ICPSR | INTER-UNIVERSITY CONSORTIUM FOR POLITICAL AND SOCIAL RESEARCH

ICPSR 37513

Well-Being and Basic Needs Survey, United States, 2017

Stephen Zuckerman
Urban Institute

User Guide

Terms of Use

Information about Copyrighted Content

Well-Being and Basic Needs Survey (WBNS)

Data Documentation and User's Guide

Prepared by the Urban Institute August 2019

Background on the Well-Being and Basic Needs Survey

Purpose and Funding

In December 2017, the Urban Institute launched the Well-Being and Basic Needs Survey (WBNS), a nationally representative, internet-based survey of nonelderly adults designed to monitor changes in individual and family well-being during a time when policymakers are considering significant changes to federal safety net programs serving low-income families. The survey collects information on a broad array of topics related to health and material hardship, including health insurance, housing, food security, employment, family income, program participation, and family financial security. The survey is made possible with funding from the Robert Wood Johnson Foundation as part of the Urban Institute's *From Safety Net to Solid Ground* project.

Sample Characteristics, Sample Size, and Timing

The WBNS is fielded annually in December.[1] For each round, a stratified random sample of approximately 7,500 adults ages 18 to 64 is drawn from the KnowledgePanel®, a probability-based, nationally representative Internet panel maintained by Ipsos (formerly maintained by GfK Custom Research). The WBNS includes an oversample of adults with household incomes under 150 percent of the federal poverty level to improve the precision of estimates for this group. The December 2018 round also included an oversample of noncitizens to support analyses of current policy issues affecting immigrant families.

Sample Frame and Survey Administration

Current recruitment into the KnowledgePanel® is based on an address-based sampling frame that covers nearly all residential addresses in the United States. In earlier periods, recruitment was carried out by telephone based on random-digit dialing sampling techniques. Approximately 55,000 U.S. households participate in the panel, including those with and without Internet access. Laptop computers and Internet access are provided to panel members if needed, and respondents can complete the online surveys in English or Spanish.

The WBNS sample is drawn from the individuals in the KnowledgePanel®. Although fresh WBNS samples are drawn each round, the same individuals may be selected for different rounds of the survey. The American Association for Public Opinion Research (AAPOR) cumulative response rate for the WBNS is

---

[1] The field period for the first round of the WBNS began on December 14, 2017, and ended January 5, 2018, with nearly 90 percent of respondents completing the survey in December.

the product of the panel household recruitment rate, the panel household profile rate, and the WBNS completion rate—roughly 4 percent each round.

It takes about 14 minutes to answer the WBNS questions in each round. Detailed sociodemographic information collected by Ipsos on panel members enables post-stratification weighting based on a broad set of measures. More information about the design of the WBNS is available at https://www.urban.org/research/publication/well-being-and-basic-needs-survey.

Construction of the WBNS Data Files

Variables for Original Survey Questions

Variables that begin with "Q" followed by a number provide data from the survey questions. Almost all survey question variables are coded as shown in the survey instrument. However, some variables have been recoded or suppressed to protect the confidentiality of respondents. Recoded variables in the 2017 WBNS include:

- Number of children in the household (Q3): Top-coded at 6 or more children. Coded as -99 if there is an invalid response due to misreporting (e.g., an unlikely large number of children).
- Number of adults ages 19 and older other than the respondent in the household (Q5): Top-coded at 6 or more adults.
- Housing type (Q7): Collapsed categories for mobile home; boat, RV, or van; and something else.
- Number of times moved in the past year (Q16a): Top-coded at two or more times.
- Number of hours worked per week at a main job (Q40) and at other jobs (Q41): Top-coded at 80 hours per week.
- Family income as a percentage of the federal poverty level (Q51): Coded as -99 if there is missing data on family composition questions (Q1, Q2, Q3, and Q4) that are used to calculate family size, which is used to identify the federal poverty level for the respondent's family.

Suppressed variables in the 2017 WBNS include:

- Age of and respondent's relationship to children in the household (Q4)
- Age of and respondent's relationship to other adults in the household (Q6)
- Variables containing write-in responses (Q7, Q32a, Q44, and Q59)
- Coverage through the Indian Health Service (Q31)

For survey questions that were not asked of respondents, variables have a missing value (.). Survey questions that a respondent did not answer are either labeled "Refused" and have a value of -1, or have a separate variable to indicate whether the question was skipped (e.g., Q17_Refused).

All survey data are subject to measurement error, and two variables in the 2017 WBNS in particular should be used with caution. The variable for number of adults (other than the respondent) living in the household (Q5) may be overreported if some respondents included themselves in that count. In addition, we do not recommend using the variable for the self-reported share of income paid toward housing costs (Q10), which does not benchmark well against federal survey data.

Variables for Household Demographic Profile Data

To participate in the panel, members are asked to complete a detailed demographic profile for themselves and their households. These data facilitate post-stratification weighting, targeted oversampling, and subgroup analysis. Most profile variables in the WBNS public use files begin with the letters "pp" and are included after the survey question variables. Exceptions include household internet access (xacsnet) and sexual orientation and gender identity (lgbt), which combines information from two profile questions into a single variable.

Suppressed profile variables include:

- Geographic data below the Census region
- Diagnosed medical conditions
- Profile variables that duplicate information collected in the survey: number of children in different age groups in the household, number of adults in the household, household size, housing type, homeownership status, marital status, and employment status
- Detailed levels of educational attainment

The profile variable for annual household income (ppincimp) should not be confused with the family income variable (Q51) collected in the survey or the imputed income variables (those beginning with "fpl"). Household income is first collected when the respondent joined the KnowledgePanel and is updated annually; it includes the income of all members of the household. In contrast, the family income data collected in the WBNS are based on income reported for the year prior to the survey and only include the income of the respondent, spouse or partner if applicable, and their own children under age 19 who are living in the household.

The profile variable for citizenship status (ppp20197) is coded as missing (.) if it was not asked and -1 if the respondent did not answer the question. The variable for sexual orientation and gender identity (lgbt) is coded as missing if the questions were not asked, the respondent did not answer one or both questions, or the respondent provided insufficient information to determine if they are gay, lesbian, bisexual, or transgender.

Constructed Variables

Although fresh WBNS samples are drawn each round, the same individuals may be selected for different rounds of the survey. We replace a panel identifier with a random identification number (id) for each respondent to protect their confidentiality. This identification number can be used to link responses of the same panel member across rounds. However, panel members who completed the survey in one round do not necessarily appear consistently in other rounds.

Family composition and family income

We impute family income as a percentage of the federal poverty level (FPL),[2] as well as variables used to define the family unit. These imputed variables include respondent marital status (married), an indicator

---

[2] FPL is based on the annual guidelines established by the U.S. Department of Health and Human Services for the 48 contiguous states.

of whether the respondent is living with a partner (partner) if not married, and the number of children under age 19 in the household (num_child_hh), if this information is missing in Q1, Q2, or Q3. We also impute the age of each child in the household, and whether the respondent is the parent or guardian of each child in the household if this information is missing in Q4, which is excluded from the public use file to protect the confidentiality of respondents. These family composition variables are used to construct family size (famsize) for respondents ages 19 or older as well as 18-year-old respondents who are not single and childless. For these respondents, the family unit includes themselves, their spouse or partner, and their own children under age 19 who are living with them. For single, childless 18-year-olds, the family unit includes themselves and the number of parents (num_parent) and number of siblings (num_sibling) under age 19 who are living with them; these variables are derived from Q4 and Q6, which are excluded to protect the confidentiality of respondents. Number of parents and number of siblings are coded as missing for all respondents who are not single, childless 18-year-olds.

Family income as a percentage of FPL is then imputed if any information used to construct family size is missing, or if family size is not missing but the respondent did not answer the question on family income (Q51). We use a regression-based multiple imputation approach, creating five imputation variables for each income/FPL category. Family income or family composition data is missing for about 6.5 percent of respondents in each round. The variable _mi_miss flags any observations with imputed data for family composition, family income, or health insurance coverage (see below). As noted above, Q51 is recoded as -99 if there is any missing data on family composition variables used to define the family unit.

For all constructed family composition variables, missing flags are provided to identify observations with imputed data. Variables ending in "_miss" indicate whether any imputation was performed for the selected variable. Variables ending in "_imp" indicate whether the observation is only coded as "1" based on imputation of number of children in the household, ages of children in the household, and/or respondent relationship to children in the household. Table 1 describes the source variables and missing data flags for each imputed family and household composition variable.

Constructed variables for family income as a percentage of FPL are provided for the following categories:

- Family income below 50% of FPL (fpl_lt50)
- Family income of 50% to less than 100% of FPL (fpl50_100)
- Family income of 100% to less than 150% of FPL (fpl100_150)
- Family income of 150% to less than 200% of FPL (fpl150_200)
- Family income of 200% to less than 250% of FPL (fpl200_250)
- Family income of 250% to less than 300% of FPL (fpl250_300)
- Family income of 300% to less than 400% of FPL (fpl300_400)
- Family income of 400% to less than 500% of FPL (fpl400_500)
- Family income of 500% to less than 600% of FPL (fpl500_600)
- Family income of 600% of FPL or higher (fpl_600pl)

Variables are provided for each imputation of family income as a percentage of FPL (e.g., _1_fpl_lt50, _2_fpl_lt50, _3_fpl_lt50, _4_fpl_lt50, _5_fpl_lt50).

<u>Table 1. Imputed Family and Household Composition Variables in the WBNS</u>

Description	Constructed variable	Source variable(s)	Missing data flag	Missing data flag interpretation
FAMILY COMPOSITION VARIABLES				
Marital status	married	Q1	mar_partner_miss	Any imputation of marital/partner status
Living with a partner	partner	Q2	mar_partner_miss	Any imputation of marital/partner status
Any own children under 19	any_ownchild	Q3, Q4b-c	ownchild_miss	Any imputation of relationship to children in household
			any_ownchild_imp	Coded "1" based on imputation of number of children in household and/or relationship to children
Number of own children under 19	num_ownchild	Q3, Q4b-c	ownchild_miss	Any imputation of relationship to children in household
Family size	famsize	Q1, Q2, Q3, Q4b-c, Q6b-c	famsize_miss	Any imputation of family composition variables Q1, Q2, Q3, or Q4b-Q4c
Any own child age 0-2	any_ownchild_0_2	Q3, Q4a-c	any_ownchild_0_2_imp	Coded "1" based on imputation of number of children in household, age of children, and/or relationship to children
Any own child age 3-5	any_ownchild_3_5	Q3, Q4a-c	any_ownchild_3_5_imp	
Any own child age 6-12	any_ownchild_6_12	Q3, Q4a-c	any_ownchild_6_12_imp	
Any own child age 13-18	any_ownchild_13_18	Q3, Q4a-c	any_ownchild_13_18_imp	
CHILDREN IN HOUSEHOLD VARIABLES				
Number of children under 19 in the household	num_child_hh	ppage, Q3	num_child_hh_miss	Any imputation of number of children in household
Any child in the household age 0-2	any_childhh_0_2	Q3, Q4a	any_childhh_0_2_imp	Coded "1" based on imputation of number of children in household and/or age of children
Any child in the household age 3-5	any_childhh_3_5	Q3, Q4a	any_childhh_3_5_imp	
Any child in the household age 6-12	any_childhh_6_12	Q3, Q4a	any_childhh_6_12_imp	
Any child in the household age 13-18	any_childhh_13_18	ppage, Q3, Q4a	any_childhh_13_18_imp	
Any child in the household age 0-6	any_childhh_0_6	Q3, Q4a	any_childhh_0_6_imp	
Any child in the household age 5-18	any_childhh_5_18	ppage, Q3, Q4a	any_childhh_5_18_imp	

Health insurance coverage status

Questions on health insurance in the WBNS are drawn from the Urban Institute's Health Reform Monitoring Survey (HRMS), which uses a modified version of a question from the American Community Survey that asks respondents whether they are covered by any of several types of coverage. Both the WBNS and HRMS also include a verification question to determine whether those who did not report any type of coverage are uninsured. Those who report having some form of coverage are asked to provide a write-in response describing the type of coverage they have. We apply an editing process to assign coverage types based on these responses.

For those who report "any other type of health insurance coverage or health coverage plan" (Q31) or that they "have some kind of health insurance" (Q32) but do not select or describe a specific, valid type of coverage, we impute coverage status by drawing on data from the previous seven rounds of the HRMS, which contains more detailed questions on health insurance. We provide variables for health insurance coverage status at the time of the survey (insured and uninsured) along with the five imputations associated with each of these variables (e.g., _1_insured, _2_insured, _3_insured, _4_insured, _5_insured). Although coverage status variables are provided, we are not as confident in the accuracy of reported coverage type, which is often a concern in surveys,[3] and have not constructed coverage type variables for the public use files.

Material hardship

Constructed variables for the seven measures of material hardship used in previous Urban Institute WBNS briefs, as well as indicators of any hardship in the past year (any_hardship) and more than one type of hardship in the past year (mult_hardship), are included in the public use file. The seven measures of material hardship are based on experiences during the past year and are grouped into four domains:

- *Housing*: 1) the household did not pay the full amount of the rent or mortgage or was late with a payment because it could not afford to pay (rentprob_y); 2) the respondent was forced to move by a landlord, bank or other financial institution, or the government (move_evic).
- *Utilities*: 3) the household was not able to pay the full amount of the gas, oil, or electricity bills (utilprob_y); 4) the gas or electric company turned off service, or the oil company would not deliver oil (utilsusp_y).
- *Food security*: 5) the household was food insecure based on responses to the six-item short form of the USDA's Household Food Security Survey Module (food_insec).
- *Health care*: 6) the respondent had unmet needs for medical care because of costs (unmetcare_y); 7) the family had problems paying medical bills (billprob_y).

Food security in the WBNS is based on the cumulative number of affirmative responses to the six-item short-form of the household food security module from the Current Population Survey (CPS) Food Security Supplement. The short form is a six-question subset of the full 18-item household food security module. Unlike the CPS, which skips many respondents out of the food security questions based on their responses to income and screener questions, the food security questions in the WBNS are asked of all respondents. The score used to determine food security ranges from zero to six. Respondents with

---

[3] Pascale, Joanne. 2008. "Measurement Error in Health Insurance Reporting." *Inquiry* 45: 422-37.

scores of two or higher are classified as "food insecure." We code respondents who reported "don't know" or "refused" to all the questions in the scale as missing when constructing the food security measure.

Using the WBNS Data Files

Weights and Survey Design Variables

Each data file includes a variable providing a post-stratification weight designed to make the sample representative of the national population of adults ages 18-64. WBNS weights reflect the probability of sample selection from the KnowledgePanel® and post-stratification to the characteristics of nonelderly adults in the United States based on benchmarks from the Current Population Survey and American Community Survey for the following measures: age by gender; race and ethnicity; educational attainment; presence of children ages 0-17 in the household; census region; residence in a metropolitan area; homeownership status; family income as a percentage of FPL; family composition (married or living with a partner by presence of children under age 19 in the family); primary language; and internet access by age. Weights are trimmed and scaled to sum to the sample size of total survey respondents. The design effect for the first round of the WBNS is 1.83.

To obtain accurate standard errors and enable significance testing, we have created a masked variable for grouping primary sampling units for variance estimation (varunit). Survey options in statistical software packages—such as the command *mi svyset varunit [pw=weight]* in Stata—can be used for producing weighted estimates and accurate standard errors, as well as including observations for which family income or health insurance coverage status have been imputed in the analysis.

Contact Information

For questions about the WBNS and the public use data files, please send an email to wbns@urban.org.