RESEARCH NOTE: TOWARD AN EFFECTIVE SUBCOUNTY SETTLEMENT CLASSIFICATION: COMPARATIVE DENSITY REVISITED

Darryl T. Cohen and Keith G. Debbage
Department of Geography
University of North Carolina at Greensboro

Abstract: The federal Office of Management and Budget (OMB) defines metropolitan and micropolitan statistical areas (MSAs) for use by federal agencies in collecting and publishing statistics. In order to ensure their continued usefulness, OMB reviews and revises the standards for defining MSAs at least once every 10 years, typically prior to each decennial census. The most recent review was one of the most thorough and introduced some significant changes. Whereas the review produced a revised set of standards in December, 2000, OMB has encouraged continued research on a number of issues that remain open for discussion. These issues include: (1) the geographic building blocks used in defining statistical areas, (2) the creation of a settlement classification that encompasses all of the nation’s territory, and (3) the categories of settlement (e.g., urban, suburban, rural) recognized within the classification. This note documents continuing research on how these issues might be addressed by using subcounty building blocks in a comparative population-density-based approach to classifying settlement.

Background on OMB’s Review of the MSA Standards

The Office of Management and Budget’s (OMB’s) classification of metropolitan and micropolitan statistical areas (MSAs) is one of the most widely used frameworks for understanding and analyzing settlement in the United States. The statistical agencies of the federal government (including the U.S. Census Bureau, Bureau of Economic Analysis, Bureau of Labor Statistics, and Bureau of Transportation Statistics) collect and publish a large volume of MSA data on a regular basis. These data are used by state and local governments, chambers of commerce, as well as private sector firms and academic think tanks in a wide variety of research and policy applications. To ensure the quality of these data, OMB conducts a review of its standards for defining MSAs at least once every 10 years, usually prior to each decennial census. The most recent review, which began in the early 1990s and concluded in December, 2000, focused on a number of key conceptual issues regarding the classification of settlement. While the final standards were announced at the end of 2000, OMB has encouraged additional research on some of the issues explored during the review. The research documented here seeks to address three of those issues.

1 Correspondence concerning this article should be addressed to Keith G. Debbage, Department of Geography, University of North Carolina at Greensboro, Greensboro, NC 27402-6170; telephone: 336-334-3911; fax: 336-334-5864; e-mail: kgdebbag@uncg.edu
The first of these issues is the geographic building block, or basic spatial unit of analysis used in defining areas. Since their inception in the late 1940s, MSAs have relied on the county as the principal building block, except in New England where minor civil divisions (MCDs)\(^2\) have been used. Each MSA, therefore, consists of one or more whole counties.\(^3\) Numerous data users have expressed concerns about the problems associated with using counties as building blocks, most notably the large geographic size of counties in some parts of the country, particularly in the western states (OMB, 1998). As a result of the use of counties, some MSAs include large areas of sparsely-settled territory that might otherwise not be considered metropolitan in character. This problem is commonly referred to by the term “overbounding.” Though the use of building blocks smaller than counties can, in many cases, allow for improved resolution in defining areas of different settlement types, OMB has continued to rely on counties mainly due to the wider variety of statistical data available at the county level than at any subcounty level of geography, such as census tracts, county subdivisions, and ZIP code areas (OMB, 2000). OMB has, however, encouraged additional research on subcounty approaches:

Interested Federal agencies should continue research on settlement patterns below the county level to describe further the distribution of population and activity throughout the Nation. (OMB, 2000, p. 82235)

The second conceptual issue is the proportion of the U.S. territory that can be meaningfully categorized within a statistical classification of settlement. Roughly 20% of the nation’s territory is included within metropolitan areas. Prior to the introduction of the new micropolitan statistical areas in 2003, the remaining 80% was designated as nonmetropolitan and remained undifferentiated by OMB’s classification (U.S. Census Bureau, 2000). Similarly, the Census Bureau’s urban and rural classification identifies about 2.6% of the U.S. territory as urban, while the remaining 97.4% is classified as rural (U.S. Census Bureau, 2000). Some data users have expressed dissatisfaction with the dichotomous nature of such approaches, and have shown interest in the development of more inclusive approaches that meaningfully address a greater range of the nation’s settlement patterns (OMB, 1998).

The third issue that has remained open for further research is the identification of a number of distinct categories of settlement. Although terms such as “suburban” and “exurban” are commonly used to describe areas of varying settlement intensity, there is no universally accepted statistical definition of what constitutes areas of these types. Data users have suggested that a single quantitative definition of such settlement types would be useful (OMB, 1998).

\(^2\)MCDs are legal, political, or administrative subdivisions of counties. See research design section for a more detailed discussion.

\(^3\)Since the June 2003 implementation of the December 2000 criteria for defining metropolitan and micropolitan statistical areas, counties have been used as building blocks nationwide, including in New England. Definitions of metropolitan areas used in this study are as of June 20, 1990, and include MCD-based areas in the New England division.
Background on the Comparative Density Approach to Classifying Settlement

During the most recent review of the MSA standards, OMB encouraged the development of alternative approaches to classifying settlement in the United States. At the outset of the review, four studies were produced under joint statistical agreements between the Census Bureau and four universities. One of these studies, prepared by John Adams of the University of Minnesota, employed a methodology that has come to be known as the comparative density approach (Adams, 1995). This approach was further explored by OMB in 1998, revisited by Adams and his colleagues in 1999, and expanded upon by Glenn Nelson (a consultant working under contract with the Census Bureau) in 2001.

One of the key assumptions underlying the comparative density methodology is that population density alone serves as a surrogate measure for other features of structure and interaction. According to Adams and his colleagues, adjacent areas of high density are inevitably accompanied by social and economic linkages:

It seems to us that high county population densities and the proximity of densely settled neighboring counties inevitably are accompanied by linkages of various kinds such as commuting, shopping, and intraurban migration. Thus, it seems reasonable that residential population density itself can be used as an appropriate and accurate surrogate for linkage. (Adams et al., 1999, p. 706)

This approach relies solely on population density data to identify areas of varying settlement intensity. This is accomplished by ranking areas (Adams et al. used counties in their study) according to their overall residential population density and calculating an average of two percentile ranks for each area. The first percentile rank for each area reflects its density rank within the nation. The second reflects its density rank within some subnational context (Adams et al. used state rankings in their study). The average percentile rank (APR) is used in order to take into account both the national and regional significance of each area. Once the APR scores are assigned, areas can then be classified into a variety of settlement categories.

RESEARCH DESIGN

Working with Subcounty Geography

While the results of applying Adams’ comparative density methodology at the county level have been well documented (Adams, 1995; OMB, 1998; Adams et al., 1999; Nelson, 2001), little attention has been devoted to exploring its vis-à-vis effectiveness with subcounty geography. Whereas numerous attempts to classify the U.S. settlement system using ZIP code areas and census tracts have been made (Berry, 1995; Morrill et al., 1999; Rain, 1999), the county subdivision has been under-researched as a potential unit of analysis in recent years. The research presented here documents the results of applying the comparative density methodology at the county subdivision level, using population density data from the 1990 decennial census (Cohen, 2002).

What is a county subdivision? Each of the nation’s 3141 counties is subdivided into a number of legal, political, administrative, or statistical units, generally referred to as
county subdivisions. There are two main types of county subdivisions—minor civil divisions (MCDs) and census county divisions (CCDs). In 1990, there were 35,298 county subdivisions, including 29,717 MCDs and 5621 CCDs and other statistical entities (U.S. Census Bureau, 1994). MCDs are legally defined entities whose boundaries and functions are prescribed by state law. In 20 states, mainly in the Northeast and Midwest, MCDs function as active units of local government and have powers to raise revenue and provide services. In eight other states, MCDs are inactive administrative units, such as election precincts or other such districts. In the remaining 21 states where MCDs are either non-existent or inadequate for statistical use, the Census Bureau, in cooperation with state and local officials, has defined CCDs for the collection and presentation of census data at the county subdivision level. Unlike MCDs, CCDs are statistical entities whose boundaries are designed to be stable and easily described, though not provided for under state law (U.S. Census Bureau, 1994).

Why use county subdivisions? The chief advantage of the county subdivision as a geographic building block is that it generally provides greater spatial resolution than the county. The smaller geographic size of county subdivisions allows for a more precise identification of the boundaries between densely and sparsely settled areas than is possible using county-based approaches. A second advantage is the wide variety of consistent, publicly available statistical data for county subdivisions compared to other sub-county units such as census tracts and ZIP code areas. Information on a variety of population and housing topics, including income, occupation, and age of housing stock, are available for all county subdivisions in the decennial census of population and housing. Moreover, an increased range of data items (e.g., journey to work by specific origin and destination) is available for MCDs in New England. Finally, county subdivisions have the advantage of being familiar, recognizable geographic units in some states, and they are typically fewer in number than other sub-county units such as census tracts and ZIP code areas. It has been noted that tracts and ZIP code areas are often too small, too numerous, and "would yield more cumbersome detail and precision than is useful" in national-level analyses (Adams et al., 1999, p. 709). Exceptions to this include low population areas, typically in the West North Central division (e.g., Nebraska, Iowa, etc.), where county subdivisions are sometimes greater in number and smaller in area than census tracts, which are, in some cases, as few as one or two per county.

Disadvantages of county subdivisions. One of the main drawbacks of using county subdivisions as geographic building blocks is the instability of their boundaries in some parts of the country. MCDs in the Northeast are generally quite stable. CCDs, which are statistical entities, generally have stable boundaries as well. In certain parts of the country, however, boundary changes due to annexations, mergers, and administrative redistricting are common among MCDs. Another noteworthy disadvantage of using county subdivisions is the more limited variety of data available for them when compared to counties. Some survey and sample data items (such as journey to work by origin and destination from the decennial census), which are available for all counties, are available only for those MCDs in New England. Despite these drawbacks, OMB has noted the potential utility of county subdivisions, stating that:

Despite variations in population size and stability of boundaries for some MCDs and CCDs, county subdivisions could provide a compromise between the
disadvantages posed by the geographic extent of counties and the more limited availability of economic data for some other subcounty geographic units. (OMB, 1998, p. 70539)

Using the Comparative Density Approach

Perhaps the most notable advantage of using the comparative density approach is its ease of implementation. Population density data are widely available and easily obtained for a variety of geographic units. The use of the percentile rank rather than the overall population density yields the added advantage of an evenly distributed data set that is free of skew and kurtosis. This overcomes the difficulties (such as the setting of class breaks between categories) associated with the uneven distribution of overall population density data, where the vast majority of observations have very low values and a relatively small number of observations (in and around such large cities as New York and Chicago) have extremely high values.

Using the average of a national and a subnational percentile rank ensures that medium-sized centers of regional importance are distinguished, because they will rank considerably higher within their regional context than within the nation. While Adams’ original (1995) study used a state ranking, the results were presented for only a selected number of states. Additional research by Nelson (2001) has shown that the use of a Census Division⁴ rank yields more useful results for nationwide analysis. Following Nelson’s work, this study uses the percentile rank within the nation and within the Census Division to calculate an APR for each MCD or CCD. For example, the Miles City, Montana CCD has a population density percentile rank of 57.21 within the nation, compared to 83.53 within the Mountain division. The average of the two percentile ranks yields a value of 70.37, which reflects both the national and local importance of Miles City in the settlement hierarchy.

Assigning Settlement Categories

As noted by Adams et al. (1999), the population density or settlement intensity within a given area is only significant relative to the densities of other areas with which it is compared. Consequently, there are no absolute levels of population density that are inherently meaningful as thresholds among settlement types. The flexibility of the comparative density approach allows the user to group observations into any number of classes or categories needed to facilitate particular research needs.

For evaluative purposes, and for the sake of simplicity, a classification of three equal intervals is used here. Those county subdivisions in the highest third of the APR distribution (with values from 66.67 to 100) are classified as densely settled. County subdivisions in the middle third (33.33 to 66.66) are classified as moderately settled, and county subdivisions in the lowest third (0.01 to 33.33) are classified as sparsely settled. These class breaks produce a relatively comparable number of observations in each class (Table 1).

---

⁴Census Divisions are statistical groupings of states defined by the Census Bureau. For a list of the nine Census Divisions and their definitions, see http://www.census.gov/geo/www/reg_div.txt
Finally, the proportion of land in densely, moderately, and sparsely settled areas identified by the comparative density approach is calculated for each MSA and CMSA\(^5\) as defined by OMB in 1990. This allows for the identification of those metropolitan areas that contain the largest and smallest proportions of densely and sparsely settled land, which can provide insights as to where the use of county subdivisions can have the greatest impact in addressing the problem of overbounding.

**FINDINGS**

*Population and Land-Area Distribution by Settlement Type*

Table 1 shows the distribution of population in 1990 categorized according to the three settlement types identified using the comparative density approach. Comparing these population figures with the land area figures in Table 2 shows that the majority of the

\(^5\)The metropolitan areas used in this study included metropolitan statistical areas (MSAs) and consolidated metropolitan statistical areas (CMSAs).
population is clustered on a relatively small proportion of land. Densely settled areas, which contain 85% of the population, account for only 11% of the nation’s land area. At the other end of the spectrum, only 3% of population is included in sparsely settled areas, which account for more than 60% of the U.S. land area. The geographic distribution of densely, moderately, and sparsely settled areas is shown in Figure 1.

**Comparative Density vs. Metropolitan Areas**

As shown in Table 2, only about 42% of the land area classified by OMB as metropolitan in 1990 is classified as densely settled using the comparative density approach at the county subdivision level. About 32% of metropolitan land is classified as moderately settled, and about 26% is sparsely settled. This finding suggests that the comparative density approach, when applied at the county subdivision level, can be useful in addressing the problem of overbounding by differentiating between the densely, moderately, and sparsely settled components of large metropolitan counties (e.g., Yuma County, AZ and San Bernardino County, CA).

Table 2 also shows that 5.5% of land classified as nonmetropolitan by OMB in 1990 is identified as densely settled using the comparative density approach. These are areas in

---

**Fig. 1.** Settlement in the United States, 1990. Densely, moderately, and sparsely settled areas identified using the comparative density method at the county subdivision level. *Source:* 1990 census and data compiled by authors.
and around small- and medium-sized cities, mainly in the South and West, that do not meet the 50,000 population threshold for metropolitan status.

An examination of the distribution of densely, moderately, and sparsely settled land for each metropolitan area reveals that metropolitan areas with the largest percentages of sparsely settled land tend to be located in the western part of the country, where counties are typically larger in land area (Fig. 2). The 10 metropolitan areas with the largest percentages of sparsely settled land are listed in Table 3. All 10 of these MSAs are located to the west of the Mississippi River.

**CONCLUSION**

The overall results indicate that a comparative-density-based, county-subdivision-level approach to classifying settlement can be effective in representing the settlement system of the United States in statistical form. The findings show that such an approach can meaningfully categorize all of the nation’s territory, while avoiding in many cases the problem of overbounding that often occurs with the use of counties. Whereas specific settlement types such as suburban and exurban may often require even greater geographic resolution than is provided at the county subdivision level, a comparative density
approach allows for the creation of as many settlement categories as the user requires, eliminating the constraints of dichotomy associated with metropolitan/nonmetropolitan, and urban/rural.

Future avenues of research should include the exploration of more refined settlement categories, and the development of a methodology for defining boundaries around individual statistical areas. Further, application of the comparative density approach to smaller scale geography, such as census tracts and ZIP code areas, may allow for the exploration of more detailed settlement types such as urban, suburban, and exurban areas. Finally, a replication of the approach presented here using 2000 census data would provide an opportunity for time-series analysis, which could yield much needed information regarding the temporal utility of the classification, especially with regard to changes in county subdivision boundaries and the changing economy.

REFERENCES


<table>
<thead>
<tr>
<th>FIPS code</th>
<th>Metropolitan area</th>
<th>Land area (sq. mi.)</th>
<th>Densely settled</th>
<th>Moderately settled</th>
<th>Sparsely settled</th>
<th>In urbanized area</th>
<th>Percent sparsely settled</th>
</tr>
</thead>
<tbody>
<tr>
<td>4080</td>
<td>Laredo, TX MSA</td>
<td>3,357</td>
<td>29</td>
<td>0</td>
<td>3,328</td>
<td>32.9</td>
<td>99.14</td>
</tr>
<tr>
<td>4120</td>
<td>Las Vegas, NV MSA</td>
<td>7,911</td>
<td>452</td>
<td>0</td>
<td>7,459</td>
<td>231.1</td>
<td>94.29</td>
</tr>
<tr>
<td>9360</td>
<td>Yuma, AZ MSA</td>
<td>5,514</td>
<td>408</td>
<td>0</td>
<td>5,106</td>
<td>26.3</td>
<td>92.60</td>
</tr>
<tr>
<td>1010</td>
<td>Bismarck, ND MSA</td>
<td>3,561</td>
<td>104</td>
<td>191</td>
<td>3,266</td>
<td>36.5</td>
<td>91.72</td>
</tr>
<tr>
<td>6690</td>
<td>Redding, CA MSA</td>
<td>3,786</td>
<td>322</td>
<td>0</td>
<td>3,464</td>
<td>61.5</td>
<td>91.49</td>
</tr>
<tr>
<td>6720</td>
<td>Reno, NV MSA</td>
<td>6,342</td>
<td>658</td>
<td>141</td>
<td>5,543</td>
<td>93.2</td>
<td>87.40</td>
</tr>
<tr>
<td>6660</td>
<td>Rapid City, SD MSA</td>
<td>2,774</td>
<td>134</td>
<td>228</td>
<td>2,412</td>
<td>46.7</td>
<td>86.95</td>
</tr>
<tr>
<td>1350</td>
<td>Casper, WY MSA</td>
<td>5,340</td>
<td>26</td>
<td>786</td>
<td>4,528</td>
<td>27.9</td>
<td>84.79</td>
</tr>
<tr>
<td>9260</td>
<td>Yakima, WA MSA</td>
<td>4,298</td>
<td>755</td>
<td>151</td>
<td>3,392</td>
<td>37.4</td>
<td>78.92</td>
</tr>
<tr>
<td>0860</td>
<td>Bellingham, WA MSA</td>
<td>2,119</td>
<td>394</td>
<td>79</td>
<td>1,646</td>
<td>30.0</td>
<td>77.68</td>
</tr>
</tbody>
</table>

Source: Data compiled by authors based on 1990 census data. Metropolitan areas defined by the Office of Management and Budget as of June 30, 1990.


