The Evolution of Urban and Metropolitan Areas: An Essay

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

May 2019

Abstract

The Bureau of the Census first defined Metropolitan Districts and set the current minimum for an area and its population to be considered urban in 1910. The 1950 census saw the first reporting of data for Standard Metropolitan Areas and Urbanized Areas, the latter being the denser, built-up portions of metropolitan areas. The definitions of these areas evolved during the remainder of the last century with the names for the metropolitan areas changing, but the basic approach remained constant. Major changes to the definitions were made around 2000. The use of small areas rather than the entire areas of cities constituted a significant improvement for Urbanized Areas. However, the rules put in place to identify separate Urbanized Areas caused fragmentation of some areas and, because of the way in which the new Metropolitan Statistical Areas were now being defined, caused the splitting up of numbers of areas that had been unified under the previous definition. The problem was compounded by the decision for 2010 to freeze Urbanized Area divisions, ending the possible combination of both those areas and the Metropolitan Statistical Areas that has occurred in the past in response to the growth and evolution of metropolitan regions.

Introduction

The Bureau of the Census has defined Urbanized Areas, the built-up portions of metropolitan areas, since 1950. The Office of Management and Budget (and its predecessor) have likewise defined the larger Metropolitan Statistical Areas (using various names) also since the middle of the last century. This was preceded by the Census reporting data for Metropolitan Districts and settling on the current threshold for identifying the urban population in 1910.

These areas are used in reporting a variety of data and therefore shape the information available for urban and metropolitan areas. Beyond that, they can influence the very nature of research. For example, a large body of research comparing characteristics of the older, inner portions of urban areas with the newer suburbs has drawn those comparisons using data for the central central cities (more recently
principal cities) identified for Metropolitan Statistical Areas and the portions of the Metropolitan Statistical Areas outside those cities. In addition, the widespread use of these areas helps to literally shape perceptions of urban and metropolitan areas and even their identities.

It is obvious that anyone using data for these areas should have an understanding of how they have been defined and the issues and limitations created by those definitions. I say “should” because I have read far too many pieces of research where the authors clearly did not have such understanding. This is especially important for research looking at urban or metropolitan areas over time. Not only do the areas change as urban areas grow and evolve but the very ways in which the areas are defined change as well. Sometimes the changes in definition are relatively minor but at other times they can be highly significant.

I have literally worked with data for these areas from the census and other sources for every census year from 1910 up through the present. This has required me to understand the definitions of the areas at each decade and how they have changed. As a result, I not only have a reasonable knowledge of how these definitions have evolved but also have formed clear opinions about the strengths and weaknesses of the definitions.

This paper is my assessment of the evolution of the definitions of urban and metropolitan areas. It is not intended as a comprehensive history that delves into all of the details. Instead I am emphasizing the issues that I think are most important to understand when using these areas and my assessment of the positive and negative aspects of the definitions. Hence my description of this paper in the title as an “essay.”

The essay begins with the definition of the urban population and Metropolitan Districts in 1910 and the relatively minor changes through 1940. The middle of the last century saw the development of Urbanized Areas and Standard Metropolitan Areas that established the basic concepts and approaches still in use today. These areas were first used in reporting decennial census data in 1950 and are discussed in the next section of the paper. In the succeeding decades up through the 1990s, the definitions evolved but continued to follow the general pattern set when they were first established, with a few changes that were more significant. This is discussed next. The year 2000 brought wholesale revision, with completely new definitions for both Urbanized Areas and Metropolitan Statistical Areas which are critical to understand when using data from before and after the changes. Finally, the 2010 definitions saw few changes—so few that I raise the question as to whether the definitions are now failing to capture the evolution of urban and metropolitan areas as they grow and change as had been the case for the prior half century. I will then briefly offer suggestions for a better way to proceed.
Urban Population and Metropolitan Districts in 1910

That the beginning of the twentieth century would be when the Census Bureau chose to establish the current definition of the urban population and provide data for Metropolitan Districts is not surprising. Cities in the United States were growing rapidly and the nation’s population was becoming much more urban. Indeed the proportion of the population considered urban under the new definition had risen from just 5 percent at the first census in 1790 to where it would become a majority at the 1920 census.

Defining the Urban Population of the United States

This section and the later sections dealing with the definition of the urban population and Urbanized Areas draws heavily on a paper discussing the census urban and rural classification since 1910 by Michael Ratcliff (2015), who was in the geography division of the Census Bureau.

The Census Bureau first started defining the urban population in 1870 as those persons living in cities and other incorporated places with populations of 8,000 or more. After using this threshold for several more censuses, the change was made in 1910 to consider as urban persons living in such places with populations of 2,500 or more (U.S. Bureau of the Census 1913). That definition of the urban population remained constant through the 1940 census. The 2,500 threshold continues to be the minimum to be considered urban.

I had long thought that the 2,500 cutoff was too low for a place to be considered urban. I have lived in a number of small communities in Wisconsin while growing up. Places with just a few thousand people never seemed very “urban” to me. As I read through background material in preparation for the writing of this paper, I changed to having mixed feelings about the 2,500 level. The census definition of urban is also their definition of rural, which is the remainder of the population that has not been classified as urban. People living in places with populations of less than 2,500 are considered rural, along with those living outside of places of any size. The smallest community I lived in had a population of just about a thousand. While I certainly did not feel I lived in an urban area, I also did not consider myself as being rural. My friends who were rural lived out in the country, on farms or isolated houses or small clusters of houses. They rode the school bus to get to school. Those of us living in town walked.

If the population is to be classified into just 2 groups of urban and rural, then any threshold will necessarily be too low for urban and/or too high for rural. So perhaps 2,500 is a reasonable compromise, if the choice is to be between only those 2 possibilities. Of course one could imagine classification into more than 2 groups, say
urban, small town, and rural. But this adds complexity and raises the problem of establishing multiple thresholds.

Metropolitan Districts

As metropolitan areas grew in the latter decades of the nineteenth century and the beginning of the twentieth century, they expanded beyond the boundaries of the large cities at their core to include increasing suburban development. Some of this growth took place in incorporated places adjacent to the cities, but additional development occurred in nearby areas that were not part of any incorporated place. Starting with the 1910 census, the Census Bureau chose to delineate and report population data for these larger areas, which were called Metropolitan Districts.

For 1910, Metropolitan Districts were delineated around cities having populations exceeding 200,000. These areas consisted of the large city, other incorporated places, and other civil divisions (subdivisions of counties having various names in different states) within 10 miles of the boundary of the city that had a population density over 150 persons per square mile. Curiously, for cities having populations between 100,000 and 200,000, the adjacent area was defined as consisting of all divisions within the same 10 miles from the boundary with no minimum population density requirement (U.S. Bureau of the Census 1913).

For the 1930 census, the 10-mile limit was dropped and Metropolitan Districts consisted of all contiguous civil division having population densities exceeding 150 persons per square mile extending out to any distance. In addition, the minimum size for the central city required to form a Metropolitan District was reduced to 50,000, though with the additional requirement that the entire area would need to have a population of at least 100,000 (U.S. Bureau of the Census 1932). This reduction of the minimum to 50,000 was important, as it established the precedent for the definition of Urbanized Areas and metropolitan areas in 1950 and continues to be used to this day.

The Metropolitan Districts were a reasonable first effort at defining the larger urban or metropolitan areas around large cities. In some respects they can be seen as something between the Urbanized Areas and Metropolitan Statistical Areas that followed. The civil divisions used in delineating the areas were larger than the enumeration districts and later blocks used for the Urbanized Areas but were smaller than the counties used for the Metropolitan Statistical Areas. The population density threshold of 150 persons per square mile was far lower than the 1,000 and later 500 cutoffs used to define urban and Urbanized Areas. And the areal extents of Metropolitan Districts puts them between Urbanized Areas and Metropolitan Statistical Areas as well.

The density criterion meant that Metropolitan Districts were defined by settlement patterns. At least as early as the 1920s, the Census Bureau recognized the
limitations inherent in such an approach. The report on Metropolitan Districts for the 1930 census includes a very long footnote addressing the problem, first quoting from an outside report that metropolitan regions should encompass the area in which “the daily economic and social life…are predominantly influenced by the central city.” Indeed, the census asked the chambers of commerce of large cities define their metropolitan areas by considering a wide range of factors from utility services, retail and newspaper delivery, and commuting, to membership in social clubs. They deemed the results inconsistent across the areas and continued with the Metropolitan District definition (U.S. Bureau of the Census 1932).

Urbanized Areas and Standard Metropolitan Areas in 1950

The 1950 decennial census was the first to report results for the newly defined Urbanized Areas and Standard Metropolitan Areas, the name originally given to what are now called Metropolitan Statistical Areas. These set the basic purposes and patterns for the areas still in use today. Urbanized Areas defined the urban population and simultaneously the extent of the more-or-less built-up urban area. Standard Metropolitan Areas identified the larger areas of central city influence and significant interaction.

Urbanized Areas

The motivation for creating Urbanized Areas was clearly stated in a report on the 1950 census: “The major objective of the Bureau of the Census in delineating these areas was to provide a better separation of urban and rural population in the vicinity of our larger cities than was provided under the old definition” (U.S. Bureau of the Census 1952). The census had recognized the presence of significant and growing numbers of people living in areas having urban densities outside of incorporated areas at the edges of large urban areas. The Urbanized Areas were intended to include these people in the count of the urban population of the country. At the same time, the Urbanized Areas defined the extent to the urban, built-up portions of metropolitan areas.

Urbanized Areas consisted of a city with a population of 50,000 or more, the adjacent incorporated places, and all contiguous census enumeration districts with a density exceeding 500 housing units per square mile. Enumeration district were small areas, larger than blocks but smaller than census tracts, used in the collection of the census data (U.S. Bureau of the Census 1952, 1994; Ratcliff 2015). I must emphasize that this description of the definition and all of the others in this essay give only the major features necessary to understand the nature of these areas and their evolution. For the precise details, see the sources cited.
In general this was a reasonable definition of the extent of urban territory around large cities. One might quibble about the density threshold, and indeed the census changed this at the next census. The inclusion of the entire areas of incorporated areas was the only feasible option given what was practical at the time working with paper maps and using planimeters to measure areas. But this did lead to the inclusion of territory within some cities that was undeveloped and in no meaningful sense urban. The extent of this over bounding varied greatly from city to city but was definitely substantial for some areas. Of course the same problem existed with the original definition of the urban population in 1910 which also included the entire populations of incorporated areas. For purposes of counting the urban population, this made little difference, as the undeveloped areas had few people. However, in delineating Urbanized Areas, which also effectively portrayed the extent of the built-up urban areas, including these undeveloped areas was more significant.

Standard Metropolitan Areas

By the 1940s, a number of federal government agencies were reporting statistics for metropolitan areas, each using its own, different definition for those areas. The Bureau of the Budget (predecessor to the Office of Management and Budget) established a committee with representatives from the various agencies to devise a common set of areas to be used for reporting metropolitan data to enhance the comparability of data from the various agencies. These areas were called Standard Metropolitan Areas (SMAs) and were first used for reporting decennial census data for the 1950 census. The census introduction described the Standard Metropolitan Areas as encompassing “the entire population in and around the city whose activities form an integrated social and economic system” (U.S. Bureau of the Census 1952).

The decision was made to use counties as the building blocks for the Standard Metropolitan Areas (except in New England) as these were the smallest units for which many types of state and local as well as federal data were available. Like Urbanized Areas, Standard Metropolitan Areas started with a city having a population of 50,000 or more and included the county or counties within which it was located. Contiguous counties were added that met criteria for social and economic integration with the central city and for metropolitan character. Integration was determined by at least 15 percent commuting between the county and the county or counties containing the
central city and/or volumes of telephone calls between counties.¹ Metropolitan
center was determined by the agricultural and nonagricultural workforce or the
population densities of minor civil divisions.

For the New England states, towns rather than counties were used as the
building blocks for Standard Metropolitan Areas. The argument was that towns were
more important than counties in these states and more data were available for these
smaller units in New England than was the case for smaller areas elsewhere. The
criteria for inclusion of towns in a Standard Metropolitan Area were also different from
those used for counties. This decision to use towns has always struck me as
unfortunate. This made the New England Standard Metropolitan Areas inconsistent
with those in the rest of the country. For one thing, being defined using the smaller
units, they typically encompassed smaller areas than would have been the case
otherwise. In addition, counties were still the smallest units for which certain types of
data were available, such as some economic data from federal agencies. Such data could
not be aggregated to and reported for the New England Standard Metropolitan Areas.
This was later addressed by defining a second set of metropolitan areas for New
England, the New England County Metropolitan Areas. But this meant there were now
2 different sets of metropolitan areas for those states.

The use of units as large as counties meant that that Standard Metropolitan Areas
only approximated the territory that could reasonably be considered metropolitan and
integrated with the central city and county. Many Standard Metropolitan Areas were
over bounded, with only portions of the counties having metropolitan character and
being responsible for the commuting to the central counties, with the remainder of the
counties being very rural. (Some under bounding was also possible when a county with
only a small portion really metropolitan was not included in the Standard Metropolitan
Area. But this was less common and a much smaller issue because the standards for
inclusion were sufficiently generous.)

These comments about Standard Metropolitan Areas frequently being over
bounded are not meant as a criticism of the use of counties as the basic building blocks.
The advantage of far greater data availability for counties makes the choice entirely
appropriate. And for the reporting of total counts for quantities related to population,
housing, employment, and the like, the additional area and the over bounding makes
little difference. The additional rural territory would have small numbers of people and
jobs, certainly relative to the totals for a metropolitan area.

¹ Various sources from the Bureau of the Census differ with respect to the criteria used for integration.
The 1950 census report (U.S. Bureau of the Census 1952) mentions both commuting and telephone calls.
The 1994 Geographic Areas Reference Manual says that commuting was the “main” measure of
integration and makes no mention of telephone calls (U.S. Bureau of the Census 1994). However a paper
by Ratcliff (2002), an employee at the Bureau of the Census, says that telephone call patterns were used as
the measure of integration for the metropolitan areas identified for the 1950 census.
But the over (and under) bounding means that the amounts of land included in Standard Metropolitan Areas can vary greatly and are not meaningful quantities. Likewise, measures dependent on land area such as densities are also meaningless. This is where I have a quarrel with the census. Starting with the 1950 census, the land areas of the Standard Metropolitan Areas and their population densities were reported. The Standard Metropolitan Areas with the second and third highest population densities, over 3,000 per square mile, were New York and Boston, not surprising as their cores are large, old, dense cities. But the Standard Metropolitan Area with the highest population density was Milwaukee, a reasonably large area to be sure but certainly not a Chicago or Philadelphia. But the Milwaukee Standard Metropolitan Area in 1950 consisted only of Milwaukee County, most of which was developed. Arguably areas in 2 adjacent counties would seem have been at least somewhat metropolitan in character (Ottensmann 1975).

This reporting of population densities was far more nonsensical for the Standard Metropolitan Areas with the lowest densities. Compared with most counties in the eastern U.S., some counties, especially in the West, can be very large. The San Bernardino Standard Metropolitan Area in 1950 consisted of San Bernardino County, the largest county in terms of land area in the lower 48 states. It has an area of over 20,000 square miles, larger than 9 states, and extends from Los Angeles County east to the Nevada and Arizona borders, in some instances over 200 miles. Most of the county consists of mountains and deserts. These areas are not metropolitan in any meaningful sense. More than just being rural, they are largely uninhabited. Not surprisingly, the San Bernardino Standard Metropolitan Area had the lowest population density of any Standard Metropolitan Area in 1950, 14 persons per square mile(!). The only information this density value provides is that the Standard Metropolitan Area includes a very large county, most of which was more-or-less empty.

This presentation of population densities for metropolitan areas by the census has continued. A recent report done by the Census Bureau examined trends in metropolitan areas over the decade between the last 2 censuses (Wilson, et al. 2012). An entire chapter was devoted to population densities. It begins by looking at overall population densities for Metropolitan Statistical Areas, including listing those areas with the highest and lowest densities. They hint that there is an issue created by larger counties in the West but continue their presentation. Only later do they acknowledge that densities “can be heavily affected by the size of the geographic units for which they are calculated” and proceed to use population-weighted densities calculated using data for census tracts, which is much more appropriate. But the damage had been done. The non-weighted population densities had been presented and reported in the first part of the chapter as if they were meaningful, and they are not.

Given that the spatial extent of Standard Metropolitan Areas is so arbitrary, the mapping of the areas included in Standard Metropolitan Areas can be quite misleading.
as well. The 1950 census report (and reports to follow) included a map of the United States showing the areas included within Standard Metropolitan Areas. Looking at this map, the largest area of contiguous Standard Metropolitan Areas is in Southern California with the Los Angeles, San Diego, and San Bernardino Standard Metropolitan Areas. This exceeds in size (area shown on the map) the line of Standard Metropolitan Areas extending from Washington through Philadelphia and New York to Hartford, Connecticut. This gives a completely incorrect impression of the amount and extent of metropolitan settlement. Of course it is desirable to provide a map showing the locations of Standard Metropolitan Areas. But such a map could use symbols to represent the Standard Metropolitan Areas rather than indicating the extent of the counties included, perhaps with larger symbols for those areas with greater populations.

Evolution of Urbanized Areas and Metropolitan Areas through the 1990s

The period after the initial definition of the Urbanized Areas and Standard Metropolitan Areas in 1950 through the 1990s can best be characterized as one of evolution. Definitions were reviewed and changes were made for every census, but most changes were gradual. The collective effect of these changes over the half century was not insubstantial, however. And a few more significant changes were made during this period. But the general approach remained fairly constant.

Urbanized Areas

From 1960 through 1990, the Urbanized Area definition evolved, changing somewhat at each census, but the fundamental approach remained unchanged. The minimum density threshold for areas to be included in the Urbanized Areas was changed to 1,000 persons per square mile, a reduction from the original level of 500 housing units per square mile. The requirement that an Urbanized Area have a city with a minimum population of 50,000 was relaxed, first allowing “twin” cities and later being eliminated in favor of a requirement of a “densely settled area” of 50,000 or more. The census defined some of the greatly over bounded cities as “extended cities,” with only the more densely settled portion being counted as urban and included in Urbanized Areas. This addressed only the extreme cases, however. Significant undeveloped rural territory remained in many other cities. And finally, census blocks came to be used in delineating Urbanized Areas (U.S. Bureau of the Census 1994; Ratcliff 2015).

This period saw the effect of one other aspect of the definition of Urbanized Areas, the merging of Urbanized Areas as areas grew together. Two or more large cities in close proximity could come to be located within an area of continuous urban
settlement as the urban areas grew together. If these were in separate metropolitan areas (Metropolitan Statistical Areas or their predecessors) the urban areas would be split into separate Urbanized Areas at the metropolitan area boundary. Then if the metropolitan areas were merged into a single area, the Urbanized Areas would likewise be merged. A good example of this is the Dallas-Fort Worth area. Through 1970, Dallas and Fort Worth were separate metropolitan areas and Urbanized Areas. As of 1980, the metropolitan areas were merged into a single combined area. The Urbanized Areas were therefore combined as well. The practice during this period was appropriate and reflected the growth and evolution of urban and metropolitan areas. However problems would arise with the new definitions of Urbanized Areas and Metropolitan Statistical Areas in 2000.

**Metropolitan Areas**

Most of the changes to the metropolitan area definition through the definition developed in 1990 and last used for reporting the 2000 census were fairly modest and evolutionary. First, the names for these areas changed several times. Starting with the 1960 census, they were called Standard Metropolitan Statistical Areas (SMSAs) to emphasize that were intended for the reporting of statistical data and were not defined to serve other purposes (though they frequently were used otherwise). In the 1990 census, the name changed again, to Metropolitan Statistical Area (MSA), which continues in use to this day. (Variants on that were also used, as will be discussed later in this section.) (U.S. Bureau of the Census 1961, 1983, 1993)

Other changes involved refinements to the data and criteria used for delineating the areas. For the Standard Metropolitan Areas in 1950, local employment data were used for the commuting criterion, but these sources were variable. The 1960 census included a question on place of work for the first time, which provided consistent data on commuting that could be used for area delineation. As with the Urbanized Areas, the requirement for qualifying as a metropolitan area were relaxed, first allowing twin cities and then an Urbanized Area having a larger population. The criteria for adding counties to a metropolitan area were expanded to allow a tradeoff between metropolitan character and integration. The standards used in the 1990s and for the 2000 census had 6 combinations of metropolitan character and percent commuting to the central counties. The commuting thresholds varied from 15 to 50 percent, with the lower cutoffs requiring greater levels of metropolitan character. Note that the 15 percent standard from 1950 still remained (U.S. Bureau of the Census 1994; U.S. Office of Management and Budget 1990).

Over the period from 1950 to 2000, the metropolitan areas evolved, not only expanding outward as they grew but also growing together. This was reflected in the delineation of the Standard Metropolitan Statistical Areas and Metropolitan Statistical
Areas. In the 1980 census, Dallas and Fort Worth became a single Standard Metropolitan Statistical Area. By 1990, the Los Angeles Standard Metropolitan Statistical Area had been combined with the adjacent Anaheim-Santa Ana, Oxnard-Ventura, and Riverside-San Bernardino areas. San Francisco-Oakland was likewise combined with San Jose (U.S. Bureau of the Census 1983, 1993).

The metropolitan area definition used for reporting the 1990 census brought the greatest changes, both to the definition of central cities and with the subdivision and naming of areas. Beginning with the original Standard Metropolitan Areas in 1950s, central cities were identified for each area. These included the largest city in an Standard Metropolitan Area and sometimes additional cities if they were also relatively large. This was formalized with the central cities consisting of the largest city and up to 2 additional cities if their populations were at least one-third that of the largest city. The census then reported population and other data for central cities (sometimes individually, sometimes combined) and for the remainder of the metropolitan area.

This division between central cities and the remainder of metropolitan areas formed the basis for large numbers of studies comparing various characteristics of these areas, often referring to these as contrasting central cities with their suburbs, the idea being that the former were the older urban cores and the latter represented the newly developed areas. These were always fraught comparisons, as metropolitan areas varied widely with respect to the proportions of their populations included in the central cities, and the central cities ranged from areas that had been fully developed by the early twentieth central to others that were continuing to see new development. Nevertheless, for the United States as a whole, such comparisons did have some value.

Starting with the 1990 census, central cities were defined as those cities with relatively greater concentrations of employment compared to their populations, without limits on their number or size in relation to the largest areas. This resulted in the addition of central cities to some Metropolitan Statistical Areas. These cities were as diverse as Cambridge, Massachusetts; East St. Louis, Illinois; Arlington, Texas; and Palm Springs, California. Not all were “central” as in being part of the center of the Metropolitan Statistical Area. To argue that they should be called “central” because of their importance as places of employment would seem to be a stretch. But perhaps most important was that this non-trivial change was made to the definition of central cities and could easily be missed or just ignored. Indeed, the census put out a short profile on metropolitan area populations after the 1990 census that included a chart showing the percentages of the United States population living in central cities and in the portions of metropolitan areas outside central cities for each census from 1900 to 1990. The chart shows a small be steady decline in the percentage in central cities from 1950 to 1980 but then a small increase from 1980 to 1990. The typical reader is likely to conclude that the pattern of decline for the central cities may have been halted and even reversed. The chart did include the disclaimer that the values reflected the definitions in effect at each
census. But it did not go on to explain that the reversal of the trend in the last decade resulted solely from the change to the definition of central cities and the addition of more new central cities. Using a consistent definition of central cities, the decline would have been seen to be continuing (Ottensmann 1996; U.S. Bureau of the Census 1991).

The 1990 census also saw another, more significant change to the metropolitan area definition that was very confusing and unfortunate and that continues to create problems for both the Urbanized Area and Metropolitan Statistical Area definitions. As mentioned, the name for the areas was changed to Metropolitan Statistical Area. In addition, the decision was made to subdivide the largest Metropolitan Statistical Areas into smaller divisions. The reason put forward was that this would provide more detailed data for those wishing to analyze portions of the Metropolitan Statistical Areas. I have always seen this as a rather weak argument. First, this assumed that the subdivisions being created corresponded to the smaller areas that people might wish to study. Second, the divisions continued to use counties, and since much of the data people might be interested in analyzing for county-based subdivisions was available for counties, it would have been trivial to aggregate the county data to such areas. The only exception might be for certain data that are not reported for counties but only for metropolitan areas because of limited sample sizes or disclosure rules (this would be mainly economic data). Then if the agencies reporting such data would choose to also report it for the subdivisions, having the subdivisions might be helpful.

I am not going to describe the criteria used for subdividing the Metropolitan Statistical Areas as I feel that those subdivisions are quite nonsensical in some cases. I see no point in delving into criteria that I believe do not produce results that make sense. Many of the subdivisions are quite reasonable, however. For those Metropolitan Statistical Areas that were the product of 2 or more metropolitan areas growing together, the subareas were essentially the originally separate areas. For example, the Dallas-Fort Worth Metropolitan Statistical Area was divided into the Dallas and Fort Worth portions. The subdivisions were called Primary Metropolitan Statistical Areas (PMSAs). This term has always bothered me, as I do not see how these areas should be considered “primary.”

But some of the subdivisions of Metropolitan Statistical Areas that had just expanded without combining with adjacent comparable Metropolitan Statistical Areas were just silly. The New York area provides a clear example. One of the subareas is the Newark Primary Metropolitan Statistical Area, consisting of the city of Newark and numbers of the surrounding counties. I suppose this is not unreasonable: Newark is certainly a large city and the counties included undoubtedly have significant integration with Newark. But this gets a little confusing in that all of this is also part of the New York area and is significantly integrated with New York City. In any event, all of the areas in New Jersey that were part of the original New York Metropolitan Statistical Area are in various New Jersey-only Primary Metropolitan Statistical Areas.
Long Island and the 2 counties that are not part of New York City but are among New York’s suburbs. This gets designated the Nassau-Suffolk Primary Metropolitan Statistical Area. How can this possibly be called any type of “metropolitan area.” It is suburban. It doesn’t have a large urban center. Indeed, that is why it was necessary to use the names of the 2 counties instead. After doing all this subdivision, we finally get to the New York Primary Metropolitan Statistical Area. It consists of New York City, of course, and the counties in New York state on the mainland north of the Bronx. The only way I can understand this areas is that it is what’s left of the original Metropolitan Statistical Area after all of the other Primary Metropolitan Statistical Areas have been carved out.

Now the final bad decision was to rename those Metropolitan Statistical Areas that had been subdivided into Primary Metropolitan Statistical Area as Consolidated Metropolitan Statistical Areas (CMSAs). These areas were not the product of any consolidation, of the Primary Metropolitan Statistical Areas or anything else other than counties, which was true for all Metropolitan Statistical Areas. They were the original Metropolitan Statistical Areas defined using exactly the same criteria that were used in defining all of the Metropolitan Statistical Areas that were not subdivided (which continued to be termed Metropolitan Statistical Areas). Giving these areas the new name Consolidated Metropolitan Statistical Area obscures (deliberately?) the fact that the Consolidated Metropolitan Statistical Areas are the areas directly comparable to the other Metropolitan Statistical Areas, not the Primary Metropolitan Statistical Areas. Especially after the next set of changes to the metropolitan area definition, I have started wondering whether some people involved in the process just disliked the fact that metropolitan areas were becoming larger and merging and wanted to do something about that.

With the existence of both the Consolidated Metropolitan Statistical Areas and the Primary Metropolitan Statistical Areas, many researchers comparing characteristics of metropolitan areas have chosen to use the Primary Metropolitan Statistical Areas with the undivided Metropolitan Statistical Areas. I have seen the choice of Primary Metropolitan Statistical Areas over Consolidated Metropolitan Statistical Areas justified (when authors have even bothered with an explanation) by saying that the Primary Metropolitan Statistical Areas were more homogeneous than the Consolidated Metropolitan Statistical Areas. Of course this is true. But this is a strange justification in

---

2 Groups in the Kansas City area anticipated the possibility of this happening. They were concerned that the subdivision of the Kansas City Metropolitan Statistical Area into 2 Primary Metropolitan Statistical Area, Kansas City MO and Kansas City KS, would have the effect of making their area(s) appear smaller on lists of metropolitan areas than if they remained an undivided Metropolitan Statistical Area. They registered complaints about the proposed subdivision of their area. The Office of Management and Budget responded by making the choice of whether a Metropolitan Statistical Area was to be subdivided a matter of local option.
the context of metropolitan area analysis, as metropolitan areas are inherently highly heterogeneous areas. Where the decision to use Primary Metropolitan Statistical Areas rather than Consolidated Metropolitan Statistical Areas is especially incongruous has been in studies aiming to provide measures of the level of urban sprawl for metropolitan areas. Sprawl is associated with suburban development. If you measure sprawl for an essentially suburban Primary Metropolitan Statistical Area such as Nassau-Suffolk, of course it will have higher levels of sprawl than most Metropolitan Statistical Areas that include a combination of more sprawling suburban areas and less-or non-sprawling large central cities. And this goes both ways. Excluding some suburban areas from a metropolitan area will almost certainly guarantee that the remaining area is seen as less sprawling than if the suburban areas had not been excluded. Finally, I have even seen several studies in which the authors used Consolidated Metropolitan Statistical Areas for one aspect of the study and Primary Metropolitan Statistical Areas for another, a practice that should be considered highly questionable at best.

Urbanized Areas and Metropolitan Statistical Areas Redefined in 2000

The 1990s brought major reconsideration of both the Urbanized Area and Metropolitan Statistical Area definitions leading to new and very different definitions for both. Some of the impetus came from the availability of both better data and more powerful technology for manipulating that data that made possible approaches that previously would have been unfeasible. As the definitions had evolved over time, things had been added and the definitions had become more complex (especially the Metropolitan Statistical Area definition). The desire was to develop new, simpler definitions from scratch. This section discusses the new definitions, with emphasis on the problems I have with them. To summarize where this is going, a lot of good ideas… to a not so good effect.

New Urbanized Areas for Census 2000

The 2000 census brought a completely new approach for defining the urban population and urban areas (U.S. Bureau of the Census 2002). Advances in technology enabled the automation of aggregations of small areas meeting a minimum population density threshold. This allowed the elimination of the use of incorporated areas for the urban definition. Instead, urban areas were defined starting with a small core and adding adjacent blocks or block groups that had a population density greater than 500 persons per square mile. The lower density cutoff was chosen to include within the urban areas territory that was not completely residential but clearly urban, such as commercial and industrial areas. Under the prior definition, most of these areas had
been included within the boundaries of the incorporated areas. The Urbanized Area
definition had always included provisions for extending across gaps having lower
population densities, possibly with nonresidential uses. The new definition increased
the size allowed for some gaps and allowed more of them. These were termed “hops”
and “jumps.”

This new definition was not limited to the delineation of Urbanized Areas. It
applied to the identification of smaller urban areas as well. Areas of urban settlement
with populations of 50,000 or more were still called Urbanized Areas. Smaller areas with
populations of at least 2,500 were called Urban Clusters. The urban population of the
United States was then the population living in the Urbanized Areas and Urban
Clusters.

Overall, I see the new urban definition as a great improvement. The boundaries
of incorporated areas alway had an arbitrary aspect to them. The new definition
eliminates the problem of over bounded urban areas resulting from undeveloped, rural
land being included within the areas of cities. I really have only 2 problems with the
definition. The first is a minor one involving the nomenclature. The Urbanized Area
name was retained for urban areas with populations over 50,000, presumably to retain
continuity with the previous areas. The smaller areas then got the somewhat
unfortunately name of Urban Clusters. This suggests a fundamental difference between
the 2 types of areas, obscuring the fact that they are defined in exactly the same way.
The only difference is population size. I think it would have been clearer to just call
them all urban areas. More data could still be reported for those areas with larger
populations, just as the census does for other types of areas.

My second problem is with one other aspect of the Urbanized Area definition
that would prove to highly significant. The automated procedure combined urban areas
that became contiguous. In some cases, this resulted in very large areas of urban
territory that obviously should be considered to be multiple Urbanized Areas. The
major example here is the urban area along the eastern seaboard stretching from
Massachusetts to Delaware (a tiny piece actually is in Maryland, but it did not extend
down to the Baltimore area). Of course one would want New York and Philadelphia
(and other areas) to be identified as separate Urbanized Areas. So some criterion for
splitting the larger areas was required. The definition specified that areas would be split
near the boundaries of the previously defined Metropolitan Statistical Areas,
Consolidated Metropolitan Statistical Areas, or Primary Metropolitan Statistical Areas
when the distance along which the areas were contiguous was less than 3 miles. The
choice of 3 miles is interesting. The lengths of contiguity between the San Francisco-
Oakland and San Jose areas and the Los Angeles and Riverside-San Bernardino-Ontario
areas were both just under that. So these areas were split. Coincidence or conscious choice by the census?

Simply from the perspective of identifying areas of urban settlement, different people may certainly have varying options on whether certain areas should be seen as separate or combined. Being very familiar with the areas around the boundaries between the San Francisco and San Jose Urbanized Areas and the Los Angeles and Riverside Urbanized Areas I can say without qualification that they feel parts of unified urban areas. One does not have any sense of going from one urban area to another when crossing the border. To not have any of the Connecticut suburbs included in the New York Urbanized Area seems quite frankly silly. However, Urbanized Areas are not used that much and do not play that significant a role in how people perceive urban and metropolitan areas. The real problem arises in the use of these Urbanized Areas in the new definition of Metropolitan Statistical Areas, since the nature of that definition makes the extent of the Urbanized Areas critical, as will now be discussed.

New Metropolitan Statistical Areas First Delineated in 2003

During the 1990s, the Office of Management and Budget constituted the Metropolitan Area Standard Review Committee to reconsider how those areas should be defined. One major concern giving rise to this review was that the standards had become too complex and ad hoc, with the multiple combinations of metropolitan character and integration for the definition of Metropolitan Statistical Areas. Also, metropolitan settlement patterns were seen as evolving, making measures of metropolitan character less relevant. Very different alternatives were considered, some focusing on integration and commuting and others on metropolitan character and density, some continuing the use of counties as building blocks while others used census tracts (Office of Management and Budget 1998, 1999, 2000a).

The issue of using measures of both metropolitan character and integration was seen as a significant inconsistency in the current definition. Indeed, that question had been raised at least as early as 1969 in an evaluation of the metropolitan area definition in use at that time (Berry, Goheen, and Goldstein 1969). The argument was made in the new review that changing patterns of settlement in metropolitan areas, with more dispersed development, made the use of metropolitan character no longer appropriate.

---

3 The definition is actually ambiguous as to whether the distance of contiguity should be the single longest contiguous distance or the sum of the contiguous distances if contiguity occurs at multiple points. The census interpreted this as the former. In the San Francisco-San Jose case, the length of the minimum point of contiguity on the peninsula between San Francisco and San Jose was about 2.5 miles near Palo Alto, under the 3-mile cutoff. However, there was another point where the urban areas were contiguous on the east side of San Francisco Bay between Oakland and San Jose with a length of about 0.7 miles. Together, the total contiguous distance was about 3.2 miles, above the threshold.
Instead, the new definition should consider integration only, with a single uniform commuting standard and no consideration of metropolitan character. It did not actually do this, as will be explained after the presentation of the new definition, but this has never been acknowledged (U.S. Office of Management and Budget 1999, 2000a).

The new definition, which was first used to delineate Metropolitan Statistical Areas in 2003, continued to use counties as the basic units. Following the example and using the results of the new Urbanized Area definition, large cities were no longer used as the starting points. Rather, a Metropolitan Statistical Area would begin with the counties within which an Urbanized Area was located. Additional counties were added that had a commuting interchange with these central counties exceeding 25 percent (U.S. Office of Management and Budget 2000b).

Before proceeding further with the Metropolitan Statistical Areas, it is necessary to describe a new addition provided for in the definition. Counties that included Urban Clusters with populations from 10,000 up to 50,000 (the Urbanized Area minimum) would become the cores of Micropolitan Statistical Areas, and additional counties with commuting interchange above 25 percent could be added to these just like the Metropolitan Statistical Areas. Together the Metropolitan Statistical Areas and the Micropolitan Statistical Areas were termed Core Based Statistical Areas (CBSAs). (I have always felt that these were not the most felicitous choices for names.)

The final part of the definition addressed the merging of areas. Adjacent Core Based Statistical Areas (metropolitan or micropolitan) would be merged if they had a commuting interchange with each other exceeding the same 25 percent minimum. For micropolitan areas adjacent to Metropolitan Statistical Areas this was completely reasonable and consistent. Indeed, many of the counties adjacent to Metropolitan Statistical Areas that would have been added had the micropolitan areas not been defined would have first been identified as micropolitan areas because they included an urban cluster with a population of at least 10,000. So now such counties first became micropolitan areas, possibly in combination with 1 or more other counties, and then were added to the Metropolitan Statistical Area.

However, for the potential merger of adjacent Metropolitan Statistical Areas that may have had very significant levels of integration, the 25 percent threshold was a nearly impossible barrier. Unlike the smaller adjacent counties added to Metropolitan Statistical Areas, both of the adjacent Metropolitan Statistical Areas would have had major centers of employment. This would have made it highly unlikely that the commuting interchange among those Metropolitan Statistical Areas would exceed 25 percent, therefore making it highly unlikely that areas that were initially defined as separate Metropolitan Statistical Areas would be merged.

This gets to the serious problem with the new definitions. The initial delineation of the Metropolitan Statistical Areas was based on the Urbanized Areas. If there are separate Urbanized Areas, the areas start as separate Metropolitan Statistical Areas.
Then when the definition makes it very unlikely that adjacent Metropolitan Statistical Areas will be merged, the final Metropolitan Statistical Areas will likewise be separate. But what determined whether areas of continuous urban settlement are considered to be 2 or more Urbanized Areas, splitting those areas? Whether the distance along which areas were contiguous was less than 3 miles near the boundary of the Metropolitan Statistical Areas or Primary Metropolitan Statistical Areas as defined using the previous, 1990s definition. In other words, the new Urbanized Area and Metropolitan Statistical Area specifications had become circular definitions! The extent of Urbanized Areas depended on the extent of the older Metropolitan Statistical Areas and Primary Metropolitan Statistical Areas, and because of the rule making merger highly unlikely, the extent of the new Metropolitan Statistical Areas was dependent on the extent of the Urbanized Areas.

The result was a set of Metropolitan Statistical Areas with some areas divided and smaller compared to the last set of Metropolitan Statistical Areas defined using the previous definition. Because Metropolitan Statistical Areas are much more widely used than Urbanized Areas, however, the negative impact is greater. Eighteen areas that had been single Metropolitan Statistical Areas (or Consolidated Metropolitan Statistical Areas) under the previous definition were split into 2 or more Metropolitan Statistical Areas (Ottensmann 2016). Here are some examples of the problems with the new Metropolitan Statistical Areas:

- The new New York Metropolitan Statistical Area does not include any areas in Connecticut. While reasonable people might differ on how far the New York metropolitan area should extend into Connecticut, a delineation of the area that does not include any portion in that state is not reasonable. The Regional Plan Association considered areas in Connecticut to be part of the New York region as early as the 1920s (Regional Plan Association 2017).
- San Francisco-Oakland and San Jose became separate Metropolitan Statistical Areas though under previous definitions they had been part of a single area as of the 1990 census. Commuter rail service connects San Francisco with San Jose. Firms such as Google run buses to transport employees from residences in San Francisco to their headquarters in the San Jose Metropolitan Statistical Area. Silicon Valley is associated with the San Jose area, but Facebook and the venture capitalists on Sand Hill Road are in the San Francisco Metropolitan Statistical Area.
- The Riverside-San Bernardino-Ontario and Oxnard Metropolitan Statistical Areas are now separate from the Los Angeles Metropolitan Statistical Area though they also were part of the same area by the 1990 census. Four freeways having 4 or more lanes in each direction connect the Riverside Metropolitan Statistical Area to the Los Angeles Metropolitan Statistical Area and are at capacity during rush hours. Three commuter rail lines connect the areas. Ontario airport, obviously in
the Riverside Metropolitan Statistical Area, had literally been owned and operated by the Los Angeles airport authority for many years.

• Raleigh-Durham North Carolina had been a single Standard Metropolitan Statistical Area/Metropolitan Statistical Area from the 1970 through the 2000 censuses. Raleigh and Durham residents have long considered themselves part of a single area (along with Chapel Hill) focused on the Research Triangle Park in the middle. The area is served by a single commercial airport, Raleigh-Durham International. The new definition has Raleigh and Durham as separate Metropolitan Statistical Areas.

The metropolitan area definitions had documented the growth and evolution of metropolitan areas since 1950, including their combination as they grew together. By splitting many of the Metropolitan Statistical Areas defined for the 2000 census, the new definition rolled back the story of this evolution by decades (U.S. Bureau of the Census 1972, 1983, 1993).

Those responsible for creating the new definition obviously realized what they had done. They attempted to address the problem by defining a new type of area, Combined Statistical Areas (CSAs). (It’s looking more and more like the new definition, instead of simplifying things, only made the complexity worse.) Combined Statistical Areas were the combination of Core Based Statistical Areas (metropolitan or micropolitan) having a minimum of 15 percent commuting interchange. The Combined Statistical Areas turned out to be very similar to the Metropolitan Statistical Areas defined for 2000 using the previous definition (Ottensmann 2016).

An effort was made to downplay the use of the Combined Statistical Areas. At one point the census included the following admonition on its 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 2014)

Now if by this they meant one should not compare the population of a Combined Statistical Area with the population of another Metropolitan Statistical Area which is part of a Combined Statistical Area, sure. Making such a comparison would be wrong. It would be like comparing the population of a Primary Metropolitan Statistical Area with the population of another Consolidated Metropolitan Statistical Area. This would not make sense. But this seems so obvious that it should not be necessary for the census to issue such a warning. I am not aware of the census ever warning against comparing
the populations of Primary Metropolitan Statistical Areas with other Consolidated Metropolitan Statistical Areas.

On the other hand, if the intent is to suggest that one should not compare the population of the Combined Statistical Areas with the population of those Metropolitan Statistical Areas that have not been combined into Combined Statistical Areas, I think this advice is clearly wrong. The Combined Statistical Areas are areas created by combining 2 or more Core Based Statistical Areas having a commuting interchange of at least 15 percent. The remaining Metropolitan Statistical Areas are those where no areas met the criteria for such combination. To see the comparison as reasonable, just make this minor tweak to the Combined Statistical Area definition. Instead of saying combine 2 or more Core Based Statistical Areas, change the definition as follows: Take the set of all Core Based Statistical Areas, combine all those areas having a commuting interchange of at least 15 percent, and use the resulting set of areas. This will include the current Combined Statistical Areas and the remaining Core Based Statistical Areas that have not been not combined. Of course that resulting set could not be called Combined Statistical Areas because not all areas would have involved combinations. So call them something else—Extended Statistical Areas or Inclusive Statistical Areas or something like that.

The choice of the 15 percent commuting threshold was very interesting and revealing. This was the minimum cutoff used for the original Standard Metropolitan Area definition in 1950 in conjunction with measures of metropolitan character. This remained as the lowest threshold, in combination with the highest criterion for metropolitan character, through the previous definition used to define Metropolitan Statistical Areas for the 2000 census. In considering the merging of Metropolitan Statistical Areas, surely any Metropolitan Statistical Area would have easily met any criterion for metropolitan character.

In returning to the use of metropolitan character in the earlier definition of metropolitan areas, the claim that the new definition considered only integration, using a commuting standard, and not metropolitan character, is not correct. To be sure, commuting was the only criterion used to add counties to the initial central counties and subsequently for the merger of areas. But the original set of counties for starting a Metropolitan Statistical Area consisted of those counties in which an Urbanized Area was located. And what is an Urbanized Area based on other than measures of urban or metropolitan character? Indeed, one of the criteria always considered for metropolitan character in the earlier metropolitan definitions was density, the basis for defining the Urbanized Areas. Defenders of the new definition might well respond that obviously the set of counties containing the Urbanized Area are integrated. In a general sense, I would agree. But that is not good enough. Integrated according to what criterion? The new Metropolitan Statistical Area definition uses as the standard for integration having a minimum commuting interchange of 25 percent. I believe it is highly likely that for
those new Metropolitan Statistical Areas that could reasonably be divided into 2 or more areas, each with a large urban center, some would have commuting interchanges less than that minimum. From that same perspective, if some of the contiguous Urbanized Areas had not been divided into separate areas and the combined areas had been considered a unified Urbanized Areas, then new definition would have considered those to be single Metropolitan Statistical Areas, even though the commuting interchanges were definitely less than 25 percent (otherwise they would have been merged into a single Metropolitan Statistical Area). Also, for those Metropolitan Statistical Areas consisting only of the central counties defined by the Urbanized Areas with no outlying counties added, integration plays no role at all in their definition (except, of course, in the negative sense of an absence of integration with adjacent counties.

Not all aspects of the new Metropolitan Statistical Area definition are bad. Not starting the definition with incorporated places makes sense here, as it does for the Urbanized Areas. The decision was finally made to use counties as the building blocks for Metropolitan Statistical Areas in New England as well as in the rest of the country, which produces welcome uniformity. Town-based areas were unfortunately retained, but only as a secondary type of area with a different name rather than being used for defining the Metropolitan Statistical Areas themselves. A smaller number of Metropolitan Statistical Areas were subdivided and mercifully those areas were called Metropolitan Divisions rather than the misleading Primary Metropolitan Statistical Areas. It still produced sets of areas that were often strange, including continuing Nassau-Suffolk as a subdivision of the New York Metropolitan Statistical Area. But at least it was no longer being called a metropolitan area. The new definition used similar criteria to those used in the previous revised central city definition that added numerous central cities. However, the new standard now calls these principal cities, which is less confusing and is more descriptive.

I do not believe that the creation of the Micropolitan Statistical Areas constituted a useful addition. I have not seen the areas used for many analyses beyond those done by the census. Having areas defined exactly like the Metropolitan Statistical Areas except for being based on Urban Clusters and then having distinct names obscures the similarity. However, I see the problem given a desire to specifically identify metropolitan areas. One issue addressed by the Metropolitan Area Standards Review Committee was whether a system should be established that would subdivide the entire area of the country. This does not do that. It created the Core Based Statistical Areas, metropolitan and micropolitan. The remainder of the counties are then what? The rest of the country? They cannot be called rural, because that has an entirely different meaning with the urban-rural definition. In addition, the Micropolitan Statistical Area definition creates yet another distinction among urban areas. All areas of contiguous urban settlement having populations of 2,500 or more are considered urban. Those with
populations of 50,000 or more are Urbanized Areas. The Urban Clusters with populations from 10,000 to 50,000 can form the core of a Micropolitan Statistical Area. So this creates a third category of urban areas, Urban Clusters with populations ranging from 2,500 to 10,000. What are these? Minor Urban Clusters?

I have only been able to find one piece in the scholarly literature addressing the new Metropolitan Statistical Area definition, by a group at Brookings (Frey et al. 2006). To make the comparison with the older Metropolitan Statistical Areas, they began with the premise that it had been better to use the Primary Metropolitan Statistical Areas rather than the Consolidated Metropolitan Statistical Areas for metropolitan analysis. So of course they saw no problem with the new Metropolitan Statistical Areas being too small, amazingly concluding that the new New York Metropolitan Statistical Area definition “is arguably more satisfying than the old one.” Of course! The old New York Primary Metropolitan Statistical Area included only New York City and counties in New York State north of the Bronx. So the new Metropolitan Statistical Area certainly comes a lot closer to encompassing the New York metropolitan area. But it still does not include any of the Connecticut suburbs.

The paper goes out of its way in support of the new Metropolitan Statistical Area definition. They discuss the splitting of numbers of the old Metropolitan Statistical Areas into multiple new Metropolitan Statistical Areas, specifically identifying the breakup of the Raleigh-Durham Metropolitan Statistical Area. They correctly note that the division reflects the higher commuting thresholds for metropolitan integration under the new definition. But they go on to add without any evidence at all that this also might reflect “an emerging economic independence separating formerly close-knit neighbors.” As the long-term trends have been towards increasing integration of ever-larger metropolitan areas over time, asserting this without any evidence is truly bizarre and unsupportable.

In my searches for commentary critiquing the new Metropolitan Statistical Areas, I did uncover discussions of the issue in a surprising venue—among those who work to edit Wikipedia pages. The most-spirited involved New York. At issue was how the metropolitan area should be defined for the New York metropolitan area entry (Wikipedia 2019). Should the (current) Metropolitan Statistical Area definition be used or should the New York metropolitan area be defined in the more generally accepted way that included areas in Connecticut? I found this discussion to be more thoughtful than anything I had encountered in the scholarly literature.

The End of Evolution in 2010?

This section is brief because there were few changes. And that is what I see as the serious problem with the 2010 Urbanized Area and Metropolitan Statistical Area
definitions. The actions taken (perhaps better considered the failure to act) has to be the worst decision in the history of the urban and metropolitan definitions.

*The Freezing of Urbanized Areas*

Most of the changes to the urban area definition for 2010 were fairly minor, technical alterations. Perhaps the most significant was the use of impervious surface data from the National Land Cover Dataset for identifying nonresidential urban territory. The significant issue for 2010 remained the question of the splitting of urban areas (U.S. Bureau of the Census 2010, 2011, 2012; Ratcliff 2015).

In the run up to the 2010 census, the Census Bureau published the proposed urban area definition for comment (U.S. Bureau of the Census 2010). With respect to the issue of splitting, the census proposed no splitting of areas with populations less than a million. Larger areas were to be split along Metropolitan Statistical Area lines. Note that any consideration of the distance of contiguity was eliminated. (If they had retained the 3-mile limit for splitting, San Francisco-San Jose and Los Angeles-Riverside would no longer have been split, as the lines of contiguity had come to exceed 3 miles in 2010 due to additional urban growth. So much for that having been a logical criterion.) The notice did include this admission:

> This approach, however, results in some circularity of outcomes—the metropolitan statistical area and NECTA definitions that would be used to split large agglomerations are those that were defined on the basis of Census 2000 data, including Census 2000 urban area definitions; the 2010 Urbanized Areas resulting from the splitting process will form the cores of metropolitan statistical areas and NECTAs.

But they leave it with that. Nothing further is said about this. Certainly no suggestion is made that this presents a problem that needs to be addressed, either in the final 2010 urban area criteria or sometime in the future.

It was the proposal to not split the smaller areas that raised a furor and many objections to the proposal. This would have meant that some previously separate Urbanized Areas would have been combined. And Urbanized Areas had formed the basis for establishing the federally mandated metropolitan planning organizations (MPOs) for transportation planning and the receipt of federal transportation funds. Fearing they might see their MPOs merged out of existence, local officials and their Congressional representatives raised strong opposition.

In response to the uproar, the Census Bureau surrendered. The final urban area definition froze in place the Urbanized Areas that were delineated at the 2000 census. New territory could be added, of course, and new Urbanized Areas could be added
when populations reached 50,000, but any Urbanized Area existing in 2000 would continue to be a separate Urbanized Area in 2010 as long as its population had not dropped below the 50,000 minimum. The final notice justified this by saying, “Adoption of these criteria will facilitate continuity and comparability between the two decades’ urban definitions” (U.S. Bureau of the Census 2011).

This freezing of the set of Urbanized Areas was unprecedented. For the entire history of the Urbanized Area and Metropolitan Statistical Area definitions, the areas had been allowed to evolve and sometimes merge as urban areas grew together. Had this approach freezing the Urbanized Areas been adopted when Urbanized Areas were first defined, Dallas and Fort Worth would have remained separate Urbanized Areas. And, of course since the Urbanized Areas were used in the definition of Metropolitan Statistical Areas, the effect spilled over to those areas as well.

*Metropolitan Statistical Areas Just Follow Along*

The metropolitan (and micropolitan) area definitions released in 2010 changed little from those put in place a decade earlier (U.S. Office of Management and Budget 2010). With their extent and whether adjacent areas were combined remaining dependent on the Urbanized Areas and with the Urbanized Areas being frozen, the set of Metropolitan Statistical Areas was frozen as well. To emphasize the sequence producing the Metropolitan Statistical Areas delineated in 2013, consider this: Whether adjacent areas were combined or separate Metropolitan Statistical Areas depended on whether the 2010 Urbanized Areas were combined or separate. Whether the 2010 Urbanized Areas were combined or separate depended on whether the 2000 Urbanized Areas were combined or separate because the extent of the Urbanized Areas was frozen. And whether continuous urban areas were combined or separate for the 2000 Urbanized Areas depended on whether the maximum distance of contiguity was less than 3 miles and on the boundaries between Metropolitan Statistical Areas, Consolidated Metropolitan Statistical Areas, and Primary Metropolitan Statistical Areas delineated using 1990 census data and the 1990 metropolitan definition. So the Metropolitan Statistical Areas delineated in 2013 depended, in part, on the 1990s Primary Metropolitan Statistical Areas.

The 2000 Metropolitan Statistical Area definition rolled back the process by which the metropolitan areas documented the growth and evolution of metropolitan areas and their combination as they grew together. With the freezing of the Urbanized Areas and the continued dependence of Metropolitan Statistical Areas on the Urbanized Areas, the possibility of the Metropolitan Statistical Areas capturing this evolution has ended, at least for now. The Metropolitan Statistical Areas are now likewise essentially frozen, with only changes to outlying counties possible.
A Modest Proposal

The key problem with the current Urbanized Area and Metropolitan Statistical Area definitions is their circular nature, with one dependent on the other. This is compounded by the use of the higher 25 percent commuting interchange standard for the combination of adjacent Metropolitan Statistical Areas rather than the lower 15 percent standard which had been used, at least for those areas with the greatest metropolitan character, since the start of the county-based metropolitan area definition in 1950.

Many of the ideas from the new Urbanized Area and Metropolitan Statistical Area definitions are very good. Defining the Urbanized Areas using only the densities of small areas rather than populations of incorporated areas makes sense. Identifying central counties of Metropolitan Statistical Areas using Urbanized Areas is a very reasonable approach, again eliminating dependence on incorporated areas. And using only a commuting interchange standard for the addition of outlying counties is a welcome simplification, with the 25 percent threshold being quite reasonable given the potentially larger sets of central counties with the new approach.

So how might one retain these features in defining the Urbanized Areas and Metropolitan Statistical Areas without resorting to either circular definitions or highly arbitrary criteria or both? I would like to make some general suggestions for how this might be accomplished. This is not intended to be a detailed definition nor am I certain it would even work. One would have to go through the process of attempting to implement it to see what refinements or details might be needed or even if it is workable at all.

The process would start with those areas delineated as urban as they are in the Urbanized Area definition but without contiguous areas being split. This yields some very large areas, most notably the Northeast urban corridor. Those counties containing these urban areas are then considered to be the central counties of Metropolitan Statistical Areas, as they are in the current definition, though with no specification yet regarding which Metropolitan Statistical Areas.

The next step is where the central counties get split up for individual Metropolitan Statistical Areas. This is where the circularity of the current definitions is avoided. This is also the point where I am not certain that the process would be workable in practice. To illustrate the proposed process, the example of establishing the boundary between the New York and Philadelphia Metropolitan Statistical Areas will be considered. Presumably those counties in New Jersey close to New York would have higher proportions of those commuting out of the county going towards New York rather than Philadelphia. Likewise, counties closer to Philadelphia would have higher proportions commuting in that direction. Then presumably (hopefully) between New York and Philadelphia will be a location where adjacent counties to the north have
slightly higher proportions commuting towards New York and the ones to the south have more commuting in the direction of Philadelphia. This would be the boundary between the central counties of the (potentially) separate Metropolitan Statistical Areas. Note that this process will almost certainly break apart areas that we currently consider (and have long considered) single Metropolitan Statistical Areas, such as Dallas and Fort Worth. Hence the designation of these as potentially separate Metropolitan Statistical Areas.

All sorts of issues are posed by this proposal. Commuting out of a county can be to many different destinations (presumably counties would be used as destinations; remember separate Metropolitan Statistical Areas have not yet been defined). How is direction of out-commuting defined? Will the locations of shifts in directions form a consistent boundary? What about inconsistencies such as a county closer to Philadelphia that has more commuting in the direction of New York or vice versa? Can these be handled in a reasonable and not completely arbitrary way?

Assuming this can be made to work, groups of counties have now been identified that are central counties of potential Metropolitan Statistical Areas. The term potential is being used because it may still be the case that some of these areas would be merged to create the final Metropolitan Statistical Areas. Outlying counties are then added to these potential Metropolitan Statistical Areas using the current criterion of 25 percent commuting interchange with the central counties. Note that this is referring to all adjacent counties and nothing is to be done yet regarding Micropolitan Statistical Areas. For getting the Metropolitan Statistical Areas right, it is better to be considering the outlying counties for addition separately from micropolitan areas. Allowing the simultaneous creation of the micropolitan areas can create situations where 2 or more counties are combined into a single Micropolitan Statistical Area which is then later considered for merger with a Metropolitan Statistical Area. In those instances, it is possible that one county would meet the commuting interchange criterion for being added to the Metropolitan Statistical Area and the other county would not. The 2 counties taken together as a micropolitan area might or might not meet same commuting criterion for merger.

Once all of the potential Metropolitan Statistical Areas have been formed, adjacent areas having a commuting interchange of at least 15 percent would be merged. Some of those areas that were separated because commuting flows in opposite directions will be merged. This will also produce the merger of some Metropolitan Statistical Areas where the Urbanized Areas are not contiguous but which have this same level of integration. (Raleigh and Durham in 2000 would have been an example.) This would produce the final set of Metropolitan Statistical Areas.

Now to continue the identification of the Urbanized Areas. The areas of continuous urban settlement had been identified. Split those areas at Metropolitan Statistical Area boundaries to produce the final set of Urbanized Areas.
Using the 15 percent criterion for the merger of these areas (or some other threshold less than 25 percent) is a critical aspect of this proposed definition. Some areas currently considered single Metropolitan Statistical Areas would undoubtedly be split due to changes in the direction of commuting. Just as the 25 percent threshold for merger kept separate some previously unified Metropolitan Statistical Areas, applying the same standard in this cases would likely allow others to remain split, worsening the problems associated with the current Metropolitan Statistical Area definition. If this can be made to work, it would provide a logical basis for establishing where both Metropolitan Statistical Areas and Urbanized Areas can be separated—at the location the dominant direction of commuting shifts. How a county interacts with other counties would determine membership in a potential Metropolitan Statistical Area, not a boundary drawn previously.

References


Ratcliff, Michael. 2015. A century of delineating a changing landscape: The Census Bureau’s urban and rural classification, 1910 to 2010. by Michael Ratcliffe. Presented at the Social
changes to the standards for defining metropolitan areas. *Federal Register* 64, 202 (October 22): 56628-56644.


