 2012. Exploring the relationship between population density and maternal health coverage. *BMC Health Services Research* 12, 1: 416

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.

Krugman, Paul. 2013. Density. *The Conscience of a Liberal. New York Times* (website). Accessed from https://krugman.blogs.nytimes.com/2013/04/16/density/?_r=1 on February 8, 2017.

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.

Lewontin, R.C. and Levins, R., 1989. On the characterization of density and resource availability. *The American Naturalist* 134, 4 (October): 513-524.

Loader, Chris. 2013. Are Australian cities becoming denser? *Charting Transport* (website). Accessed from https://chartingtransport.com/tag/population-weighted-density/ on February 8, 2017.

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.

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. 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, 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., 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.

Rappaport, Jordan. 2006. A productivity model of city crowdedness. *Research Working Papers* 06-06. Federal Reserve Bank of Kansas City, Economic Research Department. Downloaded from https://www.kansascityfed.org/publicat/reswkpap/pdf/rwp06-06.pdf on December 13, 2017.

Rappaport, Jordan. 2008. Consumption amenities and city population density. *Regional Science and Urban Economics* 38, 6 (November): 533-552.

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

U.S. Bureau of the Census. 1913. *Thirteenth Census of the United States Taken in the Year 1910*. Vol. 1. *Population 1910. General Report and Analysis.* Washington DC: Government Printing Office.

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. 2012a. 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. Census Bureau, 2012b. *Patterns of Metropolitan and Micropolitan Population Change: 2000 to 2010*. 2010 Census Special Report, C2010SR-01. Washington DC: U.S. Government Printing Office.

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.
Yglesias, Matthew. 2016. Seattle and Portland are showing us how cities can grow without sprawl. *Vox* (website). Accessed from http://www.vox.com/2016/5/19/11713668/growth-and-sprawl on February 8, 2017.
Zhao, Zhenxiang, and Robert Kaestner. 2010. Effects of urban sprawl on obesity. *Journal of Health Economics* 29, 6 (December): 779-787.

Appendix

Conventional Density, Housing-Unit-Weighted Density, and Density Variance for 59 Large Urban Areas, 2010

Urban Area	Conventional Density	Housing-Unit-Weighted Density	Density Variance (thousands)
Albany-Schenectady-Troy	798	2,093	1,034
Albuquerque	1,268	1,940	852
Atlanta	725	1,283	405
Austin	929	1,796	806
Birmingham	730	1,151	307
Boston	909	3,445	2,305
Buffalo	1,099	2,621	1,673
Charlotte	663	1,005	227
Chicago	1,342	4,856	4,715
Cincinnati	886	1,654	680
Cleveland-Akron	978	2,069	1,067
Columbus	1,090	1,939	925
Dallas-Fort Worth	1,066	1,941	932
Dayton	912	1,528	562
Denver	1,357	2,606	1,695
Detroit	1,245	2,024	969
El Paso	1,087	1,712	679
Fresno	1,210	1,881	812

Grand Rapids	833	1,460	523
Greensboro--Winston-Salem--High Point	665	972	204
Greenville-Spartanburg	546	732	102
Harrisburg-York	697	1,645	660
Hartford	792	1,858	844
Houston	1,066	1,967	960
Indianapolis	827	1,374	452
Jacksonville	983	1,308	320
Kansas City	913	1,495	531
Knoxville	631	932	190
Las Vegas	1,859	2,888	1,914
Los Angeles	1,928	4,173	4,328
Louisville	956	1,689	701
Memphis	889	1,414	466
Miami-Ft Lauderdale-W Palm Beach	1,639	3,756	3,470
Milwaukee	1,175	2,742	1,841
Minneapolis-St Paul	1,079	2,066	1,065
Nashville	717	1,135	302
New Orleans	1,692	3,209	2,567
New York	1,822	14,889	23,805
Norfolk-Virginia Beach	1,162	1,872	825
Oklahoma City	913	1,502	538
Orlando	891	1,429	480
Philadelphia	1,134	4,133	3,402
Phoenix	1,319	2,090	1,017
Pittsburgh	886	2,139	1,110
Portland	1,435	2,508	1,540
Raleigh-Durham	632	1,060	271
Richmond	801	1,462	529
Rochester	943	2,049	1,043
Sacramento	1,324	2,151	1,095
Salt Lake City-Ogden-Provo	1,055	1,819	806
San Antonio	1,141	1,738	681

San Diego	1,507	3,095	2,394
San Francisco-Oakland-San Jose	1,880	5,476	6,762
Seattle-Tacoma	1,233	2,848	1,991
St Louis	961	1,835	840
Tampa-St Petersburg	1,069	1,866	852
Tucson	1,149	1,824	775
Tulsa	910	1,513	549
Washington-Baltimore	1,382	3,731	3,247