Methods – Neighborhood Deprivation Index Data

A Neighborhood Deprivation Index (NDI) for each Census tract in the U.S. was created using factor analysis to identify key variables from 13 measures in the following dimensions of socioeconomic (SES) status: wealth and income, education, occupation, and housing conditions.[1] The specific 13 measures are:

- Wealth and income
    - Median household income (dollars)
    - Percent of households receiving dividends, interest, or rental income
    - Percent of households receiving public assistance
    - Median home value (dollars)
    - Percent of families with incomes below the poverty level
- Education
    - Percent with a high school degree or higher
    - Percent with a college degree or higher
- Occupation
    - Percent in a management, business, science, or arts occupation
    - Percent unemployed
- Housing conditions
    - Percent of households that are female headed with any children under 18
    - Percent of housing units that are owner occupied
    - Percent of households without a telephone
    - Percent of households without complete plumbing facilities

These 13 variables were obtained from the Census Bureau's 5-year American Community Survey (ACS) data for 2013-2017. Factor analysis was then used to generate the NDI. This involved the following steps:

- Log transform median household income and median home value
- Reverse code percentages so that higher values represent more deprivation. For example, the percent of housing units that are owner occupied was converted to the percent of housing units that are not owner occupied.
- Z-standardize the percentages
- Run a factor analysis using Promax (oblique) rotation and a minimum Eigenvalue of 1
- Calculate the factors using only variables with a loading score > 0.4 for the first factor (this removed three variables: the percent of housing units that are owner occupied, the percent of households without a telephone, and the percent of households without complete plumbing facilities)
- Calculate Cronbach's alpha correlation coefficient among the factors and verify values are above 0.7.
- Use the resulting calculation of the first factor as the Neighborhood Deprivation Index (NDI)

---

[1] Diez Roux A V, Mair C, Roux AVD, Mair C, Diez Roux A V, Mair C. Neighborhoods and health. *Annals of the New York Academy of Sciences* 2010; 1186: 125–45.

The final NDI calculation is based on 10 of the 13 variables included in the factor analysis. Those variables are median household income; percent of household receiving dividends interest or rental income; percent of households receiving public assistance; median home value; percent of families with incomes below the poverty level; percent with a high school degree or higher; percent with a college degree or higher; percent in a management, business, science, or arts occupation; percent unemployed; and percent of households that are female headed with any children under 18. The percent of housing units that are owner occupied, the percent of households without a telephone, and the percent of households without complete plumbing facilities loaded poorly during the factor analysis and, thus, are excluded. NDI values range from -3.6 to +2.8. Higher values indicate more neighborhood deprivation (lower socioeconomic status). We also created a categorical variable representing NDI quintiles weighted by tract population (so that 20% of the population is in each quintile group).

The NDI is based on factor analysis of tract-level variables at the national level. For research involving a particular study area, there might not be as much variation in deprivation levels as there in nationally. Also, variation in one or more of the variables may be more pronounced regionally or locally than nationally. For example, in an area with a high cost of living, regional or local variations in household income and home value may be more significant than they are at a national level. Researchers could generate a specific index for their study area (or a larger area that includes the study area) to characterize regional or local variation in deprivation. The original 13 variables are included in the dataset for this purpose. Working with a subset of the national data, researchers could generate a custom version of the NDI using the steps described above. The generated index might include a different subset of the original variables and, thus, might emphasize slightly different aspects of neighborhood deprivation.