

---

**NLSY79 APPENDIX 6:**  
**URBAN-RURAL AND SMSA-CENTRAL**  
**CITY VARIABLES**

---



## 1979-2000 URBAN-RURAL VARIABLES

### *1979-1996 Urban-Rural Residence*

Through 1996, the NLSY79 Urban-Rural Residence variables were constructed using the total and urban population data for the county of residence from the 1970 Census of Population Characteristics of the Population (for NLSY79 1979-1982) and from the 1980 Census of Population and Housing (for NLSY79 1983-1988). These data are included in the 1977 and 1983 *County and City Data Book* data files respectively.

The urban population consists of all inhabitants of urbanized areas. An urbanized area is defined as a central core or city and its adjacent, closely settled territory which have a combined total population of 50,000 or more (with exceptions in Alaska, New York, the New England states, and Wisconsin). These definitions have remained largely unchanged since 1950. For more detailed definitions and comments on exceptions, refer to the *U.S. Bureau of the Census 1970/1980 Census of Population, Characteristics of the Population, Number of Inhabitants* (Series PC80-1-A).

Calculation of the 1979–96 NLSY79 Urban-Rural Residence variables involved the following two steps:

- 1) The percent of urban population was calculated by dividing the urban population of the county by the total county population and multiplying by 100.
- 2) Rural counties for the NLSY79 variables were defined as those with between 0-49% urban population. Urban counties were defined as those with 50% or more urban population.

### *1998-2000 Urban-Rural Residence*

Beginning with the 1998 release, the information needed to calculate whether the county was urban/rural was no longer available in the *County and City Data Book* releases. For 1998, a respondent is coded urban if living in an urbanized area or in a place with greater than 2500 population. There is no regard to county.

***User Notes:*** The method of calculating urban vs rural prior to the 1998 data release designated an entire county as either urban or rural. For example, if a county contained multiple respondents, some living in an urban-like area and others in rural areas, they all received the same designation depending on the degree of urbanization of the county of residence. In 1998 and 2000, the method of calculating urban/rural allows a respondent living in a rural area of a county to be coded rural, while another respondent in an urban area of that same county can be coded urban. **The net affect appears to be an increase in respondents living in rural areas.**

For all survey years, non-interview cases are set to a -5 value on the Urban-Rural Residence variable. All other missing cases are valid skips on these variables.

For a detailed discussion of the procedures used for hand-editing and merging respondent data with other data in creating the geographic variables, see “Appendix 10: Geocode Documentation” in the NLSY79 *Geocode Codebook Supplement*.

## 1979-2000 SMSA-CENTRAL CITY VARIABLES

The NLSY79 SMSA-Central City variables are constructed using data for the SMSA/MSA and Place Description (PD) included in the City Reference File (CRF) data files. The 1973 CRF was used for NLSY79 1979-1982 SMSA-Central City variables; the 1982 CRF was used for 1983 variables; 1983 CRF was used for 1984-1987 variables; the 1987 CRF was used for the 1988-1992 variables; and the 1992 CRF was used for the 1993-1998 variables. See below for information on the 2000 variables. (For a more detailed discussion of changes in official terminology and geographic designations and comparability across years, see “Appendix 10: Geocode Documentation” in the *NLSY79 Geocode Codebook Supplement*.)

### *1979-1998 SMSA-Central City Residence*

Through 1998, calculation of the SMSA-Central City Residence variables involved the following two steps:

- 1) A respondent’s SMSA/MSA and PD were assigned based on the county, state, and zip code of current residence. (For a detailed discussion of the procedures used for hand-editing and merging respondent data with other data in creating the geographic variables, including the SMSA/MSA, see “Appendix 10: Geocode Documentation” in the *NLSY79 Geocode Codebook Supplement*.)
- 2) Based on their PD and SMSA/MSAs, respondents were assigned to one of the following categories:
  - Respondents not residing in an SMSA/MSA were defined as “not in SMSA.”
  - Respondents residing in an SMSA/MSA, but not in a central city of an SMSA/MSA according to the PD, were defined as “SMSA, not central city.”
  - Respondents for whom the PD leads to an ambiguous central city residence status were defined as “SMSA, central city not known.” These cases generally resulted from zip codes that cover more than one geographically-defined area.
  - Respondents residing in both an SMSA/MSA and the central city of an SMSA/MSA according to the PD were defined as “SMSA in central city.”

Non-interview respondents were assigned a value of -5. Respondents with an ambiguous SMSA/MSA residence status or missing values for reasons other than non-interview are valid skips on these variables.

### *2000 SMSA-Central City Residence*

Beginning with the 2000 release, the calculation of the central city variable was revised slightly. The process still consists of consists of two components. The first delineation is MSA/non-MSA residence. For respondents in an MSA, the residence is further defined by placement inside or outside a central city as defined by the Census Bureau.

Respondents who live in an MSA who have a quality of match code of either manual edit or zip centroid are evaluated before assigning the central city code. If the street (manual edit match) or the zip code area (zip centroid match) falls entirely within the boundaries of the central city they are coded as central city. If the street or zip code area fall completely outside the central city boundary, they are coded as not in a central city. Otherwise, these respondents are coded as central city unknown.

A further complication is that Maptitude, the software currently being used, does not include central city boundary files for several areas. Regardless of quality of match code, respondents in those areas are

coded as being in a MSA with an unknown central city status since the mailing address is insufficient to reveal whether the respondent resides within the central city boundaries. When updates of Maptitude are available, the city boundaries will be updated.

