



Date: December 22, 2005

TECHNICAL  
MEMORANDUM FOR: Robert A. Kominski   
Assistant Division Chief, Social and Demographic Statistics  
Population Division, U.S. Census Bureau

From: Doug Gevert   
Analyst, Social and Demographic Statistics  
Population Division, U.S. Census Bureau

Subject: Review of NCES school locale tabulation and analysis

### **Overview**

The National Center for Education Statistics' Common Core of Data (CCD) provides an indicator of school location as part of its annually updated universe of public schools and school districts. These locale assignments provide a general indicator of the type of geographic area or community where the school or district is located. NCES has periodically revised the locale classifications, criteria, and assignment methods, but the basic framework has been relatively stable since its inception in the late 1980s. Recent changes to the Census Bureau's urban criteria and to the Office of Management and Budget's metropolitan area standards (the introduction of core based statistical areas (CBSAs) composed of metropolitan Statistical Areas and micropolitan statistical areas) have complicated the way NCES assigns community locale indicators to schools and school districts. Although NCES has incorporated new terms into its locale criteria, these accommodations have not addressed inherent conflicts in the current locale framework. As part of their on-going effort to improve data quality, NCES requested that we review the technical and conceptual elements that define the CCD locale framework and recommend potential changes to improve the technical consistency, conceptual coherence, and analytic utility of these geographic classifications. This technical memorandum summarizes key findings and recommendations provided to NCES during our on-going project discussions, and it offers a brief analysis of the likely effects the recommended changes would have on the current locale distribution for CCD schools, school enrollment, and school districts. All results were based on custom tabulations and spatial analyses developed for the project.

### **Project Rationale**

Geography is one of the primary organizing features of U.S. public education. States are subdivided into school districts, school districts are subdivided into individual school attendance areas, students are typically assigned to schools based on place of residence, and local property taxes that fund most public school systems are assessed on property within specific geographic boundaries. Research and policy discussions about desegregation, school finance, school choice, and other educational issues reinforce the centrality of geography by making frequent comparisons between city, suburban, and rural schools. Discussants may disagree on equitable solutions to these policy issues, but they do tend to agree with a simple premise: When it comes to schooling, location matters.

Classification as an exercise does little to explain how or why educational systems work or what policies ought to be adopted in particular contexts, but it is an essential component for research (Bailey, 1994), and essential to NCES's mandate to analyze and report on the condition of education in the U.S.. The practical implications of locale classification and assignment are considerable for administrators, policy makers, and researchers. Federal education programs and private foundations frequently rely on the NCES locale assignments as a criterion for funding eligibility, therefore if school locale assignments are inaccurate, administrative decisions and research

based on those locale data may be unreliable – depending on the magnitude of the assignment error. This is likely to occur, at least in part, because CCD school locale classifications carry over into other NCES surveys. In other words, if a school is not assigned as City in CCD, it will not be assigned as City in other NCES surveys. This practice is designed to offer consistency across NCES surveys, but the disadvantage is that it also has the potential to reproduce non-sampling error from one source to another. This is equally true for school district locale assignments, since district locales are coded according to the assignments given to the district's schools. Less obvious, but equally important, is the use of school locales as a sampling stratifier in surveys like NELLS, SASS, and ECLS to insure samples adequately represent schools from all types of geographic areas (McMillen, Kasprzyk, & Planchon, 1994). If locale assignments in the original school universe are unreliable, then the resulting samples could be compromised as well. These issues might be less of a concern if educators and researchers didn't use the locale data – but they do.

### Locale comparisons

Geographic classifications of school location saturate the education literature, and a variety of education journals cater specifically to urban and rural researchers. From general comparisons of school counts (Hoffman, 1999a; Hoffman 1999b; Bandeira de Mello, 2000), to detailed consideration of the amount of time elementary students spend on core academic subjects (Perie, Baker & Bobbit, 1997), educators have persistently compared the conditions of schooling on the basis of school location. The significance of educational geography and the nomenclature of city and suburbs has been a mainstay in discussions of desegregation policy over the years (Jacobs, 1998; James, 1989; McDermott, 1998; Orfield & Eaton, 1996), and educational historians have regularly employed geographic indicators to describe philosophical, pedagogical, and administrative changes in public education (Cremin, 1988; Mirel, 1993). Likewise, the very notion of school choice implicitly acknowledges that school attendance is primarily a function of place of residence (i.e., geography), so it isn't surprising to see locale comparisons included in choice discussions (Carnevale & Desrochers, 1999; Friedman, Gutnick, & Aulicino, 1999; McArthur, Colopy, & Schlaline, 1995). Explicit locale comparisons are also featured in discussions of school poverty (Lippman, Burns, & McArthur, 1996), educational costs and organizational resource allocation (Ballou, 1996; Chambers & Flanagan, 1998), the distribution of school crime and safety (Kaufman, et.al, 1999; Shen, 1997), the distribution of academic achievement among Asian and Hispanic students (Kaufamn, Chavez, & Lauen, 1998; Smith, 1995), academic preparation of first-generation college students (Horn & Nunez, 2000), school size and class size (McLaughlin, Huberman, & Hawkins, 1997; Naik, 1999), school infrastructure quality (Rowand, 1999; White, 1999), teacher attrition and mobility (Boe, Bobbit, Cook, Whitener, & Weber, 1996; Ingersoll & Rossi, 1995; Whitener et.al, 1997), teacher satisfaction (Henke Choy, Chen, Geis, & Alt, 1997; Perie & Baker, 1997), the distribution of minority principals (Rossi & Dougherty, 1996), differences in technology access (Heaviside, Farris, & Malitz, 1995), the distribution of students with disabilities (Rossi, Herting, & Wolman, 1997), access to early childhood programs (DeAngelis & Rossi, 1997), and myriad other areas of educational research. The terms may differ between research studies – one uses city and fringe, another uses urban and suburban, while still another applies different definitions to the same terms – but categories of location have become familiar indicators for describing schools and the conditions of schooling in the U.S. In the same way that demographic disaggregation by race or sex is commonplace, school locale comparisons have become standard descriptive fare for many educational analyses.

Despite the widespread application of school locale data, there is a paucity of research that directly investigates the quality, appropriateness, or coherence of these geographic classifications (Stephens, 1992). Most of what little direct locale discussion is available typically describes the CCD assignment process or compares locales across NCES surveys (Johnson, 1989; Johnson, 1994; Owens, 1999). Likewise, little comment (and perhaps awareness) exists about the integrative nature of the CCD locale assignments as a source for locale data in other NCES surveys, or the potential for systematic CCD errors to compromise data quality in other NCES datasets. This inattention in the literature may occur, at least in part, because variants and combinations of locale titles in other NCES datasets obscure the linkage with CCD. But it may also occur because the fundamental geographic concepts that define the CCD locale framework are unfamiliar to most educational researchers.

### **Current criteria and core geographic concepts**

The current CCD typology relies on geographic concepts from the Census Bureau and OMB to classify U.S. territory into eight community types. The categories and criteria include:

1. *Large City*: A principal city of a metropolitan statistical area, with the city having a population greater than or equal to 250,000.
2. *Midsized City*: A principal city of a metropolitan statistical area, with the city having a population less than 250,000.
3. *Urban Fringe of Large City*: Any incorporated place, census designated place, or non-place territory within a metropolitan statistical area of a Large City and defined as urban by the Census.
4. *Urban Fringe of Midsized City*: Any incorporated place, census designated place, or non-place territory within a metropolitan statistical area of a Midsized City and defined as urban by the Census.
5. *Large Town*: An incorporated place or census designated place with a population greater than or equal to 25,000 and located outside a metropolitan statistical area.
6. *Small Town*: An incorporated place or census designated place with a population less than 25,000 and greater than 2,500 and located outside a metropolitan statistical area.
7. *Rural, Outside Metropolitan Area*: Any incorporated place, census designated place, or non-place territory not within a metropolitan statistical area and defined as rural by the Census Bureau.
8. *Rural, Inside Metropolitan Area*: Any incorporated place, census designated place, or non-place territory within a metropolitan statistical area and defined as rural by the Census Bureau.

As indicated earlier, careful consideration of the NCES locale indicators requires some familiarity with geographic concepts employed by and – in many cases – defined by the Census Bureau. The current CCD typology involves six primary concepts. They include the following:

#### Urban (urbanized areas and urban clusters)

The Census Bureau defines an urban area as a densely settled core of census block groups and census blocks that meet minimum population density requirements, along with adjacent densely settled surrounding census blocks. When a core area contains a population of 50,000 or more, it is classified as an urbanized area (UA). Core areas with population between 2,500 and 50,000 are classified as urban clusters (UC) (Federal Register, 2002).

#### Rural

The Census Bureau classifies all population and territory not included in an urbanized area or urban cluster as rural.

#### Core Based Statistical Areas

Core Based Statistical Areas (CBSA) are defined by OMB and represent county or counties associated with at least one core of 10,000 or greater population, plus adjacent counties having a high degree of social and economic integration with the core(s) as measured by commuting ties (Federal Register, 2000). CBSAs with

a population core of 50,000 or more are identified as metropolitan statistical areas (metros), and those with population cores of 10,000 to 50,000 are identified as micropolitan statistical areas (micros). Unlike urbanized areas and urban clusters that are primarily designed to reflect urban structure, metro and micro areas are primarily designed to reflect the functional relationship between urban cores and the areas surrounding them. This includes relationships between urban cores, as well as relationships between urban cores and surrounding rural areas. Since UAs and UCs are constructed from census blocks and block groups and are designed to reflect the structural effects of urbanization, and CBSAs are a separate concept constructed from counties and designed to reflect functional spatial relationships at a larger scale, it is reasonable to find urban and rural territory both inside and outside CBSAs. This relationship is illustrated for the central Arkansas region in Figure 1.

### Principal City

Principal cities include the largest place (incorporated or unincorporated) and other relatively large places that serve as the primary population and employment centers within a CBSA. Principal cities replaced the older central city term defined by OMB's 1990 metropolitan area standards, recognizing that many central cities have become much less central (functionally and structurally) within increasingly polynucleated urban areas (Muller, 1981; Stanbeck, 1991). Although principal cities are present in both metropolitan and micropolitan statistical areas, CCD City locale classifications are currently limited to principal cities of metropolitan statistical areas only.

### Place

Census places are considered to be concentrations of population that are legally bounded and incorporated. Most towns and cities fall into this category. However, many areas that look like towns and cities with commonly recognized community names are not legally incorporated. To accommodate these place-equivalent areas, the Census Bureau identifies them as census designated places or CDPs (U.S. Bureau of the Census, 1994). For most common analytic and data production purposes, places and CDPs are treated as equivalents. This was not the case prior to the 1990 census. Unless noted otherwise, any mention of place in the remainder of this discussion includes both incorporated places and census designated places.

### **Problems with the current locale framework**

NCES has incorporated new CBSA terms into its locale criteria, but these accommodations do not address many technical conflicts and conceptual shortcomings inherent in the current framework. Some of the more significant problems include the following:

#### Lack of Small City

Most principal cities in metropolitan statistical areas have a population less than 70,000, and only about 1 in 10 are large enough for CCD to class as a Large City, i.e., cities with a population greater than 250,000. This means that CCD homogenizes the overwhelming majority of principal cities in metro areas into a single Midsize City category. Unfortunately, this monolithic classification fails to recognize that school systems in small cities may face very different demographic and economic conditions than those in larger cities, and it offers little if any ability to detect systematic differences in educational systems or outcomes among these areas.

#### Suburban

Educational researchers, program administrators, and policymakers frequently refer to suburban schools, but seldom offer an explicit definition of suburban territory. This gets particularly problematic when suburban is used as part of the common tripartite scheme – urban, suburban, and rural. The urban definition noted earlier leaves little conceptual space for a third type of suburban territory that is both non-urban and non-rural. A more frequent case is made for situating suburban as a metropolitan component, i.e., the territory inside a metro area but outside the primary city or cities. However, equating the urban, suburban, rural locale trio to the three metropolitan components commonly used for statistical reporting (inside metro and inside city, inside metro and outside city, outside metro) results in dysfunctional and ill-defined categories because

CBSA criteria clearly accommodate both urban and rural territory within each of the three reported metropolitan components. Again, this occurs because the urban and metropolitan/micropolitan concepts are intentionally designed to identify different sets of spatial characteristics. A better and frequently used approach among NCES data users is to combine the two CCD Urban Fringe categories into a de facto suburban classification, since this at least limits consideration to non-city *urban* territory within a metro area. However, even this more limited definition continues to include many urban areas that are difficult to justify as suburbs.

#### Urban Fringe trumps Town

The CCD currently defines the Urban Fringe as *all* urban territory inside a metro area but outside a principal city. In other words, by definition Towns may not exist within a metro area. Unfortunately, that condition is contrary to most theoretical models of metropolitan spatial organization that suggest smaller urban cores would indeed be found near larger urban cores (Burgess, 1925; Christaller, 1966; Harris & Ullman, 1945; Hoyt, 1939; Vance, 1991). For example, note in Figure 1 how the smaller cores of Conway, Lonoke, England, and Sheriden are distinct yet proximate to the larger Little Rock – North Little Rock core area. Instead of allowing these smaller cores to be classified as towns, the current NCES locale criteria force them into the Urban Fringe of a Large City. Since counties are the building blocks of metro areas, and the boundaries of metro counties and county-clusters can be quite extensive, the current CCD criteria allow suburbs to be extremely distant from the cities with which they are supposedly associated. The addition of these outlying urban areas to the CCD Urban Fringe has little effect on the already large universe of Urban Fringe schools, but it creates a substantial undercount of Town schools and school districts.

#### Urban Fringe of Midsize City reassigned as Urban Fringe of Large City

Although large metro areas frequently contain principal cities of various sizes and multiple urban cores (UAs and UCs), the CCD locale framework requires metro areas to be classified as *either* a metro area of a Large City *or* a metro area of a Midsize City. It cannot be both. Consequently, if a metro area contains both Large and Midsize cities, *all* schools located in the Urban Fringe of those cities are classified as Urban Fringe of a Large City – regardless of how many Midsize cities may be in the metro, how many schools may be located in the Urban Fringe of those Midsize cities, or how far those cities may be from the large principal city that defined the metro. This results in an overcount of schools assigned to the Urban Fringe of a Large City, and an undercount of schools assigned to the Urban Fringe of a Midsize City.

#### Ineffective distance proxy

The 2000 OMB metropolitan area standards reaffirm the distinction between metropolitan and urban. Metro areas are not the same as urban areas. Metro areas are constructed out of county units and determined by population size and county-to-county commuting ties. Urban areas are constructed out of census blocks and block groups and are determined by population size and density. Metro areas may include both urban and rural territory, just as rural territory may exist both inside and outside of metro areas. The current CCD locale criteria recognize this rural distinction and provide separate categories for Rural – Outside Metropolitan Statistical Area and Rural – Inside Metropolitan Statistical Area. These categories may provide a useful distinction for some descriptive purposes, but they may also mislead data users into thinking that rural territory outside metro areas is somehow more rural than rural territory inside metro areas. This isn't necessarily the case. The extensive and often oddly shaped boundaries of some metro counties and county clusters allow some rural territory inside metro areas to be more distant from large or midsize urban cores than rural territory in adjacent non-metro counties. For example, Figure 1 shows that rural territory immediately outside the southwest corner of metropolitan Garland County (between Hot Springs and Malvern) is much closer to the Hot Springs, AR urbanized area than rural territory inside the northwest corner of Garland County. Also recall that the current framework excludes micropolitan statistical areas from Fringe and City assessments, so non-metro rural schools could be in areas immediately adjacent to urban cores or principal cities of micro areas. Although the inside-outside distinction offers a crude distance proxy, these categories offer very little help differentiating rural schools in remote, isolated areas from rural schools in areas much nearer to urban cores – areas that may soon be absorbed by urban expansion.

### Town size rather than relative proximity

The CCD's current exclusion of Towns from metro areas and the current delineation of Towns on the basis of size (Large and Small) have led some NCES data users to assume that Small Towns are necessarily located in remote areas and can be conveniently grouped with rural areas for analytic and descriptive purposes. Indeed, many Small Towns are located in relatively remote areas, but many others are not. For example, note the location of Beebe, AR relative to the Little Rock – North Little Rock core in Figure 1. One of the key problems with the CCD's current reliance on metropolitan criteria is that it blinds NCES data users from recognizing that many Small Towns are quite close to much larger urban cores, and this proximity may provide greater access to specialized goods and services than is available in a much larger town that may be further away. Unfortunately, the current CCD Large and Small Town categories offer little if any insight about the spatial relationships between Towns and larger urban cores.

### **Modified model**

These technical and conceptual issues appear significant enough to warrant a change in the current locale typology. NCES could address these issues in various ways, but the model proposed below would seem to resolve most of the key technical and conceptual conflicts, while maintaining the relative simplicity and parsimony of the current typology. The proposed model is constructed from the same set of standard geographic concepts used by the current typology, but it prioritizes an urban-centric rather than metro-centric approach. The model identifies four primary locales (City, Suburb, Town, Rural), each with three secondary subtypes. They include:

*City: Large:* Territory inside an urbanized area and inside a principal city with population of 250,000 or more.

*City: Midsize:* Territory inside an urbanized area and inside a principal city with population less than 250,000 and greater than or equal to 100,000.

*City: Small:* Territory inside an urbanized area and inside a principal city with population less than 100,000.

*Suburb: Large:* Territory outside a principal city and inside an urbanized area with population of 250,000 or more.

*Suburb: Midsize:* Territory outside a principal city and inside an urbanized area with population less than 250,000 and greater than or equal to 100,000.

*Suburb: Small:* Territory outside a principal city and inside an urbanized area with population less than 100,000.

*Town: Fringe:* Territory inside an urban cluster that is less than or equal to 10 miles from an urbanized area.

*Town: Distant:* Territory inside an urban cluster that is more than 10 miles and less than or equal to 35 miles from an urbanized area.

*Town: Remote:* Territory inside an urban cluster that is more than 35 miles of an urbanized area.

*Rural: Fringe:* Census-defined rural territory that is less than or equal to 5 miles from an urbanized area, as well as rural territory that is less than or equal to 2.5 miles from an urban cluster.

*Rural: Distant:* Census-defined rural territory that is more than 5 miles but less than or equal to 25 miles from an urbanized area, as well as rural territory that is more than 2.5 miles but less than or equal to 10 miles from an urban cluster.

*Rural: Remote:* Census-defined rural territory that is more than 25 miles from an urbanized area and is also more than 10 miles from an urban cluster.

### **Benefits of the Proposed Typology**

No typology will satisfy the needs of every NCES data user, but the proposed framework introduces a number of changes that should improve the usefulness of school and district CCD locale assignments for analytic and program purposes. Some of the key advantages include:

#### Urban-centric criteria

The proposed typology is constructed from urban-centric rather than metro-centric criteria, and is therefore free of the artificial constraints and problems previously imposed by metro county boundaries. This change allows Towns to be located relatively close to larger urban cores, and it prevents the creation of untenably distant suburbs. For example, instead of forcing Conway, Lonoke, England, and Sheridan to be the Urban Fringe of a Large City as done by the current CCD criteria, Figure 2 illustrates how the proposed typology offers a more reasonable Town designation for these smaller urban cores.

#### GIS

The proposed framework relies on a geographic information system (GIS) to classify territory according to the proposed criteria, and then to assess the relationship of school location relative to the classified territory. This approach not only provides the ability to identify hierarchical relationships (i.e., X is located within Y), but also provides the flexibility to identify other spatial relationships (e.g., the distance from X to Z). The current method of locale delineation relies on simple pre-defined non-spatial hierarchical data that can easily identify subsets (i.e., X is located within Y), but it has little flexibility to assess other spatial relationships.

#### Suburban

The proposed framework provides an explicit Suburban classification with clear criteria that identify a more limited and justifiable portion of urban territory than compared with the current Urban Fringe categories. This offers a convenience for data users who have regularly applied the term Suburban in spite of the CCD Urban Fringe title, and also offers a potential definition for legal references to suburban schools or school systems (No Child Left Behind Act, 2001a-e).

#### Small City

The introduction of a new Small City category offers much needed variation to the overly large set of Midsize Cities currently identified by the CCD.

#### Distance indicators

One of the primary advantages of the proposed locale framework is the use of explicit distance measures to identify Town and Rural subtypes. Unlike the current CCD framework that differentiates Towns on the basis of population size, the proposed typology classifies Towns according to their proximity to larger urban cores. This approach considers potential spatial relationships and acknowledges the likely interaction between urban cores based on their relative locations. Rural subtypes are similar in that they identify rural territory relative to urban cores. This distinction avoids the often-misleading distance proxy based on county metro status. More importantly, the explicit distance indicators offer the opportunity to identify and differentiate rural schools and school systems in relatively remote areas, from those that may be located just outside an urban core. Although Town subtypes are based on proximity to urbanized areas only, rural subtypes consider the proximity of the area relative to *both* urbanized areas and urban clusters. Since rural areas near to large urban cores are likely to have greater access to specialized goods, services and employment opportunities

than rural areas equally near to smaller urban cores, the Rural subtype distance requirements attempt to mitigate these differences by imposing larger distance requirements from urbanized areas than from urban clusters.

#### Supplemental ZIP locale assignment

A final advantage of the proposed framework is the provision of ZIP code locales to supplement missing school assignments. Unlike the current CCD supplemental assignment process that relies on place-matching and basic ZIP urban/rural conditions to supplement locales, the proposed framework directly assigns the full set of locales and subtypes to ZIP code areas based on the same process used for district locale assignments. Schools missing locale assignments can receive a substitute assignment by simply matching the ZIP reported in the school address to the supplemental locale assigned provided for that ZIP. In addition to offering a consistent method for supplemental school assignments, the ability to apply population-weighted ZIP locale assignments to a broader set of more precisely defined locales is also likely to improve the accuracy of supplemental locale assignments.

#### **Comparison: Effects of Proposed Typology**

While it's clear the CCD typology has important technical and theoretical issues that may affect the reliability of school locale assignments, it is not clear whether these shortcomings are trivial technicalities and arcane theoretical matters that affect a small set of schools, or whether these shortcomings have a substantial impact on the school locale universes. Fortunately, the proposed typology provides an opportunity to quantify the magnitude of the current assignment problem. Assuming the categories of the proposed typology offer a more appropriate spatial representation than the current classifications, the size of the locale assignment problem would be represented by the number of schools that were reclassified out of their primary CCD type (i.e., CCD City school reassigned as non-City, CCD Urban Fringe school reassigned as non-Suburban, CCD Town school reassigned as non-Town, and CCD Rural assigned as non-Rural). The following data and methods were used to conduct this comparison.

#### Data collection and preparation

The comparison was based on spatial data identifying the extent of urbanized areas, urban clusters, principal cities of metropolitan statistical areas, ZIP codes, and census blocks. These vector spatial layers are a part of the Census Bureau's TIGER (Topologically Integrated Geographic Encoding and Referencing) system, a comprehensive collection of administrative and statistical boundary files that serve as the backbone for the Census Bureau's administrative data collection efforts. Data were extracted from TIGER/Line 2004, a public spatial data product developed from the Census TIGER database. TIGER/Line 2004 did not include updated boundaries for CDPs that are principal cities, so these missing areas were supplemented with CDP data from Census 2000. Population-weighted ZIP code locale assignments were based on Census 2000 block-level population data (100 percent count).

As indicated in the earlier discussion of proposed benefits, all of the comparison data (spatial and non-spatial) were handled and integrated in a geographic information system (GIS). This is still a relatively new technology in educational research (Cobb, 1999; Gobalet, 1994; Higgenbotham, 1996; Orfield, 1997), but it serves as a primary analytic technology for urban geographers, urban planners, and other researchers whose primary need is to analyze geographic entities, spatial patterns, and spatial relationships (Dueker & DeLacy, 1990; Richter, 1992). (A notable exception to the recent introduction of GIS for educational research is the early use of Census GBF/DIME files for local educational planning (Herron, 1976; Otto, 1976; Westenhofer, 1976)). A complete discussion of the functionality of GIS is outside the scope of this brief review, and is better addressed elsewhere (Longley, Goodchild, Maguire & Rhind, 1999). Suffice it to say, this technology offers the ability to analyze school locations, administrative and statistical boundaries, land-use patterns, and various other geographic phenomena. GIS geocoding algorithms translate school addresses into physical coordinates (latitude/longitude) by matching the school address with address ranges on georeferenced street segments. In other words, if the GIS knows the physical location of a street segment, and it can match a school address with the address range provided for the street segment, it can then

interpolate the location of the school within that address range and assign location information from the street segment to the school. Once school location has been determined, attributes of other administrative and statistical areas (e.g., demographics or locale assignments) can be joined to schools based on their spatial relationship to the other geographic entities.

The proposed school locale assignments were based on latitude and longitude values provided by NCES in a preliminary CCD 2003-2004 school file. This file included the current CCD school locale assignment (LOCALE03), and it also provided supplemental school address and location information that had later been developed and added by an NCES contractor. These geocodes were primarily based on a school's physical address. However, if a physical address was not reported or uncodable, the contractor employed additional information sources and methods to determine school location (e.g., reported mailing address, commercial address lists, reverse geocoding based on reported telephone number, and manual data review to identify potential address input errors). Of the 100,594 schools on the file, our review identified only 31 that lacked the geocode information needed for a locale assignment. Manual review of these non-assigned cases resulted in supplemental locale and subtype assignments for 30 of the 31 unassigned schools. Supplemental assignments were not provided for non-geocoded schools located in the U.S. island territories.

### Locale Assignment

#### *Territory assignment*

The first and most critical step of the school locale assignment process was to assign locales and subtypes to the full extent of U.S. territory and Puerto Rico. Locales were not provided for U.S. island territory (Virgin Islands, Guam, American Samoa, and the Northern Mariana Islands). As indicated earlier, we used a GIS to evaluate the various spatial data layers according to the proposed criteria. Distances for Town and Rural subtypes were based on straight-line or Euclidean distance. Although this simple geometric measure does not account for the presence or absence of road networks that may offer point-to-point drive time estimates, it is also unaffected by short-term changes to the transportation infrastructure that could cause significant fluctuations in those estimates. More importantly, the geometric distance provides NCES data users with a simple and familiar concept that is analytically useful and relatively easy to implement. The basic unit for these distance indicators – 2.5 miles – was borrowed from the Census Bureau's criterion for connecting densely settled non-contiguous territory to a qualifying core of an urbanized area or an urban cluster during the urban delineation process (officially referred to as a 'jump'). Distances used to define locale subtypes are simple multiples of the basic distance unit (i.e., 1x, 2x, 4x, and 10x for Rural; 4x and 14x for Towns).

#### *School assignment*

The process for assigning new school locales was conceptually straightforward. First, the territory of the U.S. was classified according to the proposed locale and subtype criteria. Second, schools were spatially integrated with the territory based on school geocodes. Third, the schools were assigned a locale and subtype based on their location, i.e., they received the same assignment given to the territory where they were located. In cases where school geocodes were unavailable, we proposed to provide supplemental locale and subtype assignments based on the locale and subtype assigned to the ZIP code area identified in the school address. However, the NCES supplemented school file provided nearly universal geocode coverage, making supplemental ZIP locale assignments unnecessary. Although geocodes were available for almost all schools, and all geocoded schools were assigned a new locale, only schools with valid enrollment values (MEMBER03 >= 0) were included in the comparison of current vs. proposed locale assignments. Of the 100,594 schools on the file, 96,084 satisfied this enrollment requirement.

#### *District assignment*

School district locale and type assignments were based on enrollment-weighted locale assignments of schools within the district. First, enrollment in each locale subtype was identified for each district. Second, the district was examined to see if a single locale subtype accounted for 50 percent or more of the district enrollment. If a majority was present, the district was assigned that majority locale subtype. If the district

lacked a majority locale subtype, the locale subtypes were then aggregated into their respective locales (City, Suburb, Town, Rural), and the locales were checked for a 50 percent majority enrollment. If a majority locale was identified, then the district was assigned the locale subtype that had the plurality within the majority locale. If, however, none of the aggregate locales satisfied the 50 percent majority enrollment criterion, then the district locale assignment defaulted to the single locale subtype that accounted for the largest percentage of district enrollment. Most district assignments (97 percent) were based on the presence of a 50 percent majority locale subtype. In the small percentage of cases where school districts were composed entirely of schools reporting zero enrollment, the district assignments were based on non-weighted school counts.

#### *ZIP code assignment*

ZIP code locale assignments would have been based on Census Bureau ZCTAs, geographic entities developed by the Census Bureau and designed to approximate USPS five-digit ZIP Code service areas. ZCTAs are aggregations of census blocks that have the same predominant ZIP code associated with the residential mailing addresses in the U.S. Census Bureau's Master Address File. ZCTAs do not precisely depict ZIP code delivery areas and do not include all ZIP codes used for mail delivery. Some ZCTAs cover remote or non-residential areas such as water bodies, wilderness areas, and military installations that fall outside the scope of the ZIP codes reported by CCD schools. ZCTA codes for water bodies are indicated with the suffix 'HH' in the fourth and fifth digits, while codes for non-hydrographic uncovered areas are suffixed with 'XX'. Additionally, ZIP code boundaries are not static. Therefore locale assignments based on TIGER/Line 2004 current ZCTAs may not reflect the same geographic area presently served by the ZIP code.

ZIP code locale assignments relied on the same basic decision rules applied to CCD school districts. First, the population in each locale subtype was identified for each ZCTA. Second, the ZCTA was examined to see if a single locale subtype accounted for 50 percent or more of the population within the ZCTA. If so, the ZCTA was assigned that majority locale subtype. If the ZCTA lacked a majority locale subtype, the locale subtypes were aggregated into their respective locales (City, Suburb, Town, Rural), and the locales were checked for a 50 percent majority population. If a majority locale was identified, then the ZCTA was assigned the locale subtype that had the plurality within the majority locale. If, however, none of the aggregate locales satisfied the 50 percent majority population criterion, then the ZCTA locale assignment defaulted to the single locale subtype with the largest population percentage within the ZCTA. Most ZCTA assignments (96 percent) were based on the presence of a 50 percent majority locale subtype.

#### Comparison Results

As anticipated from the comments and criticisms noted above, the recommended changes affected some CCD locales more than others, most notably Urban Fringe/Suburban and Town classifications. Table 1 identifies changes in locale distributions for schools, school enrollment, and school districts. Table 2 identifies how CCD schools were redistributed across the proposed classifications. Although most schools retained their primary CCD locale type under the proposed typology (e.g., most CCD City schools remained City), approximately 12 percent of CCD schools were reclassified into a different primary type (e.g., City school reassigned as non-City). Many of these reclassified schools were retyped due to differences in geocoded location, but most reclassifications were a consequence of corrections to the locale criteria.

#### *City*

The new typology had minimal impact on the universe of City schools because the proposed City criteria are quite similar to the current definitions. The slightly increased percentage of City schools under the proposed typology was primarily attributable to the addition of previously non-geocoded and non-assigned schools to City areas, and the inclusion of previously borderline schools in the Urban Fringe and Rural areas that were re-geocoded to a location inside the City boundary. Approximately 28 percent of the schools that formerly lacked CCD locale assignments were assigned to Cities. Although the new criteria had minimal impact on the overall number of City schools, the introduction of the Small City classification seriously impacted the distribution of City schools, accounting for 29 percent of all City schools and 27 percent of City school

enrollment. In other words, about one out of every four City public school students attends school in a city that has a population less than 100,000.

### *Suburban*

The transition from Urban Fringe to Suburban had a sizable impact on the universe of Fringe schools. The move not only reduced the set of Fringe schools by 12 percent, it also dethroned the Urban Fringe as the locale with the largest portion of public schools. Though Suburban schools continued to account for the largest share of public school enrollment (36 percent), the proposed typology revealed that Rural areas actually account for the largest share of U.S. public schools (31 percent). This has likely been true for a while, but the Urban Fringe appeared larger because the current criteria were unable to distinguish suburban fringe areas from Towns that happened to be located in metro counties. Approximately 79 percent of CCD Urban Fringe schools were reclassified as Suburban, and 39 percent of schools that had formerly lacked a CCD locale assignment were identified as Suburban schools. Though most Urban Fringe schools were reclassified as Suburban, 13 percent were reassigned to Towns, 6 percent were reassigned to Rural areas, and 1 percent were reassigned to Cities. In addition to the impact on schools and school enrollment, the Suburban transition also resulted in a 22 percent reduction of Urban Fringe school districts.

### *Town*

As expected, the proposed typology has a massive impact on Towns. Town schools increased by 53 percent, Town enrollment increase by 56 percent, and Town districts increased by 50 percent. These results indicate that the overly restrictive metro criterion in the current CCD locale typology caused more than one-third of all Town schools, students, and districts to be incorrectly classified as Urban Fringe. Although Towns accounted for the smallest percentage of schools and students of any locale under both the current and proposed typology, the magnitude of change between these two frameworks suggests that the representation of Town schools, students, and districts in current NCES data products may be biased. In addition to the magnitude of change, the distribution of Town schools, students, and districts by subtype is quite interesting. Approximately one out of every four Town schools is located near an urban core with a population of 50,000 or more (within 10 miles), while 34 percent of schools are in more remote locations over 35 miles away from an urbanized area. Although the proposed typology does not delineate Towns by size, the comparison indicated that many of the currently classified Small Towns that get combined with Rural areas for analytic or program purposes may actually be quite close to large urban cores and more akin to suburbs than to rural areas. Only about half of CCD Small Towns were in remote locations more than 35 miles from an urbanized area.

### *Rural*

The absolute percentage change in Rural schools, students, and districts under the current and proposed typology was less than 2 percent. As with the current typology, more than half of all public school systems were assigned as Rural under the proposed typology, and about one out of every five public school students still attended a Rural school. However, the comparison offers an interesting and useful analytic distinction. Contrary to the remote, rural school stereotype, more than one-third of Rural public schools and more than one-half of Rural public school students are located quite close to a sizable urban core, and many of these rural areas have significant – if not more – access to urban amenities and specialized goods and services than many Small and Large Towns. Other rural areas clearly do not. Results of the proposed criteria located 29 percent of all Rural schools and 15 percent of Rural students in more remote territory that was 25 miles from the nearest urbanized area and more than 10 miles from an urban cluster. While the large percentage of the Rural-Fringe schools may challenge some rural stereotypes, the similarly large percentage of schools located in remote Rural areas –entirely Census-defined rural territory, not rural areas accompanied by Towns, – could be equally surprising for researchers and policymakers who may have given little attention to the unique educational needs of isolated rural schools.

## **Implications and Conclusions**

The CCD locale typology serves as a primary framework for differentiating schools, students, and districts by geographic area, and it is widely used for a variety of analytic and program purposes by federal, state, and non-governmental agencies. Despite this widespread use, the criteria that compose the CCD locale typology are not well understood, nor are the various ways CCD locale assignments are integrated with other NCES data products. As noted earlier, NCES frequently uses geographic locale (based on the CCD framework) as a stratifier for developing sample surveys, and schools that are included in those samples typically report the locale assigned to them by CCD. Therefore, if a large portion of CCD school locales were misclassified, the error would not only impact CCD data quality, it could affect the reliability of other NCES data and analytic products as well. Detailed consideration of those effects was outside the scope of this project, but some of the results presented in this analysis – particularly the potential misclassification of one out of every ten public schools, and one-third of Town schools – suggest that additional investigation in this area would be prudent.

In addition to raising interesting questions about the reliability of current data, adopting the proposed recommendations would also present a number of administrative challenges for NCES/ED programs. For example, changes to the CCD typology would create sampling changes for surveys that use a locale stratifier, changes in locale values currently reported for schools in those surveys, and changes for other administrative data like the Private School Survey (PSS) which would need to adopt the new typology to maintain locale consistency for those surveys that sample both public and private schools. Adopting the proposed recommendations could also have implications for federal initiatives like the Rural Education Achievement Program that legally require specific locale values (i.e., ‘7’ and ‘8’) rather than terms (i.e., ‘rural’) to identify rural schools. Although the proposed typology and the current typology share the same basic rural definition as well as the same basic set of schools, assigning the three proposed Rural subtypes with any code other than ‘7’ or ‘8’ could create significant conflicts with REAP administration.

Despite these short-term challenges, the proposed typology offers NCES significant long-term advantages. At a minimum, it resolves substantial technical and conceptual problems in the current locale framework that alone could justify change. But the analytic benefits are equally important. By adding Small Cities, more limited Suburbs, and Town and Rural subtypes that reflect proximity to urban cores, the proposed typology offers NCES significantly more analytic precision and flexibility than is provided by the current locale scheme. Since NCES’s primary mission is to inform policymakers and the public about the condition of education in the U.S., and geography is one of the primary organizing features of public education, it makes sense for NCES to develop and apply effective geographic indicators based on high quality spatial data. The recommendations provided in this analysis offer NCES a modest but practical step toward those ends.

## References

- Ballou, D. (1996). The condition of urban school finance: Efficient resource allocation in urban schools. In Fowler, W.J. (Ed). Selected papers in school finance, 1996. Washington, DC: National Center for Education Statistics, U.S. Department of Education. Pp. 61-83.
- Bandiera de Mello, V. (2000, January). State profiles and public elementary and secondary education, 1996-97. Washington, DC: National Center for Education Statistics, U.S. Department of Education.
- Boe, E.E., Bobbitt, S.A., Cook, L.H., Whitener, S.D., & Weber, A.L. (1996, June). Predictors of retention, transfer, and attrition of special and general education teachers: Data from the 1989 teacher followup survey. Washington, DC: National Center for Education Statistics, U.S. Department of Education.
- Burgess, E.W. (1925). The growth of the city. In R.E. Park, E.W. Burgess, & R.D. McKenzie (Eds). *The City*. Chicago: University of Chicago Press. Pp. 47-62.
- Carnevale, A. P. & Desrochers, D. M. (1999). *School satisfaction: A statistical profile of cities and suburbs*. Princeton, NJ: Educational Testing Service.
- Chambers, J.G. & Flanagan, J.C. (1998, February). Geographic variation in public schools' costs. Washington, DC: National Center for Education Statistics, U.S. Department of Education.
- Christaller, W. (1966). *Central places in southern Germany*. Translated by C.W. Baskin. Englewood Cliffs, NJ: Prentice Hall.
- Cobb, C.D. (1999, April). Using geographic information systems in education research and policy analysis. Paper presented at the Annual Meeting of the American Educational Research Association.
- Cremin, L.A. (1988). *American education: The metropolitan experience*. New York: Harper & Row. Pp. 575-604.
- DeAngelis, K. & Rossi, R. (1997, February). Schools serving family needs: Extended-day programs in public and private schools. Issue Brief. Washington, DC: National Center for Education Statistics, U.S. Department of Education.
- Dueker, K.J. & DeLacy, P.B. (1990, Autumn). GIS in the land development planning process: Balancing the needs of land use planners and real estate developers. *Journal of the American Planning Association*. Pp. 483-491.
- Federal Register. (2000, Dec. 27). Standards for Defining Metropolitan and Micropolitan Statistical Areas. 65(249). Pp.82228-82238.
- Federal Register. (2002, March 15). Urban Area Criteria for Census 2000. 67(51). Pp. 11663-11670).
- Friedman, W., Gutnick, A., & Aulicino, C. (1999). *Cities, suburbs and schools: Would citizens in Chicago, Cleveland and Milwaukee support greater collaboration?* New York, NY: Public Agenda.
- Gobalet, J.G. (1994, April). Exploring the spatial element in school district demography using GIS software. Paper presented at the 1994 Population Association of America.

- Harris, C. D., & Ullman, E. L. (1945, November). The nature of cities. *Annals of the American Academy of Political and Social Science*. 242. Pp. 7-17.
- Heaviside, S., Farris, E., & Malitz, G. (1995, February). Advanced telecommunications in U.S. public schools, K-12. NCES 95-731. Washington, D.C.: National Center for Education Statistics, U.S. Department of Education.
- Henke, R.R., Choy, S.P., Chen, X., Geis, S., and Alt, M.N. (1997, July). America's teachers: Profile of a profession, 1993-94. Washington, DC: National Center for Education Statistics, U.S. Department of Education.
- Herron, W. (1976, April). Geocoding applications in the Philadelphia school district. GBF/DIME System conference proceedings, 1975, October 30 and 31, San Francisco, California. GE 60, No.7. Washington, DC: Bureau of the Census, U.S. Department of Commerce. Pp. 94-97.
- Higginbotham, J.S. (1996, June). Mapping the future. *School Planning and Management*, 35(6). Pp. 20-23.
- Hoffman, L. (1999a, June). Overview of public elementary and secondary schools and districts: School year 1997-98. Washington, DC: National Center for Education Statistics, U.S. Department of Education.
- Hoffman, L. (1999b, October). Key statistics on public elementary and secondary schools and agencies: School year 1995-96. Washington, DC: National Center for Education Statistics, U.S. Department of Education.
- Horn, L. & Nunez, A.M. (2000, March). Mapping the road to college: First-generation students' Math track, planning strategies, and context of support. Washington, DC: National Center for Education Statistics, U.S. Department of Education.
- Hoyt, H. (1939). The structure and growth of residential neighborhoods in American cities. Washington, DC: Federal Housing Administration.
- Ingersoll, R. & Rossi, R. (1995, August). Which types of schools have the highest teacher turnover. Issue brief. Washington, DC: National Center for Education Statistics, U.S. Department of Education.
- Jacobs, G.S. (1998). Getting around Brown: Desegregation, development, and the Columbus Public Schools. Columbus, OH: Ohio State University Press.
- James, D.R. (1989). City limits on racial equality: The effects of city-suburb boundaries on public-school desegregation, 1968-1976. *American Sociological Review*, 54. Pp. 963-985.
- Johnson, F. (1989, July). Assigning Type of Locale Codes to the 1987-88 CCD Public School Universe. NCES 89-194. Washington, DC: National Center for Education Statistics, U.S. Department of Education.
- Johnson, F. (1994, July). Comparisons of school locale setting: Self-reported versus assigned. NCES Working Paper Series, 94-01. Washington, DC: National Center for Education Statistics, U.S. Department of Education.
- Kaufman, P., Chavez, L., & Lauen, D. (1998, September). Generational status and educational outcomes among Asian and Hispanic 1988 eighth graders. Washington, DC: National Center for Education Statistics, U.S. Department of Education.

- Kaufman, P., Chen, X., Choy, S.P., Ruddy, S.A., Miller, A.K., Chandler, K.A., Chapman, C.D., Rand, M.R., & Klaus, P. (1999). Indicators of school crime and safety, 1999. Washington, DC: U.S. Departments of Education and Justice.
- Lippman, L., Burns, S. & McArthur, E. (1996, June). Urban Schools: The challenge of location and poverty. NCES 96-184. Washington, DC: National Center for Education Statistics, U.S. Department of Education.
- Longley, P., Goodchild, M, Maguire, D., & Rhind, D. (1999). Geographical information systems: Principles, techniques, applications, and management. New York: John Wiley & Sons.
- McArthur, E., Colopy, K.W., & Schlaline, B. (1995, June). Use of school choice: Education policy issues, statistical perspectives. Washington, DC: National Center for Education Statistics, U.S. Department of Education.
- McDermott, K.A. (1998). Regionalism forestalled: Metropolitan fragmentation and desegregation planning in greater New Haven, Connecticut. In Stone, C.N. (Ed). Changing Urban Education. Lawrence, KS: University Press of Kansas. Pp. 45-65.
- McLaughlin, D.H., Huberman, M. B., Hawkins, E. K. (1997, May). Characteristics of small and rural school districts. Washington, DC: National Center for Education Statistics, U.S. Department of Education. P. 124.
- McMillen, M., Kasprzyk, D., & Planchon, P. (1994, July). Sampling frames at the United States National Center for Education Statistics. NCES Working Paper Series, 94-01. Washington, DC: National Center for Education Statistics, U.S. Department of Education.
- Mirel, J. (1993). The rise and fall of an urban school system : Detroit, 1907-81. Ann Arbor, Mich. : University of Michigan Press.
- Muller, Peter O. (1981). Contemporary suburban America. Englewood Cliffs, N.J. : Prentice-Hall.
- Naik, M. (1999, October). Reducing Class Size in America's Urban Schools. Council of the Great City Schools.
- No Child Left Behind Act of 2001(a), 20 U.S.C. §7245. et seq.
- No Child Left Behind Act of 2001(b), 20 U.S.C. §7247. et seq.
- No Child Left Behind Act of 2001(c), 20 U.S.C. §7269. et seq.
- No Child Left Behind Act of 2001(d), 20 U.S.C. §7275. et seq.
- No Child Left Behind Act of 2001(e), 20 U.S.C. §7283d. et seq.
- Orfield, G. & Eaton, S.E. (1996). Dismantling desegregation: The quiet reversal of Brown v. Board of Education. New York: The New Press.
- Orfield, M. (1997). Metropolitics: A regional agenda for community and stability. Washington DC: The Brookings Institution. Pp. 21-22.

- Otto, R.C. (1976, April). Putting GBF to work within a school system. GBF/DIME System conference proceedings, 1975, October 30 and 31, San Francisco, California. GE 60, No.7. Washington, DC: Bureau of the Census, U.S. Department of Commerce. Pp. 32-38.
- Owens, S. (1999, March). Classification evaluation of the 1994-95 Common Core of Data: Public/elementary/secondary education agency universe survey. NCES 1999-316. Washington, DC: National Center for Education Statistics, U.S. Department of Education.
- Perie, M. & Baker, D.P. (1997, July). Job satisfaction among America's teachers: Effects of workplace conditions, background characteristics, and teacher compensation. Washington, DC: National Center for Education Statistics, U.S. Department of Education.
- Perie, M., Baker, D.P., & Bobbit, S. (1997, February). Time spent teaching core academic subjects in elementary schools: Comparisons across community, school, teacher, and student characteristics. Washington, DC: National Center for Education Statistics, U.S. Department of Education.
- Richter, M.J. (1992, November). Mapping the future: Government use of geographic information systems. *Governing*. Pp. 50-58.
- Rossi, R. & Daugherty, S. (1996, April). Where do minority principals work? Issue brief. Washington, DC: National Center for Education Statistics, U.S. Department of Education.
- Rossi, R., Herting, J., & Wolman, J. (1997, June). Profiles of students with disabilities as identified in NELS:88. Washington, DC: National Center for Education Statistics, U.S. Department of Education.
- Rowand, C. (1999, Spring). How old are America's public schools? *Education Statistics Quarterly*, 1(1). Pp. 53-56.
- Shen, J. (1997, October). The evolution of violence in schools. *Educational Leadership*, 55(2). Pp. 18-20.
- Smith, T.M. (1995, September). Findings from the Condition of Education 1995: The Educational Progress of Hispanic Students. NCES 95-767. Washington, DC: National Center for Education Statistics, U.S. Department of Education.
- Stanbeck, T.M (1991). *The new suburbanization: Challenge to the central city*. Boulder, CO: Westview Press.
- Stephens, E.R. (1992, Fall). Mapping the research task for the construction of a federal system for classifying the nation's rural school districts. *Journal of Research in Rural Education*, 8(3). Pp. 3-28.
- U.S. Census Bureau (1994, November). *Geographic Areas Reference Manual*. Washington, DC: U.S. Department of Commerce.
- Vance, J.E. (1991). Human mobility and the shaping of cities. In J.F. Hart (Ed). *Our changing cities*. Baltimore, MD: Johns Hopkins University Press. Pp. 67-85.
- White, K.A. (1999, July 14). Summer brings building boom in city schools: Urban districts use summer to catch up on construction. *Education Week*, 18(42).
- Whitener, S.D., Gruber, K., Lynch, H., Tingos, K., Perona, M, & Fondelier, S. (1997, May). Characteristics of stayers, movers, and leavers: Results from the teacher followup survey, 1994-95. *Schools and Staffing Survey*. Washington, DC: National Center for Education Statistics, U.S. Department of Education.

Westenhofer, G.B. (1976, April). To serve the community. GBF/DIME System conference proceedings, 1975, October 30 and 31, San Francisco, California. GE 60, No.7. Washington, DC: Bureau of the Census, U.S. Department of Commerce. Pp. 32-38.