NORTHWEST AREA FOUNDATION
SOCIAL INDICATORS SURVEY
SEPTEMBER-DECEMBER 2003

VARIABLE NOTATION AND WEIGHTING
INFORMATION

Note: This is a listing of explanations for several variables beyond what is contained in the codebook. Each one is identified by its variable name and variable labe, and is presented in the same order that it appears in the datasets.

I. VARIABLE NOTATION.

RESPNUM$: Respondent number.
This respondent ID number is automatically generated by OSRL's WinCATI system.

TIME$: Duration of interview (Minutes).
The duration of the interview in minutes is automatically generated by OSRL's WinCATI system.

IACOUNTY: County (Iowa).
The variable combines the county data from COUNTY2 and COUNTY2A into a single variable.

MNCOUNTY: County (Minnesota).
The variable combines the county data from COUNTY3 and COUNTY3A into a single variable.

SDCOUNTY: County (South Dakota).
The variable combines the county data from COUNTY7 and COUNTY7A into a single variable.

COMM1X: Number of years lived in place (Categorized).
This variable collapses ratio-level data from COMM1 into six categories. The codebook presents these categories.

COMM4: One thing to change about place (Coded open-end answers).
This variable contains the categorized responses to COMM4. Codes were developed by the Principal Investigator (see Section 7 of this binder for the detailed code categories). Minus signs (-) at the end of the category indicate negative responses, while plus signs (+) indicate positive responses.

COMM4A: One thing to change about place (Recoded).
This variable collapses the data COMM4 detailed codes into general categories. The collapsed COMM4 categories are specified in parentheses at the end of the COMM4A categories.

COMM5: One best thing about living in place (Coded open-end answers).
This variable contains the categorized responses to COMM5. Codes were developed by the Principal Investigator (see Section 7 of this binder for the detailed code categories). Minus signs (-) at the end of the category indicate negative responses, while plus signs (+) indicate positive responses.

COMM5A: One best thing about living in place (Recoded).
This variable collapses the COMM5 detailed codes into general categories. The collapsed COMM4 categories are specified in parentheses at the end of the COMM5A categories.

CITIZEN2X: Years in the U.S. (Categorized).
This variable collapses ratio-level data from CITIZEN2 into six categories. The codebook presents these categories.

PARTSCAL: Civic and community participation scale.
The Civic and Community Participation Scale was calculated using the variables PART1, PART2, PART4, PART5, PART7, PART8, PART9, PART10, PART11, PART12, PART13, PART14, PART16, PART18, PART20, PART22, and PART25. For each variable, a response of "yes" was coded as 1, and all other responses were coded as zero. Each respondent's answers to each question were summed, resulting in a 0-17 scale. High scores indicate greater civic and community participation.

SKILLSCAL: Leadership skills scale.
The Leadership Skills Scale was calculated using the variables SKILL1, SKILL2, SKILL3, and SKILL4. For each variable, a response of "yes" was coded as 1, and all other responses were coded as zero. Each respondent's answers to each were summed, resulting in a 0-4 scale. High scores indicate greater leadership skill.

NEIGH4B: Worked w/neighbors to solve problem (Recoded).
For this variable, respondents who volunteered that they had no neighbors were recategorized as "NO ANSWERS".

HHNUMX: No# of people in HH (Categorized).
This variable collapses ratio-level data from HHNUM into four categories. The codebook presents these categories.

HHNUMX2: HH size (Two categories).
This variable collapses ratio-level data from HHNUM into two categories: single-person households and multiple-person households.

KIDINHH: Children present in household? (Recoded).
This variable uses data from the variables HHNUM and HHKIDS1 to create a variable that identifies households with and without children.

PERSSCAL: Financial marginalization scale.
The Financial Marginalization Scale was calculated using the variables PERS1, PERS2, and PERS3. For PERS1, the responses "not very/not at all satisfied" were coded as 1, and all other responses were coded as zero. For PERS2 and PERS3, a response of "yes" was coded as 1, and all other responses were coded as zero. Each respondent's answers to each question were summed, resulting in a 0-3 scale. High scores indicate greater financial marginalization.

AGEX: Age (Categorized).
This variable collapses ratio-level data from AGE into six categories.

RACEA: R's Race (Other=PI's assigned categories).
This variable captures the respondent's race/ethnicity as reported in the variable RACE and includes modifications requested by the Principal Investigator for respondents who volunteered answers that fell outside the standard answer categories. Interviewers classified these responses as "OTHER-SPECIFY" (code 96), and then recorded the content of the respondent's answers. After reviewing these materials, the Principal Pnvestigator identified 29 cases for review and recategorization:

Interviewer recorded as "other-specify"	Coded as
5 different types	95-Mixed Race
Amer indian/white	95-Mixed Race
American indian, caucasian, and black	95-Mixed Race
Anglo	1-White/Caucasian
Asian, american	3-Asian/Pac. Islander
Black irish	2-Black/African American
Black/white native	95-Mixed Race
Born in iran, but considers himself american, doesn't want to be classified in a group	7-Middle Eastern/Arab
Caucasian/pacific islander	95-Mixed Race
Caucasion / native american	95-Mixed Race
Declined more info	97-Refused
European descent mixed	1-White/Caucasian
Halh canjun half french	1-White/Caucasian
Heinz 54. American	1-White/Caucasian
Italian	1-White/Caucasian
Italian	1-White/Caucasian
Italian american	1-White/Caucasian
Italian-american	1-White/Caucasian
Jewish	1-White/Caucasian
Jewish	1-White/Caucasian
Middle eastern	7-Middle Eastern/Arab
Native am	5-American Indian/NativeAmerican
Norwegan	1-White/Caucasian
Pacific islander	3-Asian/Pacific Islander
Phillipino	3-Asian/Pacific Islander
Polish	1-White/Caucasian
Scandinavian american	1-White/Caucasian
Some philiphine in me not much	95-Mixed Race
White mixed	1-White/Caucasian

RACEX: R's race (Recoded).
This variable collapses the data from RACE into two categories: white, non-Hispanic and other.

EDUCB: Education--highest level completed (Recoded).
This variable collapses the data from EDUC into five categories. The codebook presents these categories.

EMPLY4AX: Miles between workplace and place (Categorized).
This variable collapses ratio-level data from EMPLOY4A into five categories. The codebook presents these categories.

INCOME: HH income.
This variable recategorizes data from INCOME1 through INCOME5 into a single variable with six categories. The codebook presents these categories.

SMPCITY: City estimate (from Genesys sample file).
This is an estimate of the city in which the randomly sampled telephone number is located. This estimate is part of the random-digit dial (RDD) sample file that is generated using the Genesys software package. The estimate is based on the city where the largest percentage of telephone exchanges are located. The percentage of the telephone exchanges that fall in a particular city is calculated by Genesys by matching ZIP codes to exchanges.

SMPSTATE: State (from Genesys sample file).

This is the state in which the respondent's telephone number is located. This information was included in the RDD sample file generated by the Genesys software package.

WTREGION: Region-Level Weight (Age/Sex/HH/State Pop).
This weight was calculated using population data from all eight states to create a region-level population estimate. See the "Weighting Documentation" section below for additional information.

WTSTATE: State-Level Weight (Age/Sex/HH).
This weight was calculated using state-level population data. See the "Weighting Documentation" section below for additional information.

II. WEIGHTING INFORMATION.

Weights were calculated on two levels (state and region) using three demographic variables (age, sex, and household size). State weights are appropriate for analyses restricted to individual states. Region weights are appropriate for analyses of all eight states combined. If other combinations of states are required, then additional weights should be calculated.

Since no one source provided all the weighting data required, two data sources were consulted in the calculation of the weights. Data on household size was taken from the 2000 US Census (table available in Summary File 1 at http://www.census.gov). Data on age and sex was obtained from the Census Bureau's Population Estimates Program's 2002 population estimates (data available at http://eire.census.gov/popest/estmates_dataset.php).

State-level weights (variable name WTSTATE) were calculated by multiplying the individual weight variables[1]:

WTSTATE = Age Weight[2] X Sex Weight X Household Size Weight[3].

To account for the difference in state population sizes and each state's proportional contribution to the regional population, an additional weight was added to the region-level weighting equation (variable name WTREGION). This weight captured the percentage of the regional population that resides in the respondent's state, and data from the Census Bureau's Population Estimates Program's 2002 population estimates was used in the development of this weight. The region-level weights were calculated by multiplying the individual-level weight variables:

WTREGION = Age Weight X Sex Weight X Household Size Weight X State Population Weight.

For cases that were missing data on one or more of the weighting variables, partial weights were calculated for cases using the data that was available. Data for two of the three weighting variables was available for twenty-five of the twenty-eight cases excluded by the first missing data strategy. For these cases, the available data was multiplied and substituted for the zero value used in the first missing data strategy. Of the three remaining cases, two had data for one of the three weighting variables. For these cases, the weight for that individual variable was substituted for the zero value used in the first missing data strategy. Finally, a single case did not have data for any of the weighting variables, and in this case the zero weight was preserved.

---

[1] Weight calculations for both the state-level and region-level weights are available on request.
[2] The following age categories were used in the weighting: 18 to 29, 30 to 39, 40 to 49, 50 to 59, 60 to 69, 70 and older.
[3] The following household size categories were used in the weighting: single-person household, multiple-person household.

The weights were checked by randomly selecting a minimum of five summary weights from each state and comparing the value in the data set against hand calculations using the raw data.

There are two important factors to note with regards to the weights. First, using the weights will change the sample size. Second, the regional data are not generalizable to any known population without implementing the region-wide weights (WTREGION).