COMPARISONS OF SCHOOL LOCALE SETTING: SELF-REPORTED VERSUS ASSIGNED

Frank Johnson, National Center for Education Statistics
U.S. Department of Education, Washington, DC 20208-5651

KEYWORDS: SASS, CCD, Urban, Rural

This paper focuses on the geographic locale settings reported in two surveys conducted by The National Center for Education Statistics (NCES), part of the U.S. Department of Education. The two surveys are the School Universe component of the Common Core of Data (CCD) survey for school year 1988-89, and the Public School component of the Schools and Staffing Survey (SASS) for school year 1990-91. Instances where the self-reported locale setting code from SASS disagree with the assigned locale setting code from CCD are analyzed.

CCD Locale Code

The Common Core of Data (CCD) School Universe Survey is an annual collection, containing a record for every public elementary and secondary school in the United States and territories. NCES assigns each school a locale code by matching each school address to Census Bureau files. Census data used in assigning locale codes are 1) population and population density, 2) Standard Metropolitan Statistical Area (SMSA) codes, and 3) a Census code characterizing places as rural or urbanized areas. All Census data used in this project are based on the 1980 Census of Population and Housing. (For more information on the locale code assignment see Johnson, 1989.) The seven CCD locale codes are:

1. Large City: Central city of an SMSA, with the city having a population greater than or equal to 400,000 or a population density greater than or equal to 6,000 people per square mile.

2. Mid-Size City: Central city of an SMSA, with the city having a population less than 400,000 and a population density less than 6,000 people per square mile.

3. Urban Fringe of Large City: Place within an SMSA of a Large City and defined as urban by Census.

4. Urban Fringe of Mid-size City: Place within an SMSA of a Mid-size City and defined as urban by Census.

5. Large Town: Town not within an SMSA, with a population greater than or equal to 25,000.

6. Small Town: Town not within an SMSA and with a population less than 25,000 and greater than or equal to 2,500 people.

7. Rural: A place with less than 2,500 people or a place having a ZIP Code designated rural by Census.

Definitions of SMSAs and urban and rural areas are given below.

Standard Metropolitan Statistical Areas (SMSA)

SMSAs are defined by the Office of Management and Budget (OMB). Each SMSA comprises a central city or urbanized area and one or more neighboring counties. In order to be classified as an SMSA, two conditions must be met; 1) the central city or urbanized area must have a population of at least 50,000, and 2) the entire metropolitan area (including the central city or urbanized area) must have a total population of 100,000 or more inhabitants (75,000 in New England). Contiguous counties are included if they have close social and economic links with the area's population nucleus. Census assigns each of these SMSAs a unique code. At the time of the 1980 census there were 318 SMSAs in the United States.

The SMSAs that are used in this typology are those defined in 1983 by the Office of

Management and Budget (OMB). Since that time, they have been updated and expanded, and are now called Metropolitan Statistical Areas (MSA).

Urban and Rural Areas

The Bureau of the Census defines urbanized areas as consisting of a central city and surrounding densely settled territory with a combined population of 50,000 or more inhabitants. Places designated as urban by Census are within these urbanized areas or in places of 2,500 or more inhabitants outside these areas. All other areas are classified as rural. The urban and rural classifications cut across the SMSA classifications. There can be both urban and rural territory within an SMSA as well as in non-SMSA areas.

SASS Community Descriptor Codes

The School and Staffing Survey (SASS), Public School component, surveys a sample of schools using the CCD file as a sampling frame. This survey received responses from 8,969 schools. In this survey, respondents to the questionnaire select a locale setting which "best describes the community in which the school is located". There are ten community descriptors ranging from "a rural or farming community" to "a very large city (over 500,000 people)". Two of these community designations are beyond the scope of this analysis. They are "military base or station" and "Indian reservation." Ninety-nine of the 8,969 schools sampled chose these descriptors as best representing their school's setting. These schools have been dropped from this analysis. The remaining community description choices are listed below.

SASS community descriptor codes

1. A rural or farming community.
2. A small city or town of fewer than 50,000 people that is not a suburb of a larger city.
3. A medium-sized city (50,000 to 100,000 people)
4. A suburb of a medium-sized city
5. A large city (100,000 to 500,000 people)
6. A suburb of a large city
7. A very large city (over 500,000 people)
8. A suburb of a very large city

Overall findings

A breakdown of the locale settings assigned and reported for the schools responding to the SASS survey is provided below.

CCD assigned locale codes	Schools	Percent
1. Large central city	633	7.14
2. Mid-size central city	1,318	14.86
3. Fringe of large city	894	10.08
4. Fringe of mid-size city	871	9.82
5. Large town	242	2.73
6. Small town	2,220	25.03
7. Rural	2,692	30.35

SASS self-reported community descriptors	Schools	Percent
1. A rural/farming community	3,336	37.62
2. A small city or town	2,231	25.15
3. A medium-sized city	737	8.31
4. A suburb of medium-sized city	403	4.54
5. A large city	797	8.99
6. A suburb of large city	589	6.64
7. A very large city	408	4.60
8. A suburb of very large city	369	4.16

A crosstabulation is presented in Table 1. The two distributions are remarkably similar, especially if one takes into consideration the differences in the definitions of the two location typologies.

Reconciling CCD and SASS Locale Codes

There are several important differences between these two coding schemes. First of all is the distinction between assigning codes based on measurable demographic data versus an individual's perception of a community setting. The choice of a locale setting is likely to differ from individual to individual. Some individuals may change their response over a brief period of time (Bushery et al, 1992). Many people do not know the population of the town they live in, and one person's suburban is another one's rural.

Though there are inherent problems in an individual's choice of locale setting, there are problems with the CCD computer assigned locale codes as well. CCD locale codes are assigned based on mailing addresses. Several of these addresses are not the street address, but are Post Office boxes in nearby towns, and some schools report the school district mailing address instead of

Table 1.--Distribution of locale setting codes from 1990-91 SASS public school survey, by CCD locale code and SASS community descriptor

SASS community descriptor codes / Frequency, Percent, Row Pct, Col Pct	1 Large City	2 Mid-Size City	3 Large City Fringe	4 Midsize City Fringe	5 Large Town	6 Small Town	7 Rural	Total
1 Rural or Farm	* 2	* 50	* 15	* 94	* 21	936	2,218	3,336
	* 0.02	* 0.56	* 0.17	* 1.06	* 0.24	10.55	25.01	37.61
	* 0.06	* 1.50	* 0.45	* 2.82	* 0.63	28.06	66.49	
	* 0.32	* 3.79	* 1.68	* 10.79	* 8.68	42.16	82.39	
2 Small City	* 6	162	* 150	* 263	174	1,141	335	2,231
	* 0.07	1.83	* 1.69	* 2.97	1.96	12.86	3.78	25.15
	* 0.27	7.26	* 6.72	* 11.79	7.80	51.14	15.02	
	* 0.95	12.29	* 16.78	* 30.20	71.90	51.40	12.44	
3 Medium City	* 23	414	* 94	* 116	34	* 31	* 25	737
	* 0.26	4.67	* 1.06	* 1.31	0.38	* 0.35	* 0.28	8.31
	* 3.12	56.17	* 12.75	* 15.74	4.61	* 4.21	* 3.39	
	* 3.63	31.41	* 10.51	* 13.32	14.05	* 1.40	* 0.93	
4 Suburb of Medium City	* 2	* 84	* 68	139	* 9	* 48	* 53	403
	* 0.02	* 0.95	* 0.77	1.57	* 0.10	* 0.54	* 0.60	4.54
	* 0.50	* 20.84	* 16.87	34.49	* 2.23	* 11.91	* 13.15	
	* 0.32	* 6.37	* 7.61	15.96	* 3.72	* 2.16	* 1.97	
5 Large City	201	446	* 67	* 71	0	* 2	* 10	797
	2.27	5.03	* 0.76	* 0.80	0.00	* 0.02	* 0.11	8.99
	25.22	55.96	* 8.41	* 8.91	0.00	* 0.25	* 1.25	
	31.75	33.84	* 7.49	* 8.15	0.00	* 0.09	* 0.37	
6 Suburb of Large City	* 50	* 75	240	146	* 3	* 37	* 38	589
	* 0.56	* 0.85	2.71	1.65	* 0.03	* 0.42	* 0.43	6.64
	* 8.49	* 12.73	40.75	24.79	* 0.51	* 6.28	* 6.45	
	* 7.90	* 5.69	26.85	16.76	* 1.24	* 1.67	* 1.41	
7 Very Large City	309	* 61	* 27	* 9	0	* 1	* 1	408
	3.48	* 0.69	* 0.30	* 0.10	0.00	* 0.01	* 0.01	4.60
	75.74	* 14.95	* 6.62	* 2.21	0.00	* 0.25	* 0.25	
	48.82	* 4.63	* 3.02	* 1.03	0.00	* 0.05	* 0.04	
8 Suburb of Very Large City	* 40	* 26	233	* 33	* 1	* 24	* 12	369
	* 0.45	* 0.29	2.63	* 0.37	* 0.01	* 0.27	* 0.14	4.16
	* 10.84	* 7.05	63.14	* 8.94	* 0.27	* 6.50	* 3.25	
	* 6.32	* 1.97	26.06	* 3.79	* 0.41	* 1.08	* 0.45	
Total	633	1,318	894	871	242	2,220	2,692	8,870
	7.14	14.86	10.08	9.82	2.73	25.03	30.35	100.00

NOTE: Asterisks indicate conflicting locale setting assignments.

their own. There are also the technical problems of matching city names to files. Spellings, abbreviations and even the entire name can differ greatly through custom and keying errors. And there are towns recognized by the Post Office which are not recognized by the Census Bureau. Whereas steps have been taken in the CCD locale code assignment process to reduce these types of errors, they have not been totally effective.

Another difference lies in the terms suburb and urban fringe. "Suburb" is a common term denoting the settled areas surrounding a city. An effort to capture this setting was made in the CCD survey by the "Fringe" designations. CCD employed the use of SMSA definitions in order to make the locale assignments more scientific and to agree with definitions used elsewhere by the federal government. SASS was seeking a concise definition understandable by their respondents. Because the SMSA boundaries are defined to include whole counties, there are areas over a hundred miles from a city which are defined as Fringe of a large or mid-size city. Without a map of SMSA boundaries it would appear logical for respondents to code such areas as rural or small city.

A final difference occurs in the breakdown of cities and their corresponding fringe/suburban areas. CCD uses the central city of an SMSA definition as its cut-off for being a city, and then arbitrarily makes a distinction between large central cities and mid-size central cities based on population and population density. SASS arbitrarily set up its three tier classification scheme based on population.

Because of these differences, it is impossible to establish a perfect one-to-one relationship between the two coding schemes. However, the following crosswalk was developed in order to make a comparison. In nearly every case, one item in one coding scheme is matched to two items on the other coding scheme. This crosswalk is presented twice below, once in order of the CCD assigned locale code and again in order of the SASS self reported community descriptor on SASS.

CCD definition	SASS definition
1. Large central city	= 7. Very large city
1. Large central city	= 5. Large city
2. Mid-size city	= 5. Large city
2. Mid-size city	= 3. Medium-sized city
2. Mid-size city	= 2. Small city or town
3. Fringe of large city	= 8. Suburb of very large city
3. Fringe of large city	= 6. Suburb of large city
4. Fringe of a mid-size city	= 6. Suburb of large city
4. Fringe of a mid-size city	= 4. Suburb of a medium-sized city
5. Large town	= 3. Medium-sized city
5. Large Town	= 2. Small city or town
6. Small town	= 2. Small city or town
6. Small town	= 1. Rural or farming community
7. Rural	= 2. Small city or town
7. Rural	= 1. Rural or farming community

SASS definition	CCD definition
1. Rural or farming community	= 7. Rural
1. Rural or farming community	= 6. Small town
2. Small city or town	= 7. Rural
2. Small city or town	= 6. Small town
2. Small city or town	= 5. Large town
2. Small city or town	= 2. Mid-size city
3. Medium-sized city	= 5. Large town
3. Medium-sized city	= 2. Mid-size city
4. Suburb of a medium-sized city	= 4. Fringe of a mid-size city
5. Large city	= 2. Mid-size city
5. Large city	= 1. Large central city
6. Suburb of a large city	= 4. Fringe of a mid-size city
6. Suburb of a large city	= 3. Fringe of a large city
7. Very large city	= 1. Large central city
8. Suburb of a very large city	= 3. Fringe of a large city

Schools with conflicting locale settings

After removing schools in which the locale settings from the two coding schemes agree based on this crosswalk, there remain 1,742 schools where the codes do not agree. These occurrences are highlighted in Table 1 by asterisks to the left of the data inside the box. This represents 20 percent of the entire SASS public school sample. More than half of these schools with conflicting locale codes were coded as urban fringe on CCD (1,007 schools or 57 percent of the 1,742). The distribution of self-reported SASS locale codes in these 1,742 schools was more even, with the greatest number being coded small city or town (419 schools or 24 percent of the 1,742). Of these SASS reported small city or town schools, all but 6 schools were coded urban fringe on the CCD file.

Another finding is the small percentage of rural schools with conflicting codes. Of the 2,692 schools coded rural on CCD only 139 (5 percent) were not coded rural or small town by SASS respondents. Of the 3,336 schools reported as rural on SASS, only 182 (5 percent) were not assigned rural or small town codes by CCD. This would agree with the findings of Huang's study (1993) of rural codes in CCD and SASS.

Reexamining locale code decisions

The above discussion has dealt primarily with the differences in the two locale coding schemes and the difficulty in comparing them. Since neither of the code assignments can be characterized as perfect, the two locale codes were checked for every school in the SASS public school survey in five states: Iowa, Maryland, Massachusetts, Oregon and Utah. These states had a total of 815 schools. Maryland was chosen because of the author's familiarity with the state, and the other four were chosen to get a sampling across the nation.

The CCD and SASS locale codes were checked against 1980 Census data. Each locale code was identified as being correct or wrong. The location and population of the towns of seven schools could not be determined, and these schools were dropped from the analysis, leaving 808 schools. Schools located in places within ten miles

of a city of greater than 50,000 people were determined to be in a suburban area in the SASS coding scheme. Schools more than 10 miles away from these cities but still in their SMSAs were counted as correct on the SASS survey if they were coded suburban or any of the appropriate city, town or rural codes depending on the place's population. Schools located in towns of greater than 10,000 people and less than 50,000 people were determined to be in a small town or city in the SASS coding scheme. The results of this study are presented in Table 2.

These results indicate that the Locale Desciptor on the SASS survey was correct for

Table 2--Verifying Locale Codes

State	Both correct	CCD correct SASS wrong	SASS correct CCD wrong	Both wrong	Total
IA	149 (84.2%)	21 (11.9%)	3 (1.7%)	4 (2.3%)	177 (100%)
MD	61 (43.3%)	58 (41.1%)	17 (12.1%)	7 (3.5%)	141 (100%)
MA	55 (35.3%)	51 (32.7%)	22 (14.1%)	28 (17.9%)	156 (100%)
OR	104 (64.2%)	34 (21.0%)	9 (5.6%)	15 (9.3%)	162 (100%)
UT	112 (65.1%)	51 (29.7%)	5 (2.9%)	4 (2.3%)	172 (100%)
Total	481 (59.5%)	215 (26.6%)	56 (6.9%)	56 (6.9%)	808 (100%)

Table 3--Counts of schools with incorrect locale codes by corrected CCD locale codes and corrected SASS community descriptor codes

Corrected SASS community descriptor codes / Frequency, Percent, Row Pct, Col Pct	1 Large City	2 Mid-Size City	3 Large City Fringe	4 Midsize City Fringe	5 Large Town	6 Small Town	7 Rural	Total
1 Rural or Farm	0 0.00 0.00 0.00	0 0.00 0.00 0.00	8 2.45 36.36 7.14	9 2.75 40.91 7.20	0 0.00 0.00 0.00	1 0.31 4.55 4.17	4 1.22 18.18 50.00	22 6.73
2 Small City	0 0.00 0.00 0.00	1 0.31 2.22 2.86	3 0.92 6.67 2.68	15 4.59 33.33 12.00	2 0.61 4.44 100.00	23 7.03 51.11 95.83	1 0.31 2.22 12.50	45 13.76
3 Medium City	0 0.00 0.00 0.00	12 3.67 80.00 34.29	2 0.61 13.33 1.79	1 0.31 6.67 0.80	0 0.00 0.00 0.00	0 0.00 0.00 0.00	0 0.00 0.00 0.00	15 4.59
4 Suburb of Medium City	0 0.00 0.00 0.00	0 0.00 0.00 0.00	0 0.00 0.00 0.00	29 8.87 93.55 23.20	0 0.00 0.00 0.00	0 0.00 0.00 0.00	2 0.61 6.45 25.00	31 9.48
5 Large City	0 0.00 0.00 0.00	19 5.81 79.17 54.29	0 0.00 0.00 0.00	5 1.53 20.83 4.00	0 0.00 0.00 0.00	0 0.00 0.00 0.00	0 0.00 0.00 0.00	24 7.34
6 Suburb of Large City	0 0.00 0.00 0.00	2 0.61 3.03 5.71	2 0.61 3.03 1.79	61 18.65 92.42 48.80	0 0.00 0.00 0.00	0 0.00 0.00 0.00	1 0.31 1.52 12.50	66 20.18
7 Very Large City	18 5.50 100.00 85.71	0 0.00 0.00 0.00	0 0.00 0.00 0.00	0 0.00 0.00 0.00	0 0.00 0.00 0.00	0 0.00 0.00 0.00	0 0.00 0.00 0.00	18 5.50
8 Suburb of Very Large City	3 0.92 2.83 14.29	1 0.31 0.94 2.86	97 29.66 91.51 86.61	5 1.53 4.72 4.00	0 0.00 0.00 0.00	0 0.00 0.00 0.00	0 0.00 0.00 0.00	106 32.42
Total	21 6.42	35 10.70	112 34.25	125 38.23	2 0.61	24 7.34	8 2.45	327 100.00

66.4 percent of the schools investigated, whereas the locale code on the CCD file was correct for 86.1 percent of the schools. Or put another way, the SASS Locale descriptor was wrong in twice as many instances as the CCD assigned locale code.

The schools which initially had incorrect locale codes assigned to them by NCES or whose respondent chose the wrong community descriptor codes were subsetted and a cross tabulation performed by the corrected locale codes. These data are presented in Table 3.

Table 3 indicates that schools located in suburban or fringe areas are more likely to be coded incorrectly. Of the 327 schools with incorrect locale codes, 237 (72.5 percent) were found to be in an SMSA outside the central city, and 203 (62.1 percent) were found to be within 10 miles of a city.

These problems are due to the difficulty in defining suburban areas. This difficulty occurs on the SASS survey when respondents do not have a common understanding of what "suburban" means. Even when there are clear operational definitions, problems exist in the CCD locale code assignment process. These problems appear to be in matching mailing addresses (i.e., suburban post offices) with census place names and identifying their central city.

References:

Bushery, J., Royce, D., and Kaspryzk, D., "The Schools and Staffing Survey: How Reinterview Measures Data Quality", Proceedings of the Section on Survey Research Methods, American Statistical Association, Alexandria, Virginia, 1992

Huang, G., Defining the Rural: Johnson Code and Self-Report by School Personnel, Unpublished manuscript, 1993

Johnson, F., Assigning Type of Locale Codes to the 1987-88 CCD Public School Universe, NCES Technical Report, CS89-194, July 1989