Urban Patterns Dataset Description

John R. Ottensmann
Indiana University-Purdue University Indianapolis
john.ottensmann@gmail.com
urbanpatternsblog.wordpress.com
October 2014

Objective

This documents the *Urban Patterns* tract dataset for the study of urban patterns and changes in urban patterns from 1940 to 2010 for large metropolitan areas in the United States.

Metropolitan Areas

The metropolitan areas to be studied are identified and the maximum extent of the areas to be included are established by the boundaries of Combined Statistical Areas (CSAs) and Metropolitan Statistical Areas (MSAs). The objective is to include both the built-up urban cores of the areas and the less-densely settled exurban territory that is linked to the urban cores of the metropolitan areas.

MSAs consist of the county or counties containing an Urbanized Area along with any additional counties having at least 25 percent commuting to or from the central counties. CSAs are combinations of Core-Based Statistical Areas (CBSAs—MSAs and Micropolitan Statistical Areas) having employment interchange of at least 15 percent (U.S. Office of Management and Budget 2010).

CSAs were chosen for those areas in which the MSAs were combined into CSAs as the CSA is a broader, more inclusive delineation that captures the full extent of sprawling metropolitan areas and better represents the areas to be considered. Three examples illustrate the benefits of using the CSAs; The New York CSA includes the suburbs of New York City in Connecticut; the New York MSA does not. The San Jose-San Francisco CSA includes the areas associated with the two MSAs; the built-up areas of these two areas are part of a single, continuous built-up area with no obvious boundary separating them. The Raleigh-Durham CSA includes the areas of two separate MSAs that form what is generally recognized as a single multi-centered metropolitan area.

Some further evidence supporting CSAs as integrated urban areas includes the following. The MSAs combined to form these CSAs are linked by commuter rail:

New York
Los Angeles
Washington-Baltimore
San Francisco-Oakland-San Jose
Boston-Providence

Seattle-Tacoma
Salt Lake City-Ogden-Provo

The following CSAs with multiple centers formed by merging those MSAs are served by a single primary commercial airport:

Raleigh-Durham (Raleigh-Durham International Airport)
Greensboro—Winston-Salem—High Point (Piedmont Triad International Airport)
Greenville-Spartanburg (Greenville-Spartanburg International Airport)

Seattle-Tacoma International Airport is the only primary commercial airport serving Tacoma. Seattle does have one other minor airport, Boeing Field.

The MSA and CSA delineations released in 2013, using the 2010 census figures, are used in this study (U.S. Bureau of the Census 2013). The areas were ranked by 2010 population for the CSA if the area was included in a CSA, otherwise by the MSA population (using 2010 county populations aggregated to the MSA and CSA areas; U.S. Bureau of the Census 2014b). Fifty-nine areas had populations greater than one million. The largest was the New York CSA with a population of 23 million. The smallest was the El Paso CSA. These areas were selected for this study. The areas are listed in the Appendix.

The Census cautions against ranking CSAs with MSAs.[1] I respectfully disagree. Starting with the largest MSA in a group of contiguous CBSAs, a CBSA is added to the MSA (forming a CSA) if the commuting interchange exceeds 15 percent. Further CBSAs are added if the commuting interchange with any of the CBSAs in the CSA exceeds that threshold. So a CSA consists of the large core MSA and any CBSAs that have been added. It seems entirely reasonable to consider a large MSA that is not part of a CSA to be equivalent to a CSA where no additional CBSAs qualified for addition. (Of course a name other than *Combined* Statistical Area would be required for such areas.)

While one might reasonably debate the choice of CSAs over MSAs for studying the patterns of entire metropolitan areas, using areas smaller than MSAs does not seem to make a lot of sense. Some studies have used Metropolitan Divisions (or earlier, Primary Metropolitan Statistical Areas), which are subdivisions of very large MSAs (e.g., Ewing, Pendall, and Chen 2002; Ewing and Himidi 2014; Lopez and Hynes 2003; Lopez 2013; Paulsen 2012). The use of some Metropolitan Divisions or PMSAs may not be unreasonable, such as Miami, Fort Lauderdale, and West Palm Beach or San Francisco and Oakland (the example cited by Lopez 2013 to justify the use of the Metropolitan Divisions). But considerably more of the Metropolitan Divisions are portions of their

[1] From the Census website: “Combined Statistical Areas (CSAs) consist of two or more adjacent CBSAs that have substantial employment interchange. The CBSAs that combine to create a CSA retain separate identities within the larger CSA. Because CSAs represent groupings of metropolitan and/or micropolitan statistical areas, they should not be ranked or compared with individual metropolitan and micropolitan statistical areas.” (U.S. Bureau of the Census 2014a)

MSAs that virtually everyone one recognizes as suburban areas within or primarily within the same Urbanized Area as the largest city in the MSA. Two examples of such Metropolitan Divisions are Nassau and Suffolk Counties in the New York MSA and Montgomery, Bucks, and Chester Counties in the Philadelphia MSA. Note that these areas are homes to two of the Levittowns built after World War II and are obviously suburban. The problem of using such Metropolitan Divisions is twofold. First, most measures will show relatively low densities and high levels of sprawl for these suburban divisions because they are, in fact, suburban and lack the traditional older urban core. Second, because such areas are excluded from the core Metropolitan Divisions such as New York and Philadelphia, measures for those core areas will show higher densities and lower levels of sprawl because some of the more sprawling suburban areas have been excluded.

Multiple Centers

Some metropolitan areas have multiple major centers such as Dallas-Fort Worth. Identifying these centers is important for two reasons. First, in years in which the urban areas associated with the centers are not contiguous, both urban areas should be included as urban areas in the dataset. This is in contrast to smaller separate areas that meet the urban density criterion but are not contiguous with a major urban areas and are therefore not considered to be part of the major urban area. Multiple major centers also need to be identified for calculation of the measures of centralization for which the measurement of distance should be to the appropriate center.

Major urban centers are identified by considering the population of potential Urbanized Areas as a percentage of the population of the largest Urbanized Area within the CSA or MSA. Each smaller Urbanized Area in 2010 was considered as a candidate. For example, the population of the Durham Urbanized Area in 2010 was 39 percent of the population of the Raleigh area. However, numbers of the potential multiple major centers were part of a single combined Urbanized Area in 2010. In those cases, the Urbanized Area population in the last year the smaller Urbanized Area was separate was compared with the population of the largest Urbanized Area in that year. For Dallas and Fort Worth, Fort Worth was a separate Urbanized Area for the last time in 1970, when its Urbanized Area population was 51 percent of the population of the Dallas area.

To identify a reasonable cutoff for including areas as major urban centers, the percentages were plotted in decreasing order for all of the potential centers. Going down the list, a group of 3 potential centers had Urbanized Area populations between 28.5 and 32 percent of the populations of the largest Urbanized Areas: Akron had 32 percent of Cleveland; Tacoma and Providence had 28.5 percent of Seattle and Boston, respectively. There was then a gap, with the next two being Kissimmee, Florida, 21 percent of Orlando, and Corcord, North Carolina, 17 percent of Charlotte. Akron, Tacoma, and Providence have long been recognized as major centers within their areas. Kissimee and Concord have not. So this seemed to be a reasonable point to make the cut for identifying major centers.

Four cases had to be considered as exceptions. Virginia Beach has never been an independent Urbanized Area separate from Norfolk. The original Norfolk Urbanized Area did not incorporate the area of Virginia Beach through 1980, and Virginia Beach was not an Urbanized Area. In 1990, the Norfolk Urbanized Area was expanded to included Virginia Beach. Since Virginia Beach is now the largest city in the Urbanized Area, while Norfolk has been the historic center, both should reasonably be considered major urban centers.

Troy, New York, was part of the Albany Urbanized Area in the first year those areas were designated, 1950. It has generally been considered as one of the centers of the Albany-Schenectady-Troy area. In the urban area delineation for this research, the Troy urban area was last distinct from the Albany area in 1950, when it had 73 percent of the number of housing units in the larger area. So it was reasonable to include Troy as a center.

Finally, Minneapolis and St. Paul have never been separate Urbanized Areas, nor have San Francisco and Oakland. The areas classified as urban in this study have never been separate for these pairs of cities. These cities have always been considered to be multiple major centers within their metropolitan areas and will likewise be considered to be so for this study for purposes of calculating measures of centralization.

The list of metropolitan areas with multiple centers is as follows:

- Albany-Schenectady-Troy
- Boston-Providence
- Cleveland-Akron
- Dallas-Fort Worth
- Greensboro—Winston-Salem—High Point
- Greenville-Spartanburg
- Harrisburg-York
- Miami-Fort Lauderdale-West Palm Beach
- Minneapolis-St Paul
- Norfolk-Virginia Beach
- Raleigh-Durham
- Salt Lake City-Ogden-Provo
- San Francisco-Oakland-San Jose
- Seattle-Tacoma
- Tampa-St Petersburg
- Washington-Baltimore

Data Sources

The primary data source used for classifying census tracts as urban and exurban is the Neighborhood Change Database for 1970 to 2000 developed by the Urban Institute and Geolytics (2003). This dataset contains census tract data from each census normalized to the 2000 census tract boundaries. In addition, data for population and housing units

from the 2010 census was aggregated to the 2000 tract boundaries to extend forward the time period covered (U.S. Bureau of the Census 2012).

The number of housing units in each census tract was the basic data element used for two reasons: First, housing units are fixed and more appropriate for the measurement of urban patterns than population, as noted by Galster *et al.* (2001) and others. Population and population densities have often been employed in describing and measuring urban patterns, but the use of population raises problems, especially when considering changes over time. The population per housing unit can change, so changes in population will not necessary be associated with changes in the distribution of housing units. Rates of occupancy and vacancy can likewise change.

The second reason for using housing units is that this allows the census data on the age of housing units to be used for estimating numbers of housing units before 1970. The census long form, including for 1970, has included the question about the year the housing unit's structure was built. This is reported as the number of units by age interval. It is therefore possible to determine the number of units in 1970 that were built before 1940, 1950, and 1960. These are then taken as estimates of the numbers of housing units in each census tract in each of those years. Of course, any units lost through demolition or merger will not be included in those figures. Likewise, units may have been added through subdivision and the conversion of nonresidential uses to residential use. Error can also be introduced by the misreporting of the age of the structure. But housing units are relatively permanent, so that these values may serve as reasonable estimates of the number of housing units.

Two approaches can allow some assessment of the errors associated with these estimates. The total number of housing units is available by county from the earlier censuses. For those counties for which all census tracts are included in the dataset, it is possible to take the totals of the estimated housing units for the tracts in the county and compare these to the numbers of units reported in the census. Another approach for assessing the error in this procedure would be to use the numbers of housing units by year built reported in the 2000 census, use those values to estimate the numbers of units for 1970 through 1990, and compare those estimates from the "known" values from the Neighborhood Change Database. Results of this analysis are reported in another research note.

A number of studies have used the census data on housing units by year built to estimate numbers of units for earlier years back to 1940. Hammer *et al.* (2004) used 1990 census data to make the estimates for split block groups in Midwestern states, and the same group (Radeloff 2005) did very much the same using 2000 census data. Theobold (2001, 2005) made similar estimates for block groups for the entire United States. Using the smaller block groups provided greater spatial resolution than census tracts (though proportional error for individual areas will be greater).

The use of the Neighborhood Change Database in the study should result is significantly lower errors, however. First, the values for 1970 through 1990, while

estimates for the 2000 tracts, are not affected by error resulting form the loss (or gain) of units or from errors in reporting the age, as they are allocations of the actual counts from those censuses. Furthermore, many tract boundaries have not changed, so in these cases there will be no error. Any error will be in the misallocation of units from one tract to a nearby tract where tract boundaries had changed. The estimates for 1940 through 1960 are mostly being made from the year-built data for 1970. They involve estimating for census years 10 to 30 years prior to the year of the data. By contrast, the estimates for 1940 through 1960 made from the 2000 census data are for times 40 to 60 years prior to when the data were reported. One can reasonably expect considerably more error in the number of units over these much longer intervals, as well as greater error in the reporting of the age for structures built that many more years before the census.

One further note on these data. Census tracts were not delineated and data were not reported for all areas of the United States in 1970 and 1980. For tracts without 1970 or 1980 data, numbers of housing units were estimated from the housing units by year built data for 1980 or 1990, whichever was the earlier census year in which tract data was reported. And estimates for the earlier years were likewise made from these same data as well. In the final dataset, 3,519 of the 38,521 census tracts covered areas that were not fully tracted in 1970 and only 581 tracts were in areas not tracted in 1980.

Urban Classification

For each census year, urban areas associated with each major center were identified consisting of those contiguous tracts with a density of at least 1 housing unit per 3 acres or 213.33 housing units per square mile. Density was computed using the land area of the census tract as reported by the census. Only census tracts within the boundaries of the metropolitan area (CSA or MSA) could be included in the urban area.

This density cutoff was selected to be comparable to the population density threshold used by the census for the delineation of Urbanized Areas. Urbanized Areas begin with a core area with a population density of at least 1,000 persons per square mile. Then contiguous or nearby census tracts or blocks are added that have a population density of at least 500 persons per square mile (U.S. Bureau of the Census 2011). For the 2000 census and the United States as a whole, the ratio of total population to total housing units was 2.34. Using this measure of population per housing unit, the housing unit density of 1 unit per 3 acres, 213.33 units per square mile, is equivalent to 499 persons per square mile, virtually the same as the census density cutoff.

When considering the housing unit density threshold of 3 acres per unit, it is important to remember than this is a measure of gross density, not the maximum lot size. The density measure includes not only the areas of lots, but the areas of common areas, streets and roads, and all nonresidential uses from commercial and industrial uses to airports and parks. Thus the maximum lot size associated with this density level would be substantially smaller.

The more detailed criteria for the delineation of the urban areas likewise follow the procedures used by the census for delineated Urbanized Areas, with some modifications to account for the exclusive use of census tracts in the classification. Tracts meeting the density threshold were added to an urban area if they were edge contiguous to the tracts in the existing urban area. Tracts were considered to be edge contiguous if they shared a portion of a common boundary that was visible at a scale of 1:100,000 (to avoid trivial edge contiguity). Tracts sharing only a common vertex were not considered to be contiguous. No provisions were made for including areas close but not contiguous to the current urban area (the “hops” and “jumps” in the Urbanized Area classification. These provisions were seen as being more appropriate to the use of blocks, not to tracts with their much larger areas. The Urbanized Area restriction limiting the addition of tracts to those with areas with less than 3 square miles was not applied here. For Urbanized Areas, this forces the switch from tracts to blocks for adding additional territory, but that was not possible here.

Tracts separated by water including lakes, bays, inlets, and major rivers, including associated flood plains, were considered to be contiguous. Tracts along major rivers that were not urban but where the majority of the tract on both sides were adjacent to urban areas were also considered to be urban.

Any tract or groups of tracts with a land area of less than 5 square miles not meeting the density criterion but totally surrounded by urban tracts was considered to be urban, as with the Urbanized Area classification. Similarly, nonurban tracts or groups of tracts surrounded by urban tracts or major areas of water were considered to be urban. In addition. In a few cases nonurban tracts or groups of tracts with areas greater than 5 square miles that were totally surrounded by urban tracts in 2010 and that were virtually completely developed, primarily with industrial development, as determined by examining satellite imagery, were considered to be urban. In earlier years, once these areas were not totally surrounded by urban tracts, they were no longer considered to be urban.

Tracts containing a major airport that were surrounded by or adjacent to urban areas and had an area of less than 10 square miles were considered to be urban. The area cutoff of 10 square miles (twice the maximum enclosure area) was used because some major airports have areas themselves that are greater than 5 square miles. In addition, airports tend to be surrounded by significant amounts of commercial and industrial development, making it even more likely that such tracts are predominantly urban in character. The tracts with major airports were identified using the point layer of airports from the National Atlas, which included airports with 250 or more emplanements per year (U.S. National Atlas 2014). In addition, five larger tracts containing airports were also included as urban. These were tracts with areas ranging from 11 to 21 square miles that were either totally or almost completely surrounded by urban tracts in 2010, that had extremely large airports (developed or expanded more recently than many airports), and where virtually all remaining area in the tract was completely developed based on satellite imagery. These airport tracts were in the Cincinnati, Dallas-Fort Worth, El Paso, Houston, and Lousiville metropolitan areas.

For census years prior to 2010, a tract would be classified as not urban when the tract was also classified as not urban in the succeeding census year, whether or not it met the urban density threshold in the earlier year. This rule is intended to exclude anomalous cases in which a tract might barely qualify as urban in one year, only to have slightly fewer units the following year because of errors in one of the counts or changes in the existing stock over the decade. This allows consistent identification of areas as to when they become urban. The decision was made to make the consistency going back it time (if not urban at any given time, not urban at the earlier times) rather than going forward (if urban at one time, urban at future times) because the more recent data are considered to be more accurate.

Comparison of these urban areas with the Urbanized Areas found that a small number of urban areas were significantly overbounded in 2010 by the inclusion of a small number of tracts with very large land areas that met the urban density threshold. These would have been tracts that had seen large amounts of development that pushed them over the urban density threshold but that were still mainly undeveloped. Analysis suggested the exclusion of tracts with areas greater than 40 square miles from the urban areas would be a reasonable modification to address this issue. The 40-square mile cutoff resulted in the deletion of 16 tracts with areas exceeding this size plus 2 additional tracts no longer contiguous to their urban areas. Moving to a lower threshold would have led to the exclusion of a significantly larger number of smaller tracts that would have been rendered no longer contiguous.

Exurban Classification

The exurban classification identifies the additional portions of the metropolitan areas beyond the urban areas that show evidence of initial urban development but that have lower densities or may include urban areas not contiguous to the major urban areas. Unlike the case of urban area delineation where the census Urbanized Area specification is broadly used and accepted, no comparable definition for exurban areas exists.

Clark *et al.* (2009) suggested that there is general agreement that “exurbia consists of low-density, urban-dependent settlement that occurs within the commutershed of urban areas.” Berube *et al.* (2006) reviewed a large number of studies that have examined exurban areas and used that to develop their own definition of exurban census tracts, which required a minimum density of 40 acres per housing unit, 25 percent commuting to the Urbanized Area, and population growth above average for the MSA and at least 10 percent.

For this study, a minimum housing unit density criterion will be used as the sole criterion for identfying exurban census tracts. Exurban tracts are limited to the area of the CSA or MSA, so they are within the broad commuting area. Data on commuting comparable to that used in other studies are not generally available for most census years (and for some years, no commuting data are available). An above-average growth criterion

would capture specific types of areas within the broader exurban landscape, but rates of growth in various areas would seem to be something that one would more appropriately consider after designating a broader exurban area. Indeed, Berube and his coauthors, after using the growth rate to define their exurban areas, go back and look at slower-growing tracts as well.

Some researchers have also employed a maximum density threshold when defining exurban tracts as well as a minimum density (see, e.g., Clark *et al.* 2009, and some of the studies reviewed in Berube *et al.* 2006). While this may make sense for identifying areas with a certain type of very low-density development in the urban fringe, it excludes areas of urban development embedded within that urban fringe. These separate urban areas are obviously an important component of the overall pattern of metropolitan development. So the definition for the exurban area for this study will include all tracts contiguous to the urban areas that meet a minimum density threshold, with no maximum.

This still leaves open the question of the minimum housing unit density to be used in designating a tract exurban. Densities used in prior studies of exurban areas have varied widely. Wolman *et al.* (2005) used a threshold of 60 units per square mile (just over 10 acres per unit) in defining their extended urban areas. Berube (2006) used 40 acres per unit, and the studies they reviewed included minimum population density thresholds of 100 and 300 persons per square mile (about 15 acres per unit and 5 acres per unit using the value of 2.34 persons per unit).

This obviously is a very wide range of possible densities. Tracts exceeding different density thresholds in 2010 were mapped to visually assess the implications of the alternative choices. Starting with the lowest density of 40 acres per unit, in the eastern half of the United States, the tracts meeting this standard tended to cover all or most of the metropolitan areas and extended far beyond. When Berube *et al.* (2006) used this level, it was in combination with the 25 percent commuting requirement, and it seems clear from their maps that the commuting requirement was most often the binding criterion limiting the size of the exurban areas. Moving down to 30 acres per unit naturally reduced the areas included somewhat, but the areas of tracts with this density still seemed to be way too large.

The next step was to compare the areas included using density thresholds of 20, 15, and 10 acres per unit. These all seemed much more reasonable. The 20-acres-per-unit standard still resulted in the exurban areas extending too far out from the urban areas in some areas with which I was familiar, while the 15-acre criterion looked more reasonable and for most areas the differences were not great. The decision was made to use as the minimum density for exurban area tracts the level of one housing unit per 15 acres, 42.67 housing units per square mile. This is equivalent to a population density of about 100 persons per square mile.

Contiguous areas of tracts meeting this density threshold around the urban areas tended to be limited to the areas of the CSAs or MSAs. For 34 of the 59 areas, the

contiguous areas of tracts with densities exceeding 15 acres per unit were either confined within those areas or, if they extended beyond the boundaries, it was into an adjacent MSA. When the area of contiguous exurban density tracts did extend beyond the CSA or MSA boundaries otherwise, it was most often in only 1 or 2 points along the boundary, with the area of contiguity most often 1 or 2 tracts beyond the CSA or MSA boundary. So this density level tends to identify areas of contiguous tracts that are largely within the metropolitan area boundaries. In any event, the exurban areas were limited to tracts within the CSA or MSA.

The more detailed criteria used for delineating the exurban areas were the same as those used for the urban areas with two exceptions. Edge-contiguity of tracts had to be observable at a scale of 1:250,000 (because of the larger tracts and larger areas involved with the exurban areas). And the classification of tracts as exurban did not take into consideration the locations of airports. The other criteria involving water and enclosed areas were the same.

Central Business Districts

Measures of centralization require the identification of the central locations for each of the major urban centers. Since the idea of centralization is closely related to the idea of the monocentric model from urban economics that assumes commuting to the central business district (CBD), the location of the CBD is used to establish the central locations.

One of the only attempts to specify the locations of all CBDs was a report from the 1982 economic censuses (U.S. Bureau of the Census 1983). These locations have continued to be used over the succeeding decades. The report lists the census tracts comprising the central business district for many (though not all) large cities. These tracts, when they still existed in the set of 2000 tracts, were designated as CBD tracts. Because tracts are split, merged, and renumbered, not all of the tracts could be found and not all of the CBDs could be designated in this manner.

For those ares where the census report did not establish the location of the CBD, online maps were used to estimate the tracts to be included. The location of the CBD is generally obvious from the location of major roads. These areas were confirmed by determining the locations of major government buildings, which are most often located within the CBD. Using such approaches, it was possible to designate CBD tracts with reasonable confidence as to accuracy for all of the areas.

The tracts (if more than one) comprising the individual CBDs were merged and the centroids of these CBD areas were taken as the point locations of the CBDs.

An alternative approach to identifying central locations would have been to base the selection on the distribution of housing units and the pattern of residential density. Paulsen (2012) computed the mean population center. The mean population center may or may not coincide with the CBD, depending upon the asymmetry of the urban area. In

particular, an urban area located on a major body of water will likely have the CBD located close to the water, as that will have been the origin of the urban area. However, since the urban area extends away from the water, the mean population center will be displaced away from the water and the CBD.

Tract Locations and Distances

The locations of tract centers are needed for the calculation of distances to the point locations of the CBDs, for the assignment of the tracts to centers, for calculation of some measures of centralization, and for the estimation of the parameters of the negative exponential density model. They are also needed for the calculation of some types of spatial weights used for certain types of analyses.

The centroids of the census tracts were used as the tract locations and the coordinates were included in the dataset. The coordinates are in meters for the National Atlas Lambert Equal Azimuthal projection, which has been used for the mapping required for the creation of the dataset. Consideration was given to the reprojection of the areas into Universal Transverse Mercator (UTM) coordinates for the appropriate zones in order the provide greater accuracy for the calculation of distances. Tests were conducted for Seattle and San Diego, two of the most extreme areas for the National Atlas projection for which one would expect the greatest errors when using those coordinates. Tracts in those areas were converted to centroids, coordinates were added, and distances to an arbitrary central point were calculated using both the National Atlas projection and the appropriate UTM projections. The distances were compared and the errors associated with the use of the National Atlas projection as compared with the UTM coordinates were calculated. The maximum absolute percentage error was 1.34 percent for San Diego, with the mean absolute percentage error of 0.78 percent. Comparable errors for Seattle were slightly lower. Given the small size of these errors compared with all of the other sources of error in the data, it was determined that use of the coordinates from the National Atlas projection was reasonable.

Assignment to Centers

For metropolitan areas with multiple centers, each tract must be assigned to a specific center for the calculation of measures of centralization (which are made with respect to the designated center). For those areas where the centers had separate Urbanized Areas in 2010, tracts were assigned based on the Urbanized Area in which the tract was located. Tracts outside the Urbanized Areas were assigned to the Urbanized Area that was closest or, in some cases, to the larger Urbanized Area when the location suggested that this area would likely have greater influence and interaction with the tract.

For those areas with multiple centers where the centers were all included in a single Urbanized Area in 2010, tracts were assigned to the center for which the CBD location was closest.

The San Francisco-Oakland-San Jose area represented a combination. San Jose is a separate Urbanized Area in 2000. Tracts in and adjacent to the San Jose Urbanized Area and those in areas extending to the west and southwest were assigned to San Jose. San Francisco-Oakland is a single Urbanized Area, so that did not provide a basis for assignment to those centers. However, the geography of the area made it easy to assign those tracts not assigned to San Jose that were west of the bay to San Francisco and those tracts east of the bay to Oakland.

There was one anomaly. For Tampa-St Petersburg, a small group of tracts across the bay from St Petersburg were actually closer to the center in St Petersburg than to the Tampa center. There is a long bridge across, but these tracts really needed to be assigned to Tampa.

References

Berube, Alan, Audrey Singer, Jill H. Wilson, and William H. Frey. 2006. *Finding Exurbia: America's Fast-Growing Communities at the Urban Fringe*. Washington, DC: The Brookings Institution.

Clark, Jill K., Ronald McChesney, Darla K. Munroe, and Elena G. Irwin. 2009. Spatial characteristics of exurban settlement pattern in the United States. *Landscape and Urban Planning* 90, 3-4 (April): 178-188.

Ewing, Reid, and Shima Himidi. 2014. *Measuring Urban Sprawl and Validating Sprawl Measures*. Prepared for National Cancer Institute, National Institutes of Health; Ford Foundation. Smart Growth America. Metropolitan Research Center, University of Utah. Downloaded from http://gis.cancer.gov/tools/urban-sprawl/ on April 8, 2014.

Ewing, Reid, Rolf Pendall, and Don Chen. 2002. *Measuring Sprawl and Its Impact*. Washington, DC: Smart Growth America. Downloaded from http://www.smartgrowthamerica.org/documents/MeasuringSprawl.PDF on April 8, 2014.

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.

Lopez, Russell. 2014. Urban sprawl in the United States: 1970-2010. *Cities and the Environment (CATE)* 7, 1: article 7. Available at: http://digitalcommons.lmu.edu/cate/vol7/iss1/7.

Lopez, Russ, and H. Patricia Hynes. 2003. Sprawl in the 1990s: Measurement, distribution, and trends. *Urban Affairs Review* 38, 3 (January): 325-355.

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

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. 2011. Urban area criteria fot 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. 2014a, *Geographical Terms and Concepts—Core Based Statistical Areas and Related Concepts*. Downloaded from http://www.census.gov/geo/reference/gtc/gtc_cbsa.html on July 8, 2014.

U.S. Bureau of the Census. 2014b. County populations from 2010 Census. Downloaded from http://factfinder2.census.gov/faces/nav/jsf/pages/index.xhtml on February 13, 2014.

U.S. National Atlas. 2014. Airports shapefile. Downloaded from http://www.nationalatlas.gov/atlasftp.html?openChapters=chptrans#chptrans on April 4, 2014.

U.S. Office of Management and Budget. 2010. 2010 Standards for Delineating Metropolitan and Micropolitan Statistical Areas. *Federal Register* 75, 128 (June 28): 37246-37252.

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

Wolman, Harold, George Galster, Royce Hanson, Michael Ratcliff, Kimberly Furdell, and Andrea Sarzynski. 2005. The fundamental challenge in measuring sprawl: Which land should be considered? *Professional Geographer* 57, 1: 94-105.

Appendix—List of Areas

Albany-Schenectady-Troy
Albuquerque
Atlanta
Austin
Birmingham
Boston-Providence
Buffalo
Charlotte
Chicago
Cincinnati
Cleveland-Akron
Columbus
Dallas-Fort Worth
Dayton
Denver
Detroit
El Paso
Fresno
Grand Rapids
Greensboro--Winston-Salem--High Point
Greenville-Spartanburg
Harrisburg-York
Hartford
Houston
Indianapolis
Jacksonville
Kansas City
Knoxville
Las Vegas
Los Angeles
Louisville
Memphis
Miami-Fort Lauderdale-West Palm Beach
Milwaukee
Minneapolis-St Paul
Nashville
New Orleans
New York
Norfolk-Virginia Beach
Oklahoma City
Orlando
Philadelphia
Phoenix
Pittsburgh

Portland
Raleigh-Durham
Richmond
Rochester
Sacramento
Salt Lake City-Ogden-Provo
San Antonio
San Diego
San Francisco-Oakland-San Jose
Seattle-Tacoma
St Louis
Tampa-St Petersburg
Tucson
Tulsa
Washington-Baltimore