

September 2017

National Financial Well-Being Survey

Public Use File User's Guide

Table of contents

Table of contents.....	1
1. Introduction.....	2
2. Description of the survey.....	5
2.1 National Financial Well-Being Survey background and objectives.....	5
2.2 Sampling, data collection, and weighting.....	7
2.3 Survey overview	31
3. How to use the Public Use File.....	33
3.1 Description of the Public Use File	33
3.2 Dataset and code.....	33
3.3 Using the study weights.....	34
3.4 Nonsubstantive value codes	41
4. Variable documentation	42
4.1 Imputation for HOUSERANGES.....	42
4.2 Data cleaning	44
4.3 Created and modified variables.....	48
Appendix A: Survey instrument	68

1. Introduction

An essential part of the Consumer Financial Protection Bureau’s (CFPB or Bureau) mission is educating and empowering consumers to take control over their financial lives. In addition to a safe, transparent marketplace, consumers need to navigate that marketplace effectively. Numerous provisions of the Dodd-Frank Wall Street Reform and Consumer Protection Act of 2010 (Dodd-Frank Act) charge the Bureau with working to improve the financial literacy of consumers in America.¹

Over the past several years, a key part of CFPB’s strategy for improving financial capability has revolved around a rigorous set of research activities designed to define and measure “success” for financial literacy initiatives.² This work has included the development of an evidence-based,

¹ An important financial literacy mandate is set forth in Section 1013(d) of the Dodd-Frank Act, which directs the CFPB, through its Office of Financial Education, to develop and implement initiatives intended to “educate and empower consumers to make better informed financial decisions” and to “develop and implement a strategy to improve the financial literacy of consumers....consistent with the National Strategy for Financial Literacy....” (12 U.S.C. § 5493(d)(1)&(2)). The Dodd-Frank Act also mandated the creation of other offices within the Bureau that are responsible for, among other things, developing financial education and policy initiatives to support the financial well-being of particular segments of the consumer population (12 U.S.C. § 5493(b),(e),(g)).

² Another key component of CFPB’s strategy to improve financial literacy is by providing consumers with tools and resources to increase their financial capability.

consumer-driven definition of financial well-being³ and a tool for measuring it, the CFPB Financial Well-Being Scale.⁴

Building on the development and validation of the financial well-being scale, the CFPB fielded the National Financial Well-Being Survey.⁵ The survey uses CFPB's validated financial well-being scale to measure the financial well-being of a national sample of adults selected to reflect the U.S. population. The survey also includes measures of individual and household characteristics, income and employment, savings and safety nets, financial experiences, and behaviors, skills and attitudes that have been hypothesized to influence adults' levels of financial well-being.

The National Financial Well-Being Survey was conducted in English and Spanish via web mode between October 27, 2016 and December 5, 2016. Overall, 6,394 surveys were completed: 5,395 from the general population sample and 999 from an oversample of adults aged 62 and older. The survey was designed to represent the adult population of the 50 U.S. states and the District of Columbia. The survey was fielded on the GfK KnowledgePanel®. The KnowledgePanel sample is recruited using address-based sampling and dual-frame landline and cell phone random digit dialing methods.

³ For information about the Bureau's definition of financial well-being and how it was created, see *Financial well-being: The goal of financial education*, available at: consumerfinance.gov/data-research/research-reports/financial-well-being

⁴ For information about the CFPB Financial Well-Being Scale, see *CFPB Financial Well-Being Scale: Scale development technical report*, available at: consumerfinance.gov/data-research/research-reports/financial-well-being-technical-report/.

⁵ The research team responsible for developing and conducting the National Financial Well-Being Survey, subsequent data analyses and reports, and the Public Use File and related documentation included CFPB staff as well as a team of research contractors led by Abt Associates and including University of Wisconsin-Madison's Center for Financial Security and GfK. The research was funded under a competitive award; contract number TPDCFPBPA130014.

This document is the User's Guide for CFPB's National Financial Well-Being Survey Public Use File (PUF). The PUF is a dataset containing (1) data collected in the National Financial Well-Being Survey, (2) data about members of the KnowledgePanel collected prior to the survey, and (3) data on poverty levels in respondents' counties of residence. No personally identifying information is available. Section 2 of the User's Guide describes the National Financial Well-Being Survey. Section 3 explains how to use the PUF. Section 4 documents variables included in the PUF that were created for analysis or otherwise modified from the form in which they were originally collected during the survey. The full survey instrument is provided in Appendix A.

See the Public Use File Codebook for the list of variables found in the PUF; a data dictionary that presents the variable names, variable labels, value labels, who was asked each survey item, and unweighted frequencies and percentages of responses; and summary statistics for variables in the PUF.⁶

⁶ Available at: consumerfinance.gov/data-research/financial-well-being-survey-data/

2. Description of the survey

2.1 National Financial Well-Being Survey background and objectives

Enhancing financial well-being is the ultimate goal of financial capability policies, programs, and interventions. Yet, despite the critical importance of this outcome for assessing the effectiveness of financial literacy and education efforts, the fields of consumer finance and financial capability have operated without an accepted definition or measure of financial well-being. Therefore, the Consumer Financial Protection Bureau (CFPB) undertook research to develop a definition of financial well-being and the development and initial validation of a survey scale to measure individuals' well-being under this definition.⁷

In Phase 1 of the project, interviews were conducted with a broad cross-section of consumers and financial practitioners. Based on the narratives that emerged from these interviews, the CFPB and its research partners developed a definition of financial well-being that includes four

⁷ The research team responsible for conducting the research and analysis to develop this definition of financial well-being, as well as to develop the scale to measure financial well-being, included Bureau staff as well as a team of research contractors led by the Corporation for Enterprise Development (CFED), including the University of Wisconsin-Madison Center for Financial Security, the Urban Institute, ICF International, and Vector Psychometric Group. The research was funded under a competitive award: contract number CFP-12-Z-00019.

basic elements: (1) having control of day-to-day and month-to-month finances; (2) having capacity to absorb a financial shock; (3) being on track to meet financial goals; and (4) having the freedom to make choices that allow enjoyment of life.⁸ In Phase 2 of the project, new, validated instruments (scales) were developed for measuring financial well-being⁹ and a newly defined construct of financial skill.

Building on the initial definitional and scale development work, CFPB began Phase 3 of the project, in which the Bureau contracted with Abt Associates¹⁰ to collect and analyze data on the state and drivers of financial well-being from a sample of U.S. adults designed and weighted to represent the U.S. adult population and key subpopulations. The National Financial Well-Being Survey had three objectives, described below.

The first objective was to deliver the first-ever picture of the current state of financial well-being of American adults overall and important subpopulations, including older and younger adults, members of different racial and ethnic groups, and economically vulnerable individuals.

Accordingly, our report *Financial Well-Being in America*¹¹ presents descriptive information from the national survey on the level and distribution of financial well-being for the U.S. adult population and that of subgroups defined by: (1) individual characteristics; (2) household and family characteristics; (3) income and employment characteristics; (4) savings and safety nets; (5) financial experiences; and (6) financial behaviors, skills and attitudes.

⁸ For more information on these qualitative interviews and CFPB's development of the definition of financial well-being, see *Financial Well-Being: The goal of financial education*, available at: consumerfinance.gov/data-research/research-reports/financial-well-being

⁹ For more information on the development of the scale to measure financial well-being, see *CFPB Financial Well-Being Scale: Scale Development Technical Report*, available at: consumerfinance.gov/data-research/research-reports/financial-well-being-technical-report/

¹⁰ The research was funded under a competitive award; contract number TPDCFPBPA130014.

¹¹ Available at: consumerfinance.gov/data-research/research-reports/financial-well-being-america/

Second, this research provided data that allows for the testing of a specific set of hypotheses developed from the Phase 1 financial narratives and literature review. These hypotheses relate to the ways financial knowledge and financial behaviors contribute to financial well-being, controlling for factors such as social context and personal traits. A quantitative understanding of the most important drivers of financial well-being and how these drivers work together to produce financial well-being, supports the development or prioritization of efforts to improve financial outcomes. Analyses of subpopulations assist in identifying not only the groups in greatest need of support, but also the drivers that are most likely to have a meaningful impact on levels of financial well-being for members of each identified subpopulation.

Third, the public use file produced from this project is intended to support additional research on the factors and circumstances that drive financial well-being and enhance familiarity with the new financial well-being measure. We hope that the initial descriptive statistics presented in *Financial Well-Being in America*, along with the PUF, will engage a broad range of researchers and policymakers in identifying the factors that drive the financial well-being of consumers, and promoting successful programmatic and policy approaches for doing so.

2.2 Sampling, data collection, and weighting

This section describes data collection for the National Financial Well-Being Survey. The sample is described first, followed by the data collection protocol, weighting, design effects and margins of error, and the final dispositions and outcome rates for the survey.

Data provided in the PUF on select individual and household variables—including federal poverty level and certain demographic characteristics such as age, race/ethnicity, educational attainment, household size, and household income—come from GfK Knowledge Panel data about panelists which were collected prior to the survey as part of GfK’s standard business operations. Variables collected on behalf of the CFPB, through the National Financial Well-Being Survey, include data on financial well-being, financial skill, financial knowledge, financial behavior, financial attitudes, financial experiences, and other related factors. A complete list of variables and their sources can be found in Section 2 of the PUF Codebook.

2.2.1 Sample

Here, we first briefly describe the population the National Financial Well-Being Survey was designed to represent, before addressing the sample design used for the survey and the sample that was actually achieved.

Population

The survey was designed to represent the noninstitutionalized adult (age 18 and older) population in the 50 U.S. states and Washington, D.C.

Sample design

The survey sample was drawn from the GfK KnowledgePanel[®], a recruited sample designed to be nationally representative of U.S. households. There are two stages to sampling for surveys using the KnowledgePanel. The first stage is the sampling design for recruitment into the KnowledgePanel. The second stage is the sample drawn of KnowledgePanel members for a specific survey. We describe both of these below, followed by the sample actually achieved.

Selection of the Knowledge Panel

The GfK panel is the largest U.S. probability-based non-volunteer Internet panel, with a total of about 55,000 panel members. The large size of the panel provides a sufficient base to allow for stratification of the sample with oversampling of key populations (e.g., ages 62 and older). GfK's strictly probability-based recruitment procedures were based on random digit dialing (RDD) using a dual-frame landline and cell phone design through 2009 before switching to address-based sampling (ABS) in 2010 in order to maintain the representativeness of the panel.¹²

¹² ABS was used to supplement the RDD frame in 2009 before replacing the RDD frame entirely in 2010. ABS selects addresses from the U.S. Postal Service Delivery Sequence File. The RDD sampling scheme used a stratified design: one stratum had a higher concentration of African American and Hispanic households relative to national estimates

Recruitment is in both English and Spanish in order to ensure that different levels of language proficiency and acculturation are represented. GfK provides non-Internet households with a web-enabled computer and free Internet service so that they can participate as online panel members. See section 2.2.2 for recruitment and panel retention rates.

In addition to the standard GfK panel, the survey drew upon GfK's KnowledgePanel LatinoSM to ensure adequate representation of the Latino population. The Latino panel is designed to represent the U.S. Hispanic population, including English-dominant, bilingual, and Spanish-dominant. The weighted GfK panel matches the U.S. adult (age 18 and older) Hispanic population on gender, age, marital status, housing ownership, education, region, Internet access, household size, language proficiency, and place of birth.¹³ It differs to some degree with respect to income and employment status.

The ABS design for KnowledgePanel recruitment and the RDD design for KnowledgePanel Latino recruitment are described below.

Address-based sampling design

The ABS sampling frame GfK uses for address selection is enhanced by appending various ancillary data to each address to facilitate a stratified design. Quarterly samples are selected using a disproportionate stratified sampling methodology across the following four strata:

1. Hispanic households with at least one 18- to 24-year-old person;
2. Remaining Hispanic households;
3. Remaining households with at least one 18- to 24-year-old person;
4. All remaining households.

from the 2000 Census and the other had a lower concentration relative to Census estimates. Telephone numbers from the first stratum were selected at approximately twice the rate of those from the second stratum.

¹³ GfK. 2012. "Research Note: KnowledgePanel LatinoSM." GfK, Palo Alto, CA. Accessed November 11, 2015. <http://www.knowledgenetworks.com/resources/docs/KnowledgePanel-LatinoSM-Demographic-Profile-Feb-2012.pdf>

Adults from sampled households are invited to join KnowledgePanel through a series of mailings, including an initial invitation letter, a reminder postcard, and a subsequent follow-up letter. Given that a subset of physical addresses can be matched to a corresponding landline telephone number, about 5 weeks after the initial mailing, telephone refusal-conversion calls are made to nonresponding households for which a telephone number is matched. Invited households can join the panel by:

- Completing and mailing back a paper form in a postage-paid envelope,
- Calling a toll-free hotline phone number maintained by GfK, or
- Going to a designated GfK website and completing the recruitment form online.

During the initial recruitment survey, attempts are made to recruit every household member who is at least 13 years of age to become an active member. For the National Financial Well-Being Survey, only adults aged 18 and older were selected. Also, only one panelist per household was selected.

KnowledgePanel Latino recruitment

Members for KnowledgePanel Latino are recruited using a dual-frame RDD sampling methodology targeting telephone exchanges associated with Census blocks that have a 65 percent or greater Latino population density (this density level covers just over 50 percent of the U.S. Hispanic population). Moreover, cellular numbers from rate centers¹⁴ with high concentrations of Hispanics are also used to improve the representation of samples. Households

¹⁴ The rate center is a basic unit of cell phone geography that ties a cell phone number to a geographic location to which the phone center was assigned. For example, in the phone number 571-434-5400, 571-434 is an exchange assigned to Verizon Virginia, Inc. associated with the Herndon, VA rate center: Number Portability Administration Center. "How LNP Works." Number Portability Administration Center, Sterling, VA. Retrieved February 2, 2017. (<https://www.npac.com/number-portability/how-lnp-works>). Rate centers are the lowest level at which cell phone sample is allocated.

are screened in the Spanish language to only recruit those homes where Spanish is spoken at least half the time. This sample supplements the Latino households (English and Spanish) that are recruited through the KnowledgePanel’s general ABS recruitment sample.

Sample drawn from the KnowledgePanel

Having discussed the selection of the KnowledgePanel, attention will now turn to the selection of the sample for the National Financial Well-Being Survey from the KnowledgePanel. The sample design called for 5,000 completed surveys of adults in proportion to the U.S. population with respect to age, race/ethnicity, and household income below 200 percent of the federal poverty level (FPL), as well as an additional oversample of 1,000 completed surveys of adults age 62 and older. Targets of completed interviews were assigned proportional to their representation in the U.S. population; in the case of the age 62 and older oversample, the distribution of targets for ages 62-74 and 75 and older was in proportion to their representation in the U.S. population.

The sample targets were specified as follows, with the percentages from the Current Population Survey (CPS) 2016 Annual Socioeconomic Supplement shown in parentheses:

Main sample (target $n = 5,000$; U.S. residents age 18 and older in the 50 states and Washington, D.C.)

Race

African American alone non-Hispanic	$n \geq 590$	(CPS: 11.8%)
Hispanic, any race	$n \geq 790$	(CPS: 15.8%)

Poverty (FPL = federal poverty level)

Less than 200 percent of FPL	$n \geq 1,432$	(CPS: 28.6%)
------------------------------	----------------	--------------

Age

18-34	$n \approx 1,495 (\pm 50)$	(CPS: 29.3%)
35-54	$n \approx 1,690 (\pm 50)$	(CPS: 33.8%)
55-61	$n \approx 615 (\pm 50)$	(CPS: 12.3%)
62-74	$n \approx 800 (\pm 25)$	(CPS: 16.0%)
75 and older:	$n \approx 400 (\pm 25)$	(CPS: 8.0%)

Age 62 and older oversample (target $n = 1,000$; U.S. residents age 62 and older in the 50 states and Washington, D.C.)

Age

62-74	$n = 666$	(CPS: 66.6%)
75 and older	$n = 334$	(CPS: 33.4%)

These targets prioritized achieving minimum sample sizes on key analytic populations over minimizing design effects for the sample as a whole.¹⁵ The overall target sample size of $n = 5,000$ (excluding the age 62 and older oversample) was large enough that sampling error would be small even with design effects higher than typical for the KnowledgePanel.

As we describe below, it was necessary to add additional sample due to lower than expected response rates for key populations: below 200 percent of the federal poverty level, African American non-Hispanic, and Hispanic.

GfK was responsible for implementing a design to yield these targets. GfK based sample selection on age groups, with each age group selected to be representative of the U.S. population in the corresponding age range. Due to known cooperation rate differences for each age group and race/ethnicity, GfK oversampled for Hispanic and African American non-Hispanic panel members in anticipation of lower cooperation rates in these groups.

Overall, 14,402 panelists were selected: 11,513 initially for the general population sample, 1,647 for the age 62 and older oversample, and another 1,242 focused on adults below 200 percent of the Federal poverty level, African American non-Hispanic, and Hispanic.

¹⁵ A design effect (abbreviated as DEFF) is a measure of the statistical efficiency of a sample compared to a simple random sample of the same size. Specifically, it is the ratio of the sampling variance for a statistic computed under the sample design to the sample variance that would have been obtained from a simple random sample of the same size (Groves, Robert M., Floyd J. Fowler, Mick P. Couper, James M. Lepkowski, Eleanor Singer, and Roger Tourangeau. 2009. *Survey Methodology*. 2nd ed. Hoboken, NJ: Wiley, p. 109).

Sample achieved

From these 14,402 panelists, 6,394 surveys were completed: 5,395 from the general population sample (5,000 from the general population sample originally drawn and 395 from the additional sample added focusing on panel members below 200 percent of the federal poverty level or who were African American non-Hispanic or Hispanic) and 999 from the age 62 and older oversample. The counts of completed surveys exclude cases that were removed for reasons of response quality. Cleaning procedures are described after Table 4. The achieved sample differed slightly from these targets due to lower than expected completion rates for certain groups: below 200 percent of the federal poverty level, African American non-Hispanic, and Hispanic.

Table 1, Table 2, Table 3, and Table 4, below, show sample targets, achieved numbers of completed interviews, amounts of sample drawn, and completion rates for the general population sample.

TABLE 1: TARGET AND ACHIEVED COMPLETED INTERVIEWS AND SAMPLE DRAWN IN THE GENERAL POPULATION SAMPLE BY RACE

Group	Target	Total Achieved	Initial Sample Achieved	Additional Sample Achieved	Initial Sample Drawn	Additional Sample Drawn	Total Completion Rate
African American non-Hispanic	590	634	434	200	1,470	585	30.9%
Hispanic	790	820	625	195	2,296	657	26.8%
White non-Hispanic	N/A	3,666	3,666	0	7,136	0	51.4%
Other non-Hispanic	N/A	275	275	0	611	0	45.0%

TABLE 2: TARGET AND ACHIEVED COMPLETED INTERVIEWS AND SAMPLE DRAWN IN THE GENERAL POPULATION SAMPLE BY POVERTY LEVEL

Group	Target	Total Achieved	Initial Sample Achieved	Additional Sample Achieved	Initial Sample Drawn	Additional Sample Drawn	Total Completion Rate
Less than 200% of FPL	1,432	1,336	1,163	173	3,833	691	29.5%
200%+ of FPL	N/A	4,059	3,837	222	7,680	551	49.3%

TABLE 3: TARGET AND ACHIEVED COMPLETED INTERVIEWS AND SAMPLE DRAWN IN THE GENERAL POPULATION SAMPLE BY AGE GROUP

Age Group	Target	Total Achieved	Initial Sample Achieved	Additional Sample Achieved	Initial Sample Drawn	Additional Sample Drawn	Total Completion Rate
18-34	1,495	1,530	1,436	94	4,454	388	31.6%
35-54	1,690	1,903	1,754	149	4,056	491	41.9%
55-61	615	708	632	76	1,118	192	54.0%
62-74	800	851	780	71	1,241	160	60.7%
75 and older	400	403	398	5	644	11	61.5%

TABLE 4: TARGET AND ACHIEVED COMPLETED INTERVIEWS AND SAMPLE DRAWN IN THE AGE 62 AND OLDER OVERSAMPLE

Age Group	Target	Achieved	Drawn	Completion Rate
62-74	666	666	1,087	61.3%
75 and older	334	333	560	59.5%

The count of completed surveys excludes 72 panel members who completed the survey but were removed due to response quality concerns (70 from the original general population sample, 2 from the age 62 and older oversample, and 0 from the additional sample added later in the field period). Those who were removed were both “speeders” and “straight-liners.”

Speeding was defined based on a measure developed by TrueSample.¹⁶ Specifically, speeders were defined as those who speeded on 40 percent or more of web survey pages. For a given page, speeding was defined as follows. These steps determine whether a given respondent completed a page in a substantially shorter than average amount of time.

1. Length of time to complete each page (t) was transformed to its natural logarithm ($\ln(t)$).¹⁷
2. Outliers are defined for step 3. Outliers are defined as cases above and below 2.5 standard deviations of the mean (μ) of $\ln(t)$.
3. The standard deviation (σ) of page completion times is calculated for transformed page completion times, less the outliers defined in step 2.
4. Page completion times are standardized to Z -scores. The Z -score for the i th case (where $i = 1, 2, \dots, n$) is calculated as $Z_i = (\ln(t_i) - \mu) / \sigma$.
5. Page-level speeders are defined as $Z < -2$ (i.e., 2 standard deviations below the mean of $\ln(t)$).

Straight-lining occurs when a respondent provides the same answer to each row of a bank of questions presented in table form (also known as grid or matrix form). In the table, response

¹⁶ TrueSample. “Determining Respondent Engagement with TrueSample: Technical Documentation.” TrueSample, Seattle, WA.

¹⁷ The reason for this step is to more closely approximate a normal distribution of length of time to complete each page. Page completion times have a “long tail” on the right hand side of the distribution: a small number of cases with very long completion times.

options are shown in columns across the horizontal axis and individual questions are shown in rows down the vertical axis and the respondent chooses one response per row. Where a respondent chooses the same option on each row, the responses appear in a straight line running down the table. The various table items in the survey were carefully reviewed to determine a straight-lining measure.¹⁸ After review, the criterion chosen was straight-lining both FWB1 and FWB2 as indicative of poor response quality.¹⁹ We selected FWB1 and FWB2 because they have reasonable numbers of rows (6 for FWB1 and 4 for FWB2) and reverse-worded items, and there was substantial diversity of responses to each item.

A total of 106 respondents were incorrectly terminated at COVERCOSTS. These respondents were sent a custom reminder asking them to log back in. All but 16 of these respondents went on to complete the survey.

Systematic item nonresponse

Due to an error, a question about monthly housing expenditures, HOUSERANGES, was only asked of respondents who reported they owned their own home. It was intended that HOUSERANGES be asked of all respondents. In total, 2,230 respondents were not asked this question. After this error was identified, a follow-up survey was fielded to collect data from those not asked the question. The survey consisted solely of HOUSERANGES and did not

¹⁸ The decision to remove a respondent for straight-lining carries the risk of false positives: cases that did straight-line, but for reasons other than poor response quality. In other words, a respondent may have carefully considered each response and legitimately selected the same response option for each question. The risk of false positives is strongly associated with the number of questions and the degree of correlation between responses. In a table item with few rows, the probability of nonmaliciously straight-lining is high. In a table with many rows, the probability of straight-lining nonmaliciously is low. The degree of uniformity of responses to each row of the grid is also very important. Where responses are very uniform across rows, there is a high probability of nonmalicious straight-lining. Where responses vary widely across rows, there is a low probability of nonmalicious straight-lining.

¹⁹ FWB1 was a table item with 6 rows and FWB2 was a table item with 4 rows. The items focused on financial well-being. See Appendix A for question wording for these items.

include any other items. In total, 1,823 of the 2,230 panel members responded to the follow-up survey for a completion rate of 81.7 percent. The item response rate for HOUSERANGES was 93.6 percent across all completed interviews, including those correctly asked HOUSERANGES in the original survey.²⁰ Values for HOUSERANGES were imputed for the remaining 407 nonrespondents to the follow-up survey. The imputation process is described in section 4.1.

Data cleaning

Outlier values (values outside of the anticipated or typical range for the question) were identified for a small number of respondents for items KIDS_1 (number of children aged below 7 respondent financially supports), KIDS_2 (number of children age 7 to 12 respondent financially supports), and KIDS_3 (number of children age 13 to 17 respondent financially supports). In total, 1 edit was made to KIDS_1, 3 edits to KIDS_2, and 6 edits to KIDS_3.

In addition, two edits were made to SOCSEC2 (age began receiving Social Security retirement benefits for individuals aged 62 and older who indicated in SOCSEC1 that they received Social Security retirement benefits). In each case, the respondent reported receiving benefits at an age more than 1 year above their current age.

We document the cleaning process in greater detail in section 4.2.

2.2.2 Final dispositions and outcome rates

Because the National Financial Well-Being Survey is drawn from an online panel, outcome rates are reported in a different fashion to other surveys because there are several levels of response

²⁰ The item response rate is defined as “the ratio of the number of respondents for whom an in-scope response was obtained (I^x for item x) to the number of respondents who were asked to answer that item.” Office of Management and Budget. 2006. “Standards and Guidelines for Statistical Surveys.” Office of Management and Budget, Washington, D.C., p. 16. If we treat the nonsubstantive responses of refused, don’t know, and prefer not to say as being out-of-scope, the item response rate is 79.2 percent for the whole sample.

that must be taken into account: recruitment into the panel and response to this specific survey. The response rate for panel recruitment is measured by the recruitment rate (RECR), 12.9 percent for this particular survey. Because panelists must complete a profile before becoming members of the KnowledgePanel, the percentage of successfully recruited panelists who complete the profile is measured by the profile rate (PROR), 62.6 percent for this particular survey. Finally, the completion rate (COMR) is the number of panelists invited to take National Financial Well-Being Survey who completed the survey, 44.4 percent in this case (6,394 out of 14,402 panelists invited to participate).²¹ The final cumulative response rate (CUMRR = RECR × PROR × COMR) is 3.6 percent. These outcome rates follow reporting guidelines from the American Association for Public Opinion Research (AAPOR) and Callegaro and DiSogra.²²

As described above, a completion rate of 81.7 percent was achieved for the follow-up survey of respondents who did not answer HOUSERANGES.

2.2.3 Data collection protocol

The study's data collection protocol is described below, beginning with the pretest and then describing the main survey.

²¹ Respondents who were cleaned from the file based on speeding and straight-lining are not included in the count of completed surveys used here, neither are those with partial responses, break-offs, nor panel members who did not consent.

²² American Association for Public Opinion Research. 2016. "Standard Definitions: Final Dispositions of Case Codes and Outcome Rates for Surveys." 9th ed. American Association for Public Opinion Research, Oakbrook Terrace, IL. Callegaro, Mario and Charles DiSogra. 2008. "Computing Response Metrics for Online Panels." *Public Opinion Quarterly* 72(5):1008-32. The AAPOR CUMRR is Callegaro and DiSogra's (2008) CUMRR1. Callegaro and DiSogra's CUMRR2 adds retention rate (RETR): $CUMRR = RECR \times PROR \times RETR \times COMR$. The retention rate for this particular survey was 29.7 percent.

Once assigned to a survey, members receive a notification email letting them know there is a new survey available for them to take. This email notification contains a link that sends them to the survey questionnaire. No login name or password is required.

GfK also operates an ongoing modest incentive program, primarily through the use of a points system, to encourage participation and create member loyalty. Members can redeem their points for cash, merchandise, gift cards, or game entries. Incentives that are not specific to a given survey are used to maintain a high degree of panel loyalty and to prevent attrition from the panel. For households without web-enabled devices or an Internet connection, the panel loyalty incentive is the web-enabled devices and the Internet connections that GfK provides for free to households without these items. Panelists who use their own personal computers and Internet service for survey participation are enrolled in a points program that is analogous to a frequent flyer program; respondents are credited with points in proportion to their regular participation in surveys.

Panelists can redeem their points at times of their own choosing, with average accumulations of \$4 to \$6 per month. Generally, panel members are invited to complete one survey per week, below the point at which survey-specific incentives are offered. On average, panel members complete two to three surveys per month with typical durations of 10 to 15 minutes per survey. Surveys of this length generally do not receive survey-specific incentives.²³ The average value of incentives for surveys of the length of the National Financial Well-Being Survey (15 to 25 minutes) is in the \$1 range, too low to affect financial well-being.²⁴

²³ Surveys that involve unusual requests, such as specimen collection, the viewing of a specific television program, or completion of a daily diary may have a survey-specific incentive even if they take less than 15 minutes to complete.

²⁴ For surveys longer than 25 minutes, respondents receive additional points. Respondents who participate in the survey are credited with 5,000 to 10,000 points, which roughly equates to \$5 to \$10 depending on the type of award selected by respondents when redeeming and the length of the survey.

Pretest

A pretest of the survey took place between October 4 and 7, 2016. A total of 126 surveys were completed. The sample of the pretest was designed to be diverse rather than representative to ensure that all groups of interest to CFPB were included in the pretest. Cases from the pretest are not included in the final dataset. A number of changes were made to the instrument after the pretest.

Main study

The main study was fielded between October 27 and December 5, 2016.

Once assigned to a survey, members received a notification email letting them know there is a new survey available for them to take. This email notification contained a link that sent them to the survey questionnaire. No login name or password was required. After three days, automatic email reminders were sent to all nonresponding panel members in the sample. To assist panel members with their survey taking, each individual has a personalized “home page” that lists all the surveys that were assigned to that member and have yet to be completed.

The invitation email to the general population sample and age 62 and older oversample was sent on October 27, 2016. Additional reminders were sent 6, 9, 13, 17, and 20 days after launch. The additional sample was released on November 23, 2016.

Median survey length was 26 minutes.

Follow-up survey for HOUSERANGES

A follow-up survey of respondents who were not asked HOUSERANGES due to an error (see p. 16) was fielded between February 21, 2017 and March 6, 2017. The survey consisted solely of HOUSERANGES and did not include any other items.

2.2.4 Weighting

Weighting took place in three steps, which are described below.

First, GfK weighted the pool of active panel members to the geodemographic benchmarks secured from the 2016 March supplement of the Current Population Survey (CPS) along several dimensions. Using the resulting weights as measure of size, in the next step a probability

proportional to size procedure was used to select study specific samples. Departures from an equal probability of selection methodology design in the selection of the sample were accounted for by adjusting the design weights in reference to the CPS benchmarks for the population of interest. The geodemographic benchmarks used to weight the active panel members for computation of size measures included:

- Sex (female, male);
- Age (18–29, 30–44, 45–59, and 60 and older);
- Race/Hispanic ethnicity (white non-Hispanic, African American non-Hispanic, other non-Hispanic, 2 or more races non-Hispanic, Hispanic);
- Education (less than high school, high school, some college, bachelor and beyond);
- Census Region (Northeast, Midwest, South, West);
- Household income (under \$10,000, \$10,000 to \$24,999, \$25,000 to \$49,999, \$50,000 to \$74,999, \$75,000 to \$99,999, \$100,000 to \$149,999, \$150,000 and above);
- Home ownership status (own, rent/other);
- Metropolitan Area (yes, no).

Second, GfK raked the sample of completed interviews to benchmarks drawn from the March 2016 CPS.²⁵ These benchmarks consisted of:

- Age;
- Race/ethnicity;
- Sex;
- Education;
- Household income;
- Poverty (less than 100% of federal poverty level, 100% to 199%, 200% or more of federal poverty level).

Third, a final set of raking adjustments were made by Abt Associates. These adjustments were designed to ensure that the weighted sample was representative with respect to key demographics:

- Age (18-39, 40-61, 62-74, 75 and older);
- Sex (female, male);
- Education (less than high school, high school graduate/GED, some college, bachelor's degree, post-baccalaureate degree);

²⁵ Raking—also known as iterative proportional fitting—is a technique for ensuring that the weighted marginal frequencies of data match those of a reference source (Deming, W. Edwards and Fredrick F. Stephan. 1940. “On a Least Squares Adjustment of a Sampled Frequency Table when the Expected Marginal Totals are Known.” *Annals of Mathematical Statistics* 11(4):427-44; Deming, W. Edwards. 1943. *Statistical Adjustment of Data*. New York: Wiley). The raking algorithm adjusts weights to ensure that the weighted frequency distribution of the sample matches that of the reference source on one dimension. The algorithm loops sequentially through each dimension in an iterative fashion, terminating when the weighted sample distribution is within a prespecified degree of tolerance of the reference distributions on all dimensions.

- Urbanicity operationalized as residence in a Metropolitan Statistical Area (MSA) or not (resident in MSA, not resident in MSA);
- Race/ethnicity (white non-Hispanic, African American non-Hispanic, Hispanic, other non-Hispanic);
- Poverty (less than 100% of federal poverty level, 100%-199% of federal poverty level, 200% or more of Federal Poverty level).

Specifically, the marginal frequencies to be raked to were:

- Region × Age;
- Sex × Education × Age;
- MSA × Age;
- Race/ethnicity × Age;
- Poverty status × Age.

Age was included in each set of marginals to ensure that the sample would be representative regardless of whether analysis was confined to adults aged 62 and older or for the entire sample.

The target values for these were drawn from the Current Population Survey Annual Socioeconomic Supplement public use file. The final round of raking converged with variation of less than 0.1 percentage points from targets.

Unweighted frequencies of completed surveys, target values, and weighted frequencies are shown for each raking dimension in the tables below.

TABLE 5: UNWEIGHTED VALUES FOR REGION X AGE

Age Group	Northeast	Midwest	South	West	Total
18-39	5.0%	7.1%	10.8%	7.6%	30.4%
40-61	6.9%	7.9%	11.7%	7.9%	34.4%
62-74	4.2%	5.5%	8.6%	5.5%	23.7%
75 and older	2.2%	2.4%	4.2%	2.8%	11.5%
Total	18.2%	22.8%	35.3%	23.7%	100.0%

TABLE 6: TARGET VALUES FOR REGION X AGE

Age Group	Northeast	Midwest	South	West	Total
18-39	6.7%	7.8%	14.3%	9.5%	38.4%
40-61	6.7%	7.9%	14.3%	8.8%	37.7%
62-74	3.0%	3.4%	6.0%	3.5%	15.9%
75 and older	1.5%	1.8%	2.9%	1.8%	8.0%
Total	17.9%	21.0%	37.5%	23.6%	100.0%

TABLE 7: WEIGHTED VALUES FOR REGION X AGE

Age Group	Northeast	Midwest	South	West	Total
18-39	6.7%	7.8%	14.3%	9.5%	38.4%
40-61	6.7%	7.9%	14.3%	8.8%	37.7%
62-74	3.0%	3.4%	5.9%	3.5%	15.9%
75 and older	1.5%	1.8%	2.9%	1.8%	8.0%
Total	17.9%	21.0%	37.4%	23.6%	100.0%

TABLE 8: UNWEIGHTED VALUES FOR SEX X EDUCATION X AGE

Sex/Age	Less than high school	High school graduate	Some college	Bachelor's degree	Graduate degree	Total
Male, 18-39	0.8%	2.9%	5.6%	4.2%	2.5%	16.0%
Male, 40-61	1.1%	4.3%	5.6%	4.7%	3.5%	19.3%
Male, 62-74	0.4%	2.2%	3.3%	2.4%	2.7%	11.0%
Male, 75 and older	0.3%	1.6%	1.4%	1.1%	1.6%	6.1%
Female, 18-39	1.3%	3.7%	4.4%	2.7%	2.2%	14.3%
Female, 40-61	1.5%	4.3%	4.4%	2.7%	2.2%	15.1%
Female, 62-74	0.8%	4.0%	4.0%	2.0%	1.9%	12.7%
Female, 75 and older	0.4%	2.3%	1.4%	0.8%	0.5%	5.5%

TABLE 9: TARGET VALUES FOR SEX X EDUCATION X AGE

Sex/Age	Less than high school	High school graduate	Some college	Bachelor's degree	Graduate degree	Total
Male, 18-39	2.5%	5.8%	5.9%	3.5%	1.5%	19.2%
Male, 40-61	2.1%	5.6%	4.7%	3.6%	2.3%	18.4%
Male, 62-74	0.8%	2.1%	1.9%	1.5%	1.1%	7.4%
Male, 75 and older	0.6%	1.1%	0.7%	0.5%	0.5%	3.4%
Female, 18-39	2.0%	4.5%	6.6%	4.1%	2.0%	19.2%
Female, 40-61	1.8%	5.2%	5.6%	4.2%	2.5%	19.3%
Female, 62-74	0.9%	2.8%	2.3%	1.4%	1.0%	8.5%
Female, 75 and older	0.9%	1.9%	1.0%	0.5%	0.3%	4.6%

TABLE 10: WEIGHTED VALUES FOR SEX X EDUCATION X AGE

Sex/Age	Less than high school	High school graduate	Some college	Bachelor's degree	Graduate degree	Total
Male, 18-39	2.5%	5.8%	5.9%	3.5%	1.5%	19.2%
Male, 40-61	2.1%	5.6%	4.7%	3.6%	2.3%	18.4%
Male, 62-74	0.8%	2.1%	1.9%	1.5%	1.1%	7.4%
Male, 75 and older	0.6%	1.1%	0.7%	0.5%	0.5%	3.4%
Female, 18-39	2.0%	4.5%	6.5%	4.1%	2.0%	19.2%
Female, 40-61	1.9%	5.2%	5.6%	4.2%	2.5%	19.3%
Female, 62-74	0.9%	2.8%	2.3%	1.4%	1.0%	8.5%
Female, 75 and older	0.9%	1.9%	1.0%	0.5%	0.3%	4.6%

TABLE 11: UNWEIGHTED VALUES FOR MSA X AGE

Age Group	Non-MSA	MSA	Total
18-39	3.6%	26.8%	30.4%
40-61	4.3%	30.1%	34.4%
62-74	3.7%	20.0%	23.7%
75 and older	1.8%	9.7%	11.5%
Total	13.4%	86.6%	100.0%

TABLE 12: TARGET VALUES FOR MSA X AGE

Age Group	Non-MSA	MSA	Total
18-39	4.4%	33.9%	38.4%
40-61	5.1%	32.6%	37.7%
62-74	2.6%	13.3%	15.9%
75 and older	1.3%	6.7%	8.0%
Total	13.5%	86.5%	100.0%

TABLE 13: WEIGHTED VALUES FOR MSA X AGE

Age Group	Non-MSA	MSA	Total
18-39	4.4%	33.9%	38.4%
40-61	5.1%	32.6%	37.7%
62-74	2.6%	13.3%	15.9%
75 and older	1.3%	6.7%	8.0%
Total	13.5%	86.5%	100.0%

TABLE 14: UNWEIGHTED VALUES FOR RACE/ETHNICITY X AGE

Age Group	African American		Other non-Hispanic		Total
	White non-Hispanic	non-Hispanic	Hispanic	Hispanic	
18-39	18.8%	3.5%	6.1%	1.9%	30.4%
40-61	23.4%	4.1%	5.2%	1.6%	34.4%
62-74	18.3%	2.3%	1.7%	1.3%	23.7%
75 and older	9.8%	0.7%	0.7%	0.4%	11.5%
Total	70.3%	10.7%	13.7%	5.3%	100.0%

TABLE 15: TARGET VALUES FOR RACE/ETHNICITY X AGE

Age Group	White non-Hispanic	African American non-Hispanic	Hispanic	Other non-Hispanic	Total
18-39	21.5%	5.1%	8.1%	3.7%	38.4%
40-61	24.6%	4.5%	5.7%	2.9%	37.7%
62-74	12.0%	1.6%	1.4%	1.0%	15.9%
75 and older	6.3%	0.7%	0.6%	0.4%	8.0%
Total	64.4%	11.8%	15.7%	8.0%	100.0%

TABLE 16: WEIGHTED VALUES FOR RACE/ETHNICITY X AGE

Age Group	White non-Hispanic	African American non-Hispanic	Hispanic	Other non-Hispanic	Total
18-39	21.5%	5.1%	8.1%	3.7%	38.4%
40-61	24.6%	4.5%	5.7%	2.9%	37.7%
62-74	11.9%	1.6%	1.4%	1.0%	15.9%
75 and older	6.3%	0.7%	0.6%	0.4%	8.0%
Total	64.3%	11.9%	15.8%	8.0%	100.0%

TABLE 17: UNWEIGHTED VALUES FOR POVERTY X AGE

Age Group	Less than 100% of federal poverty level	100% to 199% of federal poverty level	200% or more of federal poverty level	Total
18-39	5.3%	4.9%	20.2%	30.4%
40-61	3.3%	4.1%	27.0%	34.4%
62-74	1.2%	2.8%	19.7%	23.7%
75 and older	0.5%	1.7%	9.3%	11.5%
Total	10.3%	13.4%	76.2%	100.0%

TABLE 18: TARGET VALUES FOR POVERTY X AGE

Age Group	Less than 100% of federal poverty level	100% to 199% of federal poverty level	200% or more of federal poverty level	Total
18-39	5.7%	6.9%	25.9%	38.4%
40-61	3.8%	5.1%	28.8%	37.7%

Age Group	Less than 100% of federal poverty level	100% to 199% of federal poverty level	200% or more of federal poverty level	Total
62-74	1.4%	2.8%	11.7%	15.9%
75 and older	0.8%	2.2%	5.0%	8.0%
Total	11.7%	17.0%	71.3%	100.0%

TABLE 19: WEIGHTED VALUES FOR POVERTY X AGE

Age Group	Less than 100% of federal poverty level	100% to 199% of federal poverty level	200% or more of federal poverty level	Total
18-39	5.7%	6.8%	25.9%	38.4%
40-61	3.8%	5.1%	28.8%	37.7%
62-74	1.4%	2.8%	11.7%	15.9%
75 and older	0.8%	2.2%	5.0%	8.0%
Total	11.7%	17.0%	71.4%	100.0%

Although the weighting procedure described above ensures that the weighted sample matches the U.S. population on the characteristics shown above, coverage error or nonresponse error can result in differences between the sample and the U.S. population that are not corrected for by using weights.

2.2.5 Design effect and margins of error

The overall sample size achieved was 6,394. Weighting and survey design features that depart from simple random sampling tend to result in an increase in the variance of survey estimates. This increase, known as the design effect (DEFF), should be incorporated into the margin of error, standard errors, and tests of statistical significance. See Section 3.3 for descriptions of how to use weights in various statistical software packages in a way that allows for the correct calculations of margins of error, standard errors, and tests of statistical significance.

The overall design effect for a survey is commonly approximated as $1 + CV^2$, where CV is the coefficient of variation of the weights. For this survey, this apparent design effect is 1.343. The margin of error (half-width of the 95 percent confidence interval) incorporating the design effect for full-sample cross-sectional estimates at 50 percent is ± 1.4 percentage points.²⁶ For the age 62 and older population, sample size achieved was 2,253, DEFF was 1.294, and the margin of error was 2.3 percent. Sample sizes, design effects, and margins of error for these and other populations of interest are shown in Table 20.

²⁶ It is important to remember that random sampling error is only one possible source of the total error in a survey estimate. Other sources, such as question wording and reporting inaccuracy, may contribute additional nonsampling error.

TABLE 20: SAMPLE SIZE, DESIGN EFFECT, AND MARGIN OF ERROR FOR SELECTED POPULATIONS

Sample	Nominal Sample Size	DEFF	Effective Sample Size	Margin of Error
All	6,394	1.343	4,762	1.4%
Age 62 and older	2,253	1.294	1,741	2.3%
Age 18-34	1,530	1.367	1,119	2.9%
Age 35-44	828	1.118	697	3.7%
Age 45-54	1,075	1.152	933	3.2%
Age 55-61	708	1.187	596	4.0%
Age 62-74	1,517	1.186	1,279	2.7%
Age 75 and older	736	1.500	491	4.4%
White non-Hispanic	4,498	1.278	3,519	1.7%
African American non-Hispanic	685	1.325	517	4.3%
Hispanic	875	1.352	647	3.9%
Less than 200% of federal poverty level	1,520	1.253	1,213	2.8%
Less than 100% of federal poverty level	661	1.302	508	4.3%
100%-199% of federal poverty level	859	1.217	706	3.7%
200% or more of federal poverty level	4,874	1.363	3,576	1.6%

2.3 Survey overview

In selecting measures for the survey besides the core measure of financial well-being, the CFPB considered content areas of importance to the field of consumer finance and sought measures of individual, household, and situational factors that our prior research suggested were likely related to an individual's level of financial well-being.²⁷ These include measures of individual and household characteristics, income and employment, savings and safety nets, financial experiences (including financial socialization), and behaviors, skills and attitudes that have been hypothesized to influence adults' levels of financial well-being. In consultation with a panel of experts in the multiple domains of interest,²⁸ an initial inventory of possible measures was constructed and then narrowed to fit roughly 20 minutes of survey time.

Wherever possible, items and scales were selected that had been used in other surveys or validated through prior research, though many items and scales needed to be significantly adapted to fit the specific constructs of interest and time constraints of this survey. For scales, trimming was accomplished using a variety of methods. First, where available, original scale development articles were reviewed to identify item loading. Second, where data were available from previous survey waves, analyses were conducted to identify the subset of items that would correlate most highly with scores from the full scale. Finally, items were reviewed for redundancy, keeping the most distinct items from each scale. These methods were used as general tools to reduce the number of items.

Four scales are included in the survey in their entirety, with scores provided in the PUF:

²⁷ See *Financial well-being: The goal of financial education*, available at: consumerfinance.gov/data-research/research-reports/financial-well-being/, for an overview of these factors.

²⁸ We are grateful for the insights provided by members of the survey development expert panel: Adele Atkinson (OECD); Martha J. Deevy (Stanford Center on Longevity); Keith S. Ernst (FDIC); Carol Graham (Brookings Institution); Joanne W. Hsu (Federal Reserve Board) Clinton Key (Pew Charitable Trusts); Helen Levy (University of Michigan); Gary Mottola (FINRA Foundation); and Ellen Peters (Ohio State University).

- The 10-item CFPB Financial Well-Being Scale;²⁹
- The 10-item CFPB Financial Skill Scale;³⁰
- The 10-item Knoll and Houts Financial Knowledge Scale;³¹ and
- The 3-item Lusardi and Mitchell Financial Knowledge Scale.³²

The complete survey instrument is attached as Appendix A, and a full list of variables included in the PUF is available in the PUF Codebook.

²⁹ See *CFPB Financial Well-Being Scale: Scale development technical report*, available at: consumerfinance.gov/data-research/research-reports/financial-well-being-technical-report/

³⁰ Unpublished scale developed in prior phase of CFPB financial well-being research project.

³¹ This 10 item scale is an unpublished short version of Knoll, Melissa A. and Carrie R. Houts. 2012. "The Financial Knowledge Scale: An Application of Item Response Theory to the Assessment of Financial Literacy." *Journal of Consumer Affairs* 46(3): 381-410.

³² Lusardi, Annamaria and Olivia S. Mitchell. 2008. "Planning and Financial Literacy: How Do Women Fare?" NBER Working Paper No. 13750. National Bureau of Economic Research, Cambridge, MA.

3. How to use the Public Use File

The following sections describe how to use the Public Use File (PUF) for CFPB's National Financial Well-Being Survey. Section 3.1 briefly describes the PUF. Section 3.2 describes the dataset and the codes that can be used to import and label the dataset into various software packages. Section 3.3 describes the use of weights in various datasets. Section 3.4 details the meanings of missing or non-substantive value codes.

3.1 Description of the Public Use File

To minimize disclosure risk, no direct identifiers are in the PUF, and all indirect identifiers have been either removed or masked through collapsing smaller categories into broader ones (e.g., years of age into age bands) in creating the PUF.

The PUF consists of a single comma-separated values (CSV) dataset and accompanying code to import and label data for the study in a variety of software packages. These are described next.

3.2 Dataset and code

3.2.1 Dataset

The PUF is provided in a comma-separated values CSV dataset. The file name is `NFWBS_PUF_2016_data.csv` and contains 217 variables and 6,394 records. The 6,394 records

include only those who completed the survey and do not include panel members selected for the survey who did not complete it nor the 72 respondents who were removed during cleaning.

3.2.2 Code

The CSV dataset can be read in to SAS, SPSS, Stata, R or Python statistical software using the appropriate sample program, as shown below.

Software package	File name
SAS	NFWBS_PUF_2016_read_in_SAS.sas
SPSS	NFWBS_PUF_2016_read_in_SPSS.sps
Stata	NFWBS_PUF_2016_read_in_Stata.do
R	NFWBS_PUF_2016_read_in_R.r
Python 2.7	NFWBS_PUF_2016_read_in_PY27.py

In addition to reading in the datasets, these programs also fully label the variables and their values (with the exception of the programs for Python and R, which instead provide templates for remapping the values from the dummy variables to the values provided in the codebook).

Please note that each program requires the user to enter the directory in which the CSV file has been saved. Upon opening each sample program, the user will find clearly indicated guidance to copy and paste their specific file path into the program. Unless otherwise specified, the SAS, SPSS, or Stata format dataset will be saved to the same folder as the CSV file.

Users of other statistical packages may adapt the code from these programs for their preferred software.

3.3 Using the study weights

Analyses should always be run with study weights (see description of weights in section 2.2.4). Using the study weights is essential to ensure that results are correctly reported. Weights are extremely important for several reasons. First, the sample was not selected with equal probability. Most obviously, the study included an oversample of adults aged 62 and older. However, GfK also adjusted probabilities of selection for different age and race/ethnicity groups based on expected response rates. The weights, among other things, account for variations in the probability of selection (see the discussion of the sample design in section 2.2.1). Second, the

weights also adjust the final sample to match the U.S. population with respect to age, sex, race/ethnicity, poverty, and education (see discussion of weighting in section 2.2.4), which, among other things, adjusts for differential nonresponse across different socioeconomic groups. Unweighted analyses will not account for differences in the probability of selection of different groups nor for the adjustments that ensure that the achieved sample matches the U.S. population on the various parameters described.

A single weight is included in the file for analyses of the sample as a whole, inclusive of the age 62 and older subsample. This weight is called “finalwt.”

3.3.1 SAS

The SAS code provided does not enable weights by default. A WEIGHT statement should be included within the appropriate procedure. All analyses using weights should use survey procedures. A simple example is given below using PROC SURVEYFREQ for variable FWB1_1.

```
PROC SURVEYFREQ DATA=datasetname ;  
TABLES FWB1_1 ;  
WEIGHT finalwt ;  
RUN ;
```

When conducting analyses on subpopulations, the variable or variables defining the subsample should be interacted with the variable of interest using the * operator rather than using a where statement. The use of where will result in SAS incorrectly assuming that sample size is only that of the subpopulation, not the entire sample.

Correctly specified to yield correct standard errors from the subsample of interest:³³

```
PROC SURVEYFREQ DATA=datasetname ;  
TABLES agecat*fwb1_1 ;
```

³³ If the user wishes to analyze only where agecat ≥ 6 (i.e., ages 62 and older), a new variable would need to be created in the DATA step: e.g., setting $1 \leq \text{agecat} \leq 5$ equal to 0 and $\text{agecat} \geq 6$ equal to 1.

WEIGHT finalwt;
 RUN;

Table of agecat by FWB1_1

agecat	FWB1_1	Frequency	Weighted Frequency	Std Err of Wgt Freq	Percent	Std Err of Percent
18-24	Response not written to database	0
	Refused	2	10.16055	7.51454	0.1589	0.1174
	Not at all	97	139.62358	16.40497	2.1837	0.2543
	Very little	85	121.84909	14.99073	1.9057	0.2328
	Somewhat	147	240.95595	23.36445	3.7685	0.3585
	Very well	57	84.81522	12.37097	1.3265	0.1926
	Completely	26	27.72902	6.00378	0.4337	0.0939
	Total	414	625.13341	35.21199	9.7769	0.5254
25-34	Response not written to database	0
	Refused	3	3.71877	2.47993	0.0582	0.0388
	Not at all	232	300.29442	22.24622	4.6965	0.3427
	Very little	188	217.73589	17.88094	3.4053	0.2775
	Somewhat	376	456.41887	26.66119	7.1382	0.4081
	Very well	197	230.81332	18.97371	3.6098	0.2938
	Completely	120	148.37026	15.94038	2.3205	0.2474
	Total	1116	1357	43.85639	21.2285	0.6371
35-44	Response not written to database	0
	Refused	1	1.21035	1.21035	0.0189	0.0189
	Not at all	119	137.16736	13.48208	2.1453	0.2103
	Very little	151	165.28104	14.34699	2.5849	0.2239
	Somewhat	283	313.76206	19.97331	4.9071	0.3106
	Very well	180	185.98422	15.04617	2.9087	0.2349
	Completely	94	100.51813	11.31026	1.5721	0.1766
	Total	828	903.92315	32.31266	14.1371	0.4976
45-54	Response not written to database	0
	Refused	1	0.44724	0.44724	0.0070	0.0070
	Not at all	187	218.71971	16.68871	3.4207	0.2600
	Very little	194	225.79258	17.14781	3.5313	0.2670
	Somewhat	358	395.40904	21.76688	6.1841	0.3389
	Very well	224	255.07622	18.35031	3.9893	0.2854
	Completely	111	119.60798	12.08437	1.8706	0.1888
	Total	1075	1215	36.75214	19.0030	0.5619
55-61	Response not written to database	1	1.25396	1.25396	0.0196	0.0196
	Refused	1	1.38136	1.38136	0.0216	0.0216
	Not at all	96	110.37784	12.29659	1.7263	0.1918
	Very little	113	132.20628	13.47522	2.0677	0.2100
	Somewhat	221	241.89413	17.70003	3.7831	0.2755
	Very well	183	186.77471	14.65334	2.9211	0.2291
	Completely	93	89.34985	9.58345	1.3974	0.1501
	Total	708	763.23813	29.75736	11.9368	0.4602
62-69	Response not written to database	0
	Refused	1	0.96266	0.96266	0.0151	0.0151
	Not at all	101	85.46766	9.05883	1.3367	0.1421
	Very little	136	99.81709	9.26589	1.5611	0.1457
	Somewhat	325	210.62731	12.09494	3.2941	0.1931
	Very well	292	178.00864	10.83736	2.7840	0.1729
	Completely	166	106.75680	8.81122	1.6696	0.1392
	Total	1021	681.64017	21.30542	10.6606	0.3539
70-74	Response not written to database	0
	Refused	0
	Not at all	38	33.72247	5.98527	0.5274	0.0937
	Very little	50	41.37920	7.34262	0.6472	0.1148

	Somewhat	172	120.29198	9.82537	1.8813	0.1549
	Very well	152	92.30855	7.87975	1.4437	0.1245
	Completely	84	48.33006	5.53512	0.7559	0.0871
	Total	496	336.03227	16.30178	5.2554	0.2605

75+	Response not written to database	0
	Refused	1	0.56732	0.56732	0.0089	0.0089
	Not at all	57	71.18590	10.98999	1.1133	0.1713
	Very little	72	63.14567	8.86507	0.9876	0.1386
	Somewhat	221	165.34546	13.19859	2.5859	0.2069
	Very well	246	142.62841	10.35858	2.2307	0.1639
	Completely	139	68.75582	6.52948	1.0753	0.1031
	Total	736	511.62857	22.19228	8.0017	0.3508

Total	Response not written to database	1	1.25396	1.25396	0.0196	0.0196
	Refused	10	18.44826	8.21011	0.2885	0.1282
	Not at all	927	1097	38.17959	17.1498	0.5712
	Very little	989	1067	35.96238	16.6908	0.5461
	Somewhat	2103	2145	47.37532	33.5425	0.6891
	Very well	1531	1356	36.46789	21.2138	0.5677
	Completely	833	709.41791	27.19637	11.0951	0.4262
	Total	6394	6394	46.81056	100.000	

Incorrectly specified subsample:

```
PROC SURVEYFREQ DATA=datasetname;
WHERE agecat>=6;
TABLES fwbl_1;
WEIGHT finalwt;
RUN;
```

FWB1_1	Frequency	Frequency	Wgt Freq	Percent	Percent
Refused	2	1.52998	1.11718	0.1000	0.0730
Not at all	196	190.37603	14.99181	12.4486	0.9287
Very little	258	204.34196	14.22436	13.3618	0.8914
Somewhat	718	496.26475	17.92404	32.4504	1.1186
Very well	690	412.94560	14.85512	27.0022	1.0009
Completely	389	223.84268	11.47134	14.6369	0.7670
Total	2253	1529	17.47320	100.000	

3.3.2 SPSS

The SPSS code provided does not enable weights by default. To enable weights, the following command should be given:

```
WEIGHT BY finalwt.
```

Users should be aware that all analyses will be run with weights on until the following command is given:

```
WEIGHT OFF.
```

Because the weights sum to the final sample size of 6,394, simply looking at frequencies may not be sufficient to identify whether the weights are being used. If in doubt, users should look at the bottom right-hand corner of the Data Editor. A small box to the right of the notification that “IBM SPSS Statistics Processor is ready” will say “Weight On” when the weights are on and will be empty when weights are off.

Estimates of sampling error and other inferential statistics calculated in SPSS without the Complex Samples package may not be calculated correctly, particularly for subsamples. Users are advised to cross-check results with another statistical package.

3.3.3 Stata

For Stata, analyses should be set up using the `svyset` command. Specifically:

```
svyset [pweight=finalwt]
```

This is set and saved in the supplied Stata code.

In Stata, analyses should be run using commands that allow the `svy` prefix. This will ensure that point estimates are correct (i.e., weighted) and that measures of sampling error (e.g., margin of error) and inferential statistics (e.g., statistical tests) will be calculated correctly.

Analyses of subsamples should use the `subpop` option for `svy`, not `if`. For instance, a tabulation of `FWB1_1` for adults ages 62 and older would be `svy, subpop(if agecat>=6): tab FWB1_1` and not `svy: tab FWB1_1 if agecat>=6`. The use of `if` outside of `subpop` will result in Stata incorrectly assuming that sample size is only that of the subpopulation, not the entire sample. In the following examples, note the slightly larger confidence intervals on the second, incorrectly specified, tabulation below.

Correctly specified subsample:

```
. svy, subpop(if agecat==1): tab FWB1_1, ci format(%10.4fc)
(running tabulate on estimation sample)
```

```
Number of strata   =           1                Number of obs       =       6,394
```

```

Number of PSUs      =      6,394
Population size     =      6,394
Subpop. no. obs    =      414
Subpop. size       = 625.133408
Design df          =      6,393

```

```

-----
I could
handle a
major
unexpecte
d expense |
          | proportion          lb          ub
-----+-----
Refused   |      0.0163      0.0038      0.0661
Not at a  |      0.2234      0.1807      0.2727
Very lit  |      0.1949      0.1555      0.2415
Somewhat  |      0.3854      0.3310      0.4429
Very wel  |      0.1357      0.1030      0.1766
Complete  |      0.0444      0.0290      0.0673
Total     |      1.0000
-----

```

```

Key:  proportion = cell proportion
      lb         = lower 95% confidence bound for cell proportion
      ub         = upper 95% confidence bound for cell proportion

```

Incorrectly specified subsample:

```

. svy: tab FWB1_1 if agecat==1, ci format(%10.4fc)
(running tabulate on estimation sample)

```

```

Number of strata   =      1
Number of PSUs    =      414
Number of obs     =      414
Population size    = 625.133408
Design df         =      413

```

```

-----
I could
handle a
major
unexpecte
d expense |
          | proportion          lb          ub
-----+-----
Refused   |      0.0163      0.0038      0.0664
Not at a  |      0.2234      0.1805      0.2729
Very lit  |      0.1949      0.1553      0.2417
Somewhat  |      0.3854      0.3308      0.4431
Very wel  |      0.1357      0.1029      0.1768
Complete  |      0.0444      0.0290      0.0674
Total     |      1.0000
-----

```

```

Key:  proportion = cell proportion
      lb         = lower 95% confidence bound for cell proportion
      ub         = upper 95% confidence bound for cell proportion

```

3.3.4 Python

The python stub code provided will load the data from CSV format in to a format usable for subsequent python code and analysis. The code provides a template for remapping the values

from the dummy variables to the values provided in the codebook. Please adapt the code from the other software packages for use in remapping values as well as for applying study weights.

3.3.5 R

The R stub code provided will load the data from CSV format in to a format usable for subsequent R code and analysis. The code provides a template for remapping the values from the dummy variables to the values provided in the codebook. Please adapt the code from the other software packages for use in remapping values as well as for applying study weights.

3.4 Nonsubstantive value codes

Variables have been assigned the following missing or nonsubstantive value codes.

TABLE 21: NONSUBSTANTIVE VALUE CODES

Value	Label(s)	Description
-5	County not known	Applies to PCTLT200FPL. County was not known for these respondents and details on the county-level degree of poverty could not be merged in for these cases.
-4	Response not written to database	Applies to variables in the SWB and FWB series. A database error occurred for a single respondent whose responses were not recorded and could not be recovered.
-3	Invalid response	Applies to SOCSEC2 and SOCSEC3. For SOCSEC2, respondents who gave answers of less than 62 were coded as -3. For SOCSEC3, a value of -3 was entered for respondents where PPAGE (respondent age) is two or more years greater than the response to SOCSEC3.
-2	Question not asked because respondent not in item base	Applies to PAIDHELP, HOUSERANGES, VALUERANGES, MORTGAGE, SOCSEC1, SOCSEC2, SOCSEC3, LIFEEXPECT, and RETIRE. The respondent was correctly skipped over this item based on responses to previous questions or panel data on respondent age.
-1	Refused	Applies to most items. The respondent did not answer the item. This is not an explicit refusal.
8	I can't recall	Applies to HSLOC. An explicit "I can't recall" option was given for this item.
98	Don't know	Applies to HOUSERANGES, VALUERANGES, MORTGAGE, SAVINGSRANGES. An explicit "I don't know" option was given for these items.
99	Prefer not to say	Applies to HOUSERANGES, VALUERANGES, MORTGAGE, SAVINGSRANGES. An explicit "I prefer not to say" refusal option was given for these items.
99	Refused	Applies to EMPLOY. Assigned where respondent did not select any item in the EMPLOY1 series in the instrument. (EMPLOY is a composite variable constructed from items in the EMPLOY1 and EMPLOY2 series.)

4. Variable documentation

This section documents several created variables included in the PUF. Imputation for the HOUSERANGES variable is documented first, followed by cleaning for the KIDS_1, KIDS_2, KIDS_3, SOCSEC2 and SOCSEC3 variables. Variables constructed for the purpose of analysis are then documented.

4.1 Imputation for HOUSERANGES

As described in section 2.2.1, imputation of some values of HOUSERANGES was required because respondents were erroneously only asked this question if they reported owning their own home. After this error was identified, a follow-up survey was fielded to collect data from those not asked the question. In total, 1,823 of the 2,230 panel members responded to the follow-up survey. Values for HOUSERANGES were then imputed for the remaining 407 nonrespondents to the follow-up survey.

A regression-based imputation approach was selected. For cases that responded to the recontact survey, the response to HOUSERANGES was regressed on demographic variables (age, sex, and race/ethnicity), number of household members, and tract-level median rent from the American

Community Survey (ACS) 2015 5-Year Summary File using an unweighted ordinal logit model.³⁴ Cases with nonsubstantive responses (e.g., don't know, prefer not to say, and refusals) were excluded from the model.

It was decided that the deterministic nature of a regression-based imputation approach, which suppresses variance, was tolerable for this application, particularly in light of the small proportion of the total sample (6.4 percent) for which imputation was required.³⁵

It was necessary to make two additional imputation steps with respect to tract-level median rent. Median rent was missing for 2.8 percent of tracts in the ACS Summary File dataset.³⁶ For these tracts, the median gross rent of the immediately preceding tract was substituted. This procedure was used in order to take advantage of the relationship between tract ID and geography and assumes that median gross rent would be similar in nearby tracts. Overall, 1.4 percent of survey respondents resided in tracts where median gross rent was imputed. A second imputation step was required for the 396 respondents (6.2 percent) who could not be matched

³⁴ Weights were not used because the objective of the model was to make inference to the sample rather than the U.S. population as a whole. In addition, the model included the variables used in sample stratification—age, sex, and race/ethnicity—thus controlling for sample selection.

³⁵ Regression-based imputation is deterministic in that the same imputation value will be arrived at each time. See, e.g., Särndal, Carl-Erik and Sixten Lundström. 2005. *Estimation in Surveys with Nonresponse*. Hoboken, NJ: Wiley, regarding the impact of deterministic imputation on variance.

³⁶ This occurred in tracts where sample size was too low for reliable estimates or population size was small enough to raise privacy concerns. See U.S. Census Bureau. 2016. American Community Survey 2015 5-Year Summary File [MRDF]. Washington, D.C.: U.S. Census Bureau [producer][distributor]. Accessed March 1, 2017 from https://www2.census.gov/programs-surveys/acs/summary_file/2015/data/5_year_entire_sf/Tracts_Block_Groups_Only.tar.gz

to tracts from the ACS file.³⁷ For these cases, the grand median of median gross rent (i.e., a median of medians) in the survey dataset was imputed.

4.2 Data cleaning

This section describes steps to identify potentially problematic values and cleaning steps taken to address problematic responses. Cleaning of KIDS_1, KIDS_2, and KIDS_3 is described first, followed by SOCSEC2 and SOCSEC3.

4.2.1 KIDS_1, KIDS_2, and KIDS_3

KIDS_1, KIDS_2, and KIDS_3 ask for the number of children of specified age ranges (0 to 6, 7 to 12, and 13 to 17 years-old, respectively) that the respondent supports financially. In some cases, unreasonably high totals (e.g., 20 children) were found. All these items were numeric entry (i.e., the respondent entered digits), where measurement error is more likely than for close-ended types of items due to the uncontrolled nature of data entry.

Identification of outliers

Responses of 4 children and above were examined in combination with household size, respondent age, and panel data on the number of household members of various ages (0-1, 2-5, 6-12, and 13-17). As is described below, none of these allowed for direct and fully accurate checks; the exercise was one of pattern matching.

- Although household size is not a perfect match with the construct of number of children financially supported, it is indicative given that in most cases the children financially

³⁷ For 395 respondents who could not be matched, no tract was available. The other respondent had a tract not found in the ACS Summary File.

supported will be members of the respondent's household. Where 4 or more children were reported, and the number of children supported financially exceeded household size, the likelihood of measurement error was assumed to be greater.

- Respondent age was included because reports of financially supporting children or the presence of children at ages that would have required becoming a parent at a very young age were assumed to be indicative of potential measurement error.
- Panel data on the number of household members of various ages was included because of the close relationship that would be expected between the number of children of a given age in a household and the number of children supported financially. Where the number of children in a given age range was large, it was assumed that financially supporting a larger number of children would be more reasonable.

Based on examination of the patterns of response, there appeared to be two different etiologies of measurement error. One pattern appeared to be mispunches of numbers, e.g., entering "0" where "1" was intended. The other pattern appeared to be that some respondents misunderstood the question and thought it was asking for the age of the child in a given age range rather than the number of children, e.g., entering "10" for the age of the child between the ages of 7 and 12.

The possibility of malicious misentry was considered (i.e., deliberately providing unreasonable responses) but rejected on the assumption that a respondent who was deliberately misentering responses would be likely to provide unreasonable responses for each age group. There were no such cases.

In all but one case in KIDS_2, it was decided to either let the values stand or edit the response to the most likely answer. In the one exceptional case, the error could have been either a mispunch or because of misunderstanding and the type of error would have yielded different edit values. This case was given a value of -3 and labeled "Cleaned and no value substituted."

4.2.2 SOCSEC2 and SOCSEC3

For SOCSEC2, the following sets of cases were examined:

- Cases where the respondent reported collecting Social Security retirement benefits before the age of 62.

- Cases where the respondent reported starting collecting Social Security retirement benefits at an age older than their current age. (Only people who reported collecting Social Security retirement benefits in item SOCSEC1 were asked SOCSEC2.)

For SOCSEC3, we examined:

- Cases where respondents reported that they expected to begin collecting Social Security retirement benefits at an age younger than their current age.

Each of these scenarios is described in greater depth below.

With respect to respondents who reported collecting Social Security retirement benefits at an age younger than 62 in SOCSEC2, these values were set to -3 “Invalid response” because the earliest age one may receive Social Security retirement benefits is age 62.³⁸ In some cases, the respondent reported that they were permanently sick, disabled, or unable to work, suggesting that they may have been Social Security disability insurance recipients. In other cases, the respondent may have been counting the age they were when their spouse began receiving Social Security retirement benefits. In other cases, the respondent may have confused Social Security retirement benefits with other benefits.

Similar logic applied to respondents who reported expecting to start collecting Social Security retirement benefits at an age younger than their current age in SOCSEC3. A substantial number of these reported that they were permanently sick, disabled, or unable to work. In other cases, it seemed likely that the respondent was confusing Social Security retirement benefits with other types of benefits (e.g., SNAP or unemployment insurance). In no cases was evidence strong enough to support making edits other than to collapse the responses indicating anticipated Social Security retirement receipt at an age less than 62 into a single category of “61 and below.”

³⁸ There were 66 such cases. Another 2 cases were classified as “Invalid response” because the age they reported starting to receive Social Security retirement benefits was 2 or more years older than their current age. This is described below.

Two edits were made to SOCSEC2 for cases where the reported age of Social Security retirement benefits receipt was greater than the respondent's current age, as mentioned in footnote 38. The discrepancy in both cases exceeded one year, and it was not possible to infer what the correct response would be. Both cases had their values set to -3 and labeled "Invalid response." The remaining cases had discrepancies of one year and were not edited because the discrepancy was likely to be between the respondent's true age at the time of responding to that question and their reported age.³⁹

³⁹ The age variable, PPAGE, is not included in the PUF. GfK calculated age based upon a common date in the field period and not the respondent's exact age at the time they started or completed the survey.

4.3 Created and modified variables

This section describes variables created or modified from one or more survey or panel variables as part of the construction of the PUF. A number of the variables described in the coding process below are not included in the PUF, therefore, readers should refer to the survey instrument in Appendix A to identify the corresponding questions.

Modified variable: FWBscore

Description of variable: CFPB Financial Well-Being Scale score

Description of changes:

These changes assigned values and labels to cases that would otherwise have missing values.

- Set FWBscore = -4 “Response not written to dataset” where FWB1_1 to FWB2_3 were all equal to -4 “Response not written to dataset.” (See discussion of -4 in section 3.4.)
- Set FWBscore = -1 “Refused” where FWB1_1 to FWB2_3 were all equal to -1 “Refused.”
- Assigned appropriate value labels where FWBscore = -4 or FWBscore = -1.

Modified variable: FSscore

Description of variable: CFPB Financial Skill Scale score

Description of changes:

This change assigned values and labels to cases that would otherwise have missing values.

- Set FSscore = -1 “Refused” where FS1_1 to FS2_3 were all equal -1 “Refused.”
- Assigned appropriate value labels where FSscore = -1.

Modified variables: MANAGE1_1 to MANAGE1_4

Description of variables: Measures of financial management.

Description of recoding:

These changes collapsed values of “Not applicable” and “Never” into a single value.

- For each variable of MANAGE1_1 to MANAGE1_4:
- Set value to 1 “Not applicable or never” where values were equal to 1 “Not applicable” and 2 “Never”.
- Set value to 2 “Seldom” where value was equal to 3 “Seldom”.
- Set value to 3 “Sometimes” where value was 4 “Sometimes”.
- Set value to 4 “Often” where value was 5 “Often”.
- Set value to 5 “Always” where value 6 “Always”.
- Assigned appropriate value labels.

Created variable: LMscore

Description of variable: Lusardi and Mitchell scale score.

Description of construction:

This change created a summative scale score.

- FK1correct through FK3correct were summed for each respondent to create the LMscore variable.
- Label LMscore “Lusardi and Mitchell financial knowledge score”.

Created variables: FK1correct to FK3correct

Description of variables: Variables identifying whether the respondent correctly answered FINKNOWL1 through FINKNOWL3.

Description of recoding:

These changes created variables FK1correct to FK3correct where the correct response had a value of 1 and incorrect responses and refusals had values of 0.

- Set FK1correct = 0 “No” where FINKNOWL1 was equal to -1 “Refused”, 2 “Exactly \$102”, or 3 “Less than \$102”.
- Set FK1correct = 1 “Yes” where FINKNOWL1 = 1 “More than \$102”.

- Label FK1correct “FINKNOWL1 answered correctly”.
- Set FK2correct = 0 “No” where FINKNOWL2 was equal to -1 “Refused”, 1 “More than today”, or 2 “Exactly the same”.
- Set FK2correct = 1 “Yes” where FINKNOWL2 = 3 “Less than today”.
- Label FK2correct “FINKNOWL2 answered correctly”.
- Set FK3correct = 0 “No” where FINKNOWL3 was equal to -1 “Refused” or 1 “True”.
- Set FK3correct = 1 “Yes” where FINKNOWL3 = 2 “False”.
- Label FK3correct “FINKNOWL3 answered correctly”.
- Assign appropriate value labels to FK1correct to FK3correct.

Created variables: KH1correct to KH9correct

Description of variables: Variables identifying whether the respondent correctly answered KHKNOWL1 through KHKNOWL9.

Description of recoding:

These changes created variables KH1correct to KH9correct where the correct response had a value of 1 and incorrect responses and refusals had values of 0.

- Set KH1correct = 1 “Yes” where KHKNOWL1 = 3 “Stocks”.
- Set KH1correct = 0 “No” where KHKNOWL1 was equal to -1 “Refused”, 1 “Savings accounts”, or 2 “Bonds”.
- Label KH1correct “KHKNOWL1 answered correctly”.
- Set KH2correct = 1 “Yes” where KHKNOWL2 = 3 “Stocks”.
- Set KH2correct = 0 “No” where KHKNOWL2 was equal to -1 “Refused”, 1 “Savings accounts”, or 2 “Bonds”.
- Label KH2correct “KHKNOWL2 answered correctly”.
- Set KH3correct = 1 “Yes” where KHKNOWL3 = 2 “Decrease”.

- Set KH3correct = 0 “No” where KHKNOWL3 was equal to -1 “Refused”, 1 “Increase”, or 3 “Stay the same”.
- Label KH3correct “KHKNOWL3 answered correctly”.
- Set KH4correct = 1 “Yes” where KHKNOWL4 = 1 “True”.
- Set KH4correct = 0 “No” where KHKNOWL4 was equal to -1 “Refused” or 2 “False”.
- Label KH4correct “KHKNOWL4 answered correctly”.
- Set KH5correct = 1 “Yes” where KHKNOWL5 = 1 “True”.
- Set KH5correct = 0 “No” where KHKNOWL5 was equal to -1 “Refused” or 2 “False”.
- Label KH5correct “KHKNOWL5 answered correctly”.
- Set KH6correct = 1 “Yes” where KHKNOWL6 = 2 “False”.
- Set KH6correct = 0 “No” where KHKNOWL6 was equal to -1 “Refused” or 1 “True”.
- Label KH6correct “KHKNOWL6 answered correctly”.
- Set KH7correct = 1 “Yes” where KHKNOWL7 = 4 “Never, you will continue to be in debt”.
- Set KH7correct = 0 “No” where KHKNOWL7 was equal to -1 “Refused”, 1 “Less than 5 years”, 2 “Between 5 and 10 years”, or 3 “Between 10 and 15 years”.
- Label KH7correct “KHKNOWL7 answered correctly”.
- Set KH8correct = 1 “Yes” where KHKNOWL8 = 2 “Yes”.
- Set KH8correct = 0 “No” where KHKNOWL8 = 1 “No”.
- Label KH8correct “KHKNOWL8 answered correctly”.
- Set KH9correct = 1 “Yes” where KHKNOWL9 = 1 “No”.
- Set KH9correct = 0 “No” where KHKNOWL9 = 2 “Yes”.
- Label KH9correct “KHKNOWL9 answered correctly”.
- Assign appropriate value labels to KH1correct to KH9correct.

Created variables: ON2correct and ON1correct

Description of variables: Variables identifying whether the respondent correctly answered OBJNUMERACY2 and OBJNUMERACY1.⁴⁰

Description of recoding:

These changes created variables ON2correct and ON1correct where the correct response had a value of 1 and incorrect responses and refusals had values of 0.

- Set ON2correct = 1 “Yes” where OBJNUMERACY2 = 10.
- Set ON2correct = 0 “No” where OBJNUMERACY2 was equal to -1 “Refused”, 0 to 9, or 11 to 1000.
- Label ON2correct “OBJNUMERACY2 answered correctly”.
- Set ON1correct = 1 “Yes” where OBJNUMERACY1 = 2 “10%”.
- Set ON1correct = 0 “No” where OBJNUMERACY1 was equal to -1 “Refused”, 1 “1%”, or 3 “5%”.
- Label ON1correct “OBJNUMERACY1 answered correctly”.
- Assign appropriate value labels to ON2correct and ON1correct.

Modified variable: HOUSERANGES

Description of variable: Monthly amount paid for housing.

⁴⁰ OBJNUMERACY2 is not provided in the PUF.

Description of recoding: Top codes original version of HOUSERANGES at \$2,000 or more to reduce the risk of identification. These changes collapsed monthly amount paid for housing for ranges of \$2,000 or more.

- Set HOUSERANGES = 7 “\$2,000 or more” where HOUSERANGES was equal to 7 “\$2,000-2,999”, 8 “\$3,000-4,999”, or 9 “\$5,000 or more”.
- Assign appropriate value where HOUSERANGES = 7.

Modified variable: VALUERANGES

Description of variable: Estimated worth of home if sold today.

Description of recoding:

The changes collapsed home value ranges to reduce the risk of identification.

- Set VALUERANGES = 1 “Less than \$150,000” where VALUERANGES is equal to 1 “Less than \$50,000”, 2 “\$50,000-99,999”, or 3 “\$100,000-149,999”.
- Set VALUERANGES = 2 “\$150,000-249,999” where VALUERANGES is equal to 4 “\$150,000-199,999” or 5 “\$200,000-249,999”.
- Set VALUERANGES = 3 “\$250,000-399,999” where VALUERANGES is equal to 6 “\$250,000-299,999”, 7 “\$300,000-349,999”, or 8 “\$350,000-399,999”.
- Set VALUERANGES = 4 “\$400,000 or more” where VALUERANGES is equal to 9 “\$400,000-449,999”, 10 “\$450,000-\$499,999”, 11 “\$500,000-549,999”, 12 “\$550,000-599,999”, 13 “\$600,000-649,999”, or 14 “More than \$650,000”.
- Assign appropriate value labels where VALUERANGES is between 1 and 4.

Modified variable: MORTGAGE

Description of variable: Amount owed on home today.

Description of recoding:

These changes created broader value ranges for the MORTGAGE item to reduce the risk of identification and risk from disclosure.

- Set MORTGAGE = 2 “\$50,000-199,999” where MORTGAGE is equal to 2 “\$50,000-99,999”, 3 “\$100,000-149,999”, or 4 “\$150,000-199,999”.
- Set MORTGAGE = 3 “\$200,000 or more” where MORTGAGE is equal to 5 “\$200,000-249,999”, 6 “\$250,000-299,999”, 7 “\$300,000-349,999”, 8 “\$350,000-399,999”, 9 “\$400,000-449,999”, 10 “\$450,000-499,999”, 11 “\$500,000-549,999”, 12 “\$550,000-599,999”, 13 “\$600,000-649,999”, or 14 “More than \$650,000”.
- Assign appropriate value labels where MORTGAGE is 2 or 3.

Modified variable: SAVINGSRANGES

Description of variable: Amount of money in savings today.

Description of recoding:

These changes created broader value ranges for the SAVINGSRANGES item to reduce risk of identification and potential risk from disclosure.

- Set SAVINGSRANGES = 2 “\$1-99” where SAVINGSRANGES is equal to 2 “\$1-49” or 3 “\$50-99”.
- Set SAVINGSRANGES = 3 “\$100-999” where SAVINGSRANGES is equal to 4 “\$100-249” or 5 “\$250-499”, or 6 “\$500-999”.
- Set SAVINGSRANGES = 4 “\$1,000-4,999” where SAVINGSRANGES is equal to 7 “\$1,000-1,999” or 8 “\$2,000-4,999”.
- Set SAVINGSRANGES = 5 “\$5,000-19,999” where SAVINGSRANGES is equal to 9 “\$5,000-9,999” or 10 “\$10,000-19,999”.
- Set SAVINGSRANGES = 6 “\$20,000-74,999” where SAVINGSRANGES is equal to 11 “\$20,000-49,999” or 12 “\$50,000-74,999”.
- Set SAVINGSRANGES = 7 “\$75,000 or more” where SAVINGSRANGES is equal to 13 “\$75,000 or more”.
- Assign appropriate value labels where SAVINGSRANGES is 2 to 7.

Modified variable: HSLOC

Description of variable: Where respondent attended high school.

Description of recoding:

These changes collapsed U.S. states, the District of Columbia, and other U.S. territories and possessions into a single category and adjusted values where the respondent attended high school outside the U.S. or could not recall where they attended high school.

- Set HSLOC = 1 “U.S. and territories” where HSLOC is greater than or equal to 1 and less than or equal to 52 (i.e., “Alabama” to “Other Territories and Possessions”).
- Set HSLOC = 2 “Outside the U.S.” where HSLOC = 53 “Outside the U.S.”.
- Set HSLOC = 8 “I can’t recall” where HSLOC = 54 “I can’t recall”.
- Assign appropriate value label where HSLOC is between 1 and 8.

Modified variable: PAREDUC

Description of variable: Highest level of education of person/people who raised respondent.

Description of recoding:

These changes collapsed some college and Associate degree into a single category and adjusted values where PAREDUC was equal to 4 or 5.

- Set PAREDUC = 3 “Some college/Associate” where PAREDUC is equal to 3 “Some college” or 4 “Associates’ degree”.
- Set PAREDUC = 4 “Bachelor’s degree” where PAREDUC = 5 “Bachelor’s degree”.
- Set PAREDUC = 5 “Graduate/professional degree” where PAREDUC = 6 “Graduate/professional degree”.
- Assign appropriate value labels where PAREDUC is between 3 and 5.

Modified variable: SOCSEC2

Description of variable: Age began receiving Social Security retirement benefits.

Description of recoding:

These changes set values 19 to 61 to missing and collapsed values of 70 and older.

- Set SOCSEC2 = -3 “Cleaned and no value substituted” where SOCSEC2 is between 19 and 61 (ages 19 and 61).
- Set SOCSEC2 = 70 where SOCSEC2 is between 70 and maximum (age 70 and older).
- Assign appropriate value labels where SOCSEC2 is equal to -3 or 70.

Modified variable: SOCSEC3

Description of variable: Age expect to begin receiving Social Security retirement benefits.

Description of recoding: Set value to missing where value of original version of variable is 2 or more years less than current age then bottom-code age 61 and below and top-code age 71 and older. These changes set SOCSEC3 = -3 where the age the respondent reported receiving Social Security retirement benefits was two or more years above their current age and collapsed ages 61 and below and 71 and older.

- Set SOCSEC3 = -3 “Cleaned and no value substituted” where PPAGE (respondent age) minus SOCSEC3 is greater than or equal to 2 and SOCSEC3 is not missing and SOCSEC3 is greater than 0.
- Set SOCSEC3 = 61 “61 and below” where SOCSEC3 is between 18 and 61.
- Set SOCSEC3 = 71 “71 and above” where SOCSEC3 is between 71 and 99.
- Assign appropriate value labels where SOCSEC3 is equal to -3, 61 or 71.

Modified variable: HHEDUC

Description of variable: Highest level of education of all household members.

Description of recoding:

These changes collapsed some college and Associate degree into a single category and adjusted values where HHEDUC was equal to 4 or 5.

- Set HHEDUC = 3 “Some college/Associate” where HHEDUC is equal to 3 “Some college” or 4 “Associates’ degree”.
- Set HHEDUC = 4 “Bachelor’s degree” where HHEDUC = 5 “Bachelors’ degree”.
- Set HHEDUC = 5 “Graduate/professional degree” where HHEDUC = 6 “Graduate/professional degree”.
- Assign appropriate value labels where HHEDUC is between 3 and 5.

Modified variables: KIDS_NoChildren

Description of variable: Respondent does not financially support children.

Description of changes:

The first two changes assigned missing values for cases where there was no evidence that the respondent answered either the KIDS_NoChildren or KIDS_1 to KIDS_4 items and assigned missing values for one case where the respondent left a missing value for KIDS_1 and values of 0 for KIDS_2 to KIDS_4. The third change set KIDS_NoChildren to 0 where the respondent had clearly indicated he or she supported children in this age group and only the number of children supported was unclear.

- Set KIDS_NoChildren = -1 “Refused” where responses to KIDS_1 to KIDS_4 were left at default values of 0 and KIDS_NoChildren was missing.
- Set KIDS_NoChildren = -1 “Refused” for a single case where KIDS_1 = -1 “Refused” and KIDS_2 = 0 and KIDS_3 = 0 and KIDS_4 = 0.
- Set KIDS_NoChildren = 0 “Respondent financially supports children” for a single case where KIDS_2 = -3.
- Assign appropriate value labels.

Modified variables: KIDS_1 to KIDS_4

Description of variables: Number of children of specified ages financially supported by respondent.

Description of recoding:

This change set values of 2 and above to 2 to reduce the risk of identification.

- Set value of KIDS_1 to KIDS_4 to 2 “2+” where the value for that variable was greater than or equal to 2.
- Assign appropriate value labels where KIDS_1 to KIDS_4 are equal to 2.

Created variable: EMPLOY

Description of variable: Primary or only employment status.

Description of changes:

A new variable called EMPLOY was created. If only one variable was selected in EMPLOY1_1 to EMPLOY1_8, that variable was recoded into EMPLOY. If more than one variable was selected in EMPLOY1_1 to EMPLOY1_8, the option selected from EMPLOY2_1 to EMPLOY2_8 is recorded in EMPLOY. If EMPLOY1_9 was selected (i.e., no selection in EMPLOY1_1 to EMPLOY1_8), EMPLOY was set to 99 and labeled a refusal.

- Set EMPLOY = 1 “Self-employed” where EMPLOY1_1 = 1 and EMPLOY1_2 to EMPLOY1_9 are all equal to 0.
- Set EMPLOY = 2 “Work full-time for an employer or the military” where EMPLOY1_2 = 1 and EMPLOY1_1 and EMPLOY1_3 to EMPLOY1_9 are all equal to 0.
- Set EMPLOY = 3 “Work part-time for an employer or the military” where EMPLOY1_3 = 1 and EMPLOY1_1, EMPLOY1_2 and EMPLOY1_4 to EMPLOY1_9 are all equal to 0.
- Set EMPLOY = 4 “Homemaker” where EMPLOY1_4 = 1 and EMPLOY1_1 to EMPLOY1_3 and EMPLOY1_5 to EMPLOY1_9 are all equal to 0.
- Set EMPLOY = 5 “Full-time student” where EMPLOY1_5 = 1 and EMPLOY1_1 to EMPLOY1_4 and EMPLOY1_6 to EMPLOY1_9 are all equal to 0.
- Set EMPLOY = 6 “Permanently sick, disabled or unable to work” where EMPLOY1_6 = 1 and EMPLOY1_1 to EMPLOY1_5 and EMPLOY1_7 to EMPLOY1_9 are all equal to 0.
- Set EMPLOY = 7 “Unemployed or temporarily laid off” where EMPLOY1_7 = 1 and EMPLOY1_1 to EMPLOY1_6, EMPLOY1_8, and EMPLOY1_9 are all equal to 0.

- Set EMPLOY = 8 “Retired” where EMPLOY1_8 = 1 and EMPLOY1_1 to EMPLOY1_7 and EMPLOY1_9 are all equal to 0.
- Set EMPLOY = 1 “Self-employed” where EMPLOY2_1 = 1.⁴¹
- Set EMPLOY = 2 “Work full-time for an employer or the military” where EMPLOY2_2 = 1.
- Set EMPLOY = 3 “Work part-time for an employer or the military” where EMPLOY2_3 = 1.
- Set EMPLOY = 4 “Homemaker” where EMPLOY2_4 = 1.
- Set EMPLOY = 5 “Full-time student” where EMPLOY2_5 = 1.
- Set EMPLOY = 6 “Permanently sick, disabled or unable to work” where EMPLOY2_6.
- Set EMPLOY = 7 “Unemployed or temporarily laid off” where EMPLOY2_7 = 1.
- Set EMPLOY = 8 “Retired” where EMPLOY2_8 = 1.
- Set EMPLOY = 99 “Refused” where EMPLOY1_9 = 1 and EMPLOY1_1 to EMPLOY1_8 are all equal to 0.
- Assign appropriate value labels.
- Label EMPLOY “Primary or only employment status”.

Modified variable: MILITARY

Description of variable: Variable identifying whether respondent or spouse/dependent is a current or former member of the U.S. Armed Forces.

⁴¹ The EMPLOY2 series of variables are not provided in the PUF.

Description of recoding:

These changes collapsed all respondents who were a service member or veteran or were a dependent or spouse of a service member or veteran into a single category to reduce the risk of identification. The value for respondents who were neither a service member or veteran nor the dependent or spouse of a service member or veteran was set to 0.

- Set MILITARY = 1 “Yes” where MILITARY was between 1 and 3 (i.e., respondent was a service member or veteran, was a dependent or spouse of a service member or veteran, or both).
- Set MILITARY = 0 “No” where MILITARY = 4 “Neither”.
- Assign appropriate value labels.

Created variable: Military_Status

Description of variable: Records military status of respondents and spouse/dependent at a higher level of generality to reduce risk of identification. Respondents who refused MILITARY or indicated that they or their spouse/dependent had served but did not select an item in the MILITARYSTATUS or SPOUSEMILSTATUS banks of items were set to -1 “Refused”.⁴²

Description of construction:

These changes assigned respondents into five categories: -1 = Refused either MILITARY or selected no response in MILITARYSTATUS or SPOUSEMILSTATUS banks; 1 = Currently serving in the U.S. Military, the Reserves, or the National Guard (regardless of status of spouse/dependent); 2 = Retired from the U.S. Military or a veteran (regardless of status of spouse/dependent); 3 = Spouse/dependent is currently serving in the U.S. Military, the Reserves, or the National Guard (respondent is neither currently serving nor retired/a veteran);

⁴² The MILITARYSTATUS and SPOUSEMILSTATUS series of variables are not included in the PUF.

4 = Spouse/dependent is retired from the U.S. military or a veteran (respondent is neither currently serving nor retired/a veteran); 5 = Neither respondent nor spouse/dependent has served in the U.S. Military. These changes were made to reduce the risk of identification.

- Set Military_Status = -1 “Refused” where MILITARY = -1 “Refused”.
- Set Military_Status = -1 “Refused” where Military_Status was missing and either MILITARYSTATUS6 = 1 (respondent did not select any MILITARYSTATUS items if these were asked) or SPOUSEMILSTATUS6 = 1 (respondent did not select any SPOUSEMILSTATUS items if these were asked).
- Set Military_Status = 1 “Active military (active, reserve, natl guard)” where PPAGE (i.e., respondent age) was less than 60 and any of MILITARYSTATUS1 to MILITARYSTATUS3 (i.e., respondent was active duty, in the reserves, or the National Guard) were equal to 1 and any of MILITARYSTATUS4 (i.e., respondent is retired) and MILITARYSTATUS5 (i.e., respondent is a veteran) were equal to 1.
- Set Military_Status = 2 “Veteran (veteran or retired)” where PPAGE \geq 60 and any of MILITARYSTATUS1 to MILITARYSTATUS3 were equal to 1 and any of MILITARYSTATUS4 and MILITARYSTATUS5 were equal to 1.
- Set Military_Status = 1 “Active military (active, reserve, natl guard)” where any of MILITARYSTATUS1 to MILITARYSTATUS3 were equal to 1 and MILITARYSTATUS4 \neq 1 and MILITARYSTATUS5 \neq 1.
- Set Military_Status = 2 “Veteran (veteran or retired)” where MILITARYSTATUS1 \neq 1 and MILITARYSTATUS2 \neq 1 and MILITARYSTATUS3 \neq 1 and either MILITARYSTATUS4 or MILITARYSTATUS5 was equal to 1.
- Set Military_Status = 3 “Spouse/dependent Active” where Military_Status was missing and PPAGE \leq 60 and any of SPOUSEMILSTATUS1 to SPOUSEMILSTATUS3 (i.e., spouse or dependent is active duty, in the reserves, or the National Guard) was equal to 1 and any of SPOUSEMILSTATUS4 (i.e., spouse or dependent is retired) and SPOUSEMILSTATUS5 (i.e., spouse or dependent is a veteran) was equal to 1.
- Set Military_Status = 4 “Spouse/dependent Veteran” where Military_Status was missing and PPAGE \geq 60 and any of SPOUSEMILSTATUS1 to SPOUSEMILSTATUS3 were equal to 1 and any of SPOUSEMILSTATUS4 and SPOUSEMILSTATUS5 were equal to 1.

- Set Military_Status = 3 “Spouse/dependent Active” where Military_Status was missing and any of SPOUSEMILSTATUS1 to SPOUSEMILSTATUS3 were equal to 1 and SPOUSEMILSTATUS4 ≠ 1 and SPOUSEMILSTATUS5 ≠ 1.
- Set Military_Status = 4 “Spouse/dependent Veteran” where Military_Status was missing and SPOUSEMILSTATUS1 ≠ 1 and SPOUSEMILSTATUS2 ≠ 1 and SPOUSEMILSTATUS3 ≠ 1 and either of SPOUSEMILSTATUS4 or SPOUSEMILSTATUS5 were equal to 1.
- Set Military_Status = 5 “Neither Active nor Veteran” where MILITARY = 4 “Neither”.
- Assign appropriate value labels.
- Label Military_Status “Military status”.

Created variable: agecat

Description of variable: Age categories for descriptive analysis: 18-24, 25-34, 35-44, 45-54, 55-61, 62-69, 70-74, and 75+ years of age. These age categories were created from the continuous age variable, PPAGE, which was calculated based on date of birth at the beginning of the survey field period and is not included in the PUF.

Description of recoding:

These changes recoded the panel variable PPAGE into broader age categories to reduce the risk of identification.

- Set agecat = 1 “18-24” where PPAGE (i.e., respondent age) is between 18 and 24.
- Set agecat = 2 “25-34” where PPAGE is between 25 and 34.
- Set agecat = 3 “35-44” where PPAGE is between 35 and 44.
- Set agecat = 4 “45-54” where PPAGE is between 45 and 54.
- Set agecat = 5 “55-61” where PPAGE is between 55 and 61.
- Set agecat = 6 “62-69” where PPAGE is between 62 and 69.
- Set agecat = 7 “70-74” where PPAGE is between 70 and 74.
- Set agecat = 8 “75+” where PPAGE is ≥ 75.

- Assign appropriate value labels.
- Label agecat “Age”.

Created variable: generation

Description of variable: Generations for descriptive analysis: Millennial (ages 18-35), Gen X (ages 36-51), Boomer (ages 52-70), and Pre-Boomer (ages 71 and older), which represent respondent generation at the beginning of the survey field period in fall 2016. These age categories were created from the continuous age variable, PPAGE, which was calculated based on date of birth at the beginning of the survey field period and is not included in the PUF.

Description of recoding:⁴³

These changes recoded the panel variable PPAGE into age categories aligned to generations for user convenience.

- Set generation = 4 “Millennial” where PPAGE (i.e. respondent age) is between 18 and 35.
- Set generation = 3 “Gen X” where PPAGE is between 36 and 51.
- Set generation = 2 “Boomer” where PPAGE is between 52 and 70.
- Set generation = 1 “Pre-Boomer” where PPAGE \geq 71.
- Assign appropriate value labels.

Modified variable: PPEDUC

Description of variable: Respondent’s education.

⁴³ PPAGE is not included in the PUF.

Description of recoding:

These changes collapsed categories of the panel variable PPEDUC for consistency with other education items and to reduce the risk of identification.

- Set PPEDUC = 1 “Less than high school” where PPEDUC is equal to between 1 “No formal education” and 8 “12th grade NO DIPLOMA”.
- Set PPEDUC = 2 “High school degree/GED” where PPEDUC = 9 “HIGH SCHOOL GRADUATE - high school DIPLOMA”.
- Set PPEDUC = 3 “Some college/Associate” where PPEDUC is equal to 10 “Some college, no degree” or 11 “Associate degree”.
- Set PPEDUC = 4 “Bachelor’s degree” where PPEDUC = 12 “Bachelors degree”.
- Set PPEDUC = 5 “Graduate/professional degree” where PPEDUC is equal to 13 “Masters degree” or 14 “Professional or Doctorate degree”.
- Assign appropriate value labels.

Modified variable: PPETHM

Description of variable: Respondent’s race/ethnicity.

Description of recoding:

These changes collapsed the 2+ races non-Hispanic category with the other non-Hispanic category from the panel variable PPETHM variable to reduce the risk of identification.

- Set PPETHM = 3 “Other, Non-Hispanic” where PPETHM is equal to 3 “Other, Non-Hispanic” or 5 “2+ Races, Non-Hispanic”.

Modified variable: PPHHSIZE

Description of variable: Household size.

Description of recoding:

This change top-coded household sizes 5 and above in the panel variable PPHHSIZE to reduce the risk of identification.

- Set PPHHSIZE = 5 “5+” if PPHHSIZE ≥ 5.
- Assign appropriate value labels where PPHHSIZE = 5.

Modified variable: PPINCIMP

Description of variable: Household income.

Description of recoding:

This change collapsed categories of the panel variable PPINCIMP to reduce risk of identification.

- Set PPINCIMP = 1 “Less than \$20,000” where PPINCIMP is between 1 “Less than \$5,000” and 6 “\$15,000 to \$19,999”.
- Set PPINCIMP = 2 “\$20,000 to 29,999” where PPINCIMP is equal to 7 “\$20,000 to \$24,999” or 8 “\$25,000 to \$29,999”.
- Set PPINCIMP = 3 “\$30,000 to 39,999” where PPINCIMP is equal to 9 “\$30,000 to \$34,999” or 10 “\$35,000 to \$39,999”.
- Set PPINCIMP = 4 “\$40,000 to 49,999” where PPINCIMP = 11 “\$40,000 to \$49,999”.
- Set PPINCIMP = 5 “\$50,000 to 59,999” where PPINCIMP = 12 “\$50,000 to \$59,999”.
- Set PPINCIMP = 6 “\$60,000 to 74,999” where PPINCIMP = 13 “\$60,000 to \$74,999”.
- Set PPINCIMP = 7 “\$75,000 to 99,999” where PPINCIMP is equal to 14 “\$75,000 to \$84,999” or 15 “\$85,000 to \$99,999”.
- Set PPINCIMP = 8 “\$100,000 to 149,999” where PPINCIMP is equal to 16 “\$100,000 to \$124,999” or 17 “\$125,000 to \$149,999”.
- Set PPINCIMP = 9 “\$150,000 or more” where PPINCIMP is equal to 18 “\$150,000 to \$174,999” to 21 “\$250,000 or more”.
- Assign appropriate value labels.

Modified variable: PPMARIT

Description of variable: Marital status of respondent.

Description of recoding:

These changes collapsed divorced and separated.

- Set PPMARIT = 3 “Divorced/Separated” where PPMARIT is equal to 3 “Divorced” or 4 “Separated”.
- Set PPMARIT = 4 “Never married” where PPMARIT = 5 “Never married”.
- Set PPMARIT = 5 “Living with partner” where PPMARIT = 6 “Living with partner”.
- Assign appropriate value labels where PPMARIT is between 3 and 5.

Modified variables: PPT01, PPT25, PPT612, PPT1317

Description of variables: Number of children of various ages in the household.

Description of recoding:

These changes top-coded the PPT01 to PPT1317 variables at 1 or above to reduce the risk of identification.

- For each variable of PPT01 to PPT1317:
- Set the value of the variable equal to 1 “1+” where the value is 1 or higher.
- Assign appropriate value labels.

Modified variable: PPT18OV

Description of variable: Number of household members aged 18 and older.

Description of recoding:

This change top coded the panel variable PPT18OV at 4 and above to reduce the risk of identification.

- Set PPT18OV = 4 “4+” where $PPT18OV \geq 4$.
- Assign appropriate value labels.

Modified variable: PCTLT200FPL

Description of variable: Percentage of respondent's county of residence below 200% of the federal poverty level.

Description of recoding:

The changes recoded PCTLT200FPL into a binary variable around 40 percent of population of a county being below 200 percent of the federal poverty level to reduce the risk of identification.

- Set PCTLT200FPL = 0 "Less than 40% of county population below 200% of poverty level" where $PCTLT200FPL \leq 40$.
- Set PCTLT200FPL = 1 "40% or more of county population below 200% of poverty level" where $PCTLT200FPL > 40$.
- Set missing values of PCTLT200FPL to -5 "County not known".
- Assign appropriate value labels.

Appendix A: SURVEY INSTRUMENT

This appendix reproduces the survey instrument in its entirety.

Words in bold black text in brackets are programming instructions. Programming instructions include:

- Eligibility: [TERMINATE] for respondents who did not consent; [CONTINUE] for respondents who provided consent.
- Display order: [RANDOMIZE] to randomize the order in which statements were shown; [RANDOMIZE AND RECORD ORDER] to randomize the order in which response options are shown.
- Prompting in the case of nonresponse: [PROMPT] to indicate that a soft prompt should be used for that item. A soft prompt allows the respondent to dismiss the prompt and continue the survey without answering.
- Question type: [GRID] for table items; [SINGLE SELECT PER ROW] for single select (i.e., multiple selections are not allowed) per row in table items; [SINGLE SELECT] for single select for stand-alone (i.e., not table) items; [MULTIPLE SELECT] for multiple select stand-alone items; [DROP DOWN] for drop-down items; [NUMERIC ENTRY, RANGE] for numeric entry items specifying the allowable range for numeric entries.
- Display: [DISPLAY] for text to be displayed where no question was asked; [INSERT PAGE BREAK] for page breaks; [NEED SPACE HERE] for line breaks between substantive and nonsubstantive options in stand-alone items. DEFAULT is specified where a specific value was shown as the default for a numeric entry item.

Question names in bold can be used to connect items on this survey instrument to information on created and modified variables in section 4.3 of this PUF User's Guide, and to variable names in the PUF Codebook. Not all items are included in the PUF. In some cases, the name of a question was changed subsequent to being programmed. The new name is shown in bold red font.

In table items, statements in rows are numbered. These are combined with question names to provide the name for that specific item. For instance, in question SWB, the second statement ("I am optimistic about my future") would be named SWB_2. Similarly, response options in

multiple select items are numbered. For instance, in question EMPLOY1, the eighth option (“Retired”) would be named EMPLOY1_8.

Response options are numbered with the numeric values found in the dataset.

The respondents only saw the questions and response options; they did not see these instructions or question labels. Questions are listed below in the order in which they were presented to respondents. Not all questions were shown to all respondents. The skip patterns used to determine which respondents saw each question are listed as the “Base” above each question.

CFPB Study

Consent

CONSENT 1

For the following survey, please choose an option below.

1. I consent to participating in this survey without reading the consent
2. I would like to review the consent to see how my responses will be used and shared
3. I do not agree to participate in this survey

[TERMINATE IF CONSENT1 REFUSED or #3]

Base: IF CONSENT 1 = 2

CONSENT 2

CFPB National Financial Well-Being Survey Informed Consent

What is the purpose of the CFPB National Financial Well-Being Survey?

The purpose of this study is to examine how consumers make financial decisions and how those decisions might relate to the well-being of household members. As part of this study, you will be asked to answer a short online survey asking you about your financial decision-making and well-being. We expect the survey will take about 20 minutes. The data collected may be used for future research.

What is the CFPB?

The Consumer Financial Protection Bureau (CFPB) is a Federal agency created in 2010. One mission of the CFPB is to empower consumers to take control over their financial lives. This study will help the CFPB better understand how consumers make financial decisions and how those decisions may relate to financial well-being.

Who will see my responses, and how will my responses be used?

Your responses, along with those of approximately 6,000 other respondents, will be used by the study team and other researchers to understand consumers' experiences with financial decision-making and financial well-being. The study team includes staff at Abt Associates, Abt SRBI, GfK, the Center for Financial Security (CFS) at the University of Wisconsin Madison and the CFPB. The CFPB may also make an anonymous version of the survey data publicly available. These publicly available files will not contain any information that will personally identify you.

Who will have access to information that identifies me?

No one outside of GfK will be provided information that identifies you personally. Any personally-identifiable information that exists is available only to GfK. You will be asked no direct identifying information in the survey.

Risks

The risks to you of participating in this study are minimal. This study involves you answering questions about your financial decision-making and well-being. We recognize that you might be uncomfortable answering some questions about finances in general.

Benefits

There are no anticipated benefits to you. However, by sharing your thoughts and ideas you will help us better understand how to improve the financial well-being of consumers in America like you.

Right to Refuse or Withdraw

You may choose not to participate in the survey at any time. You may choose not to answer a question for whatever reason by skipping the question and/or selecting a 'refuse to answer' option. Participation is completely voluntary. The points you receive from GfK are, however, dependent on your completion of the survey.

Paperwork Reduction Act

According to the Paperwork Reduction Act of 1995, an agency may not conduct or sponsor, and a person is not required to respond to a collection of information unless it displays a valid OMB control number. The OMB control number for this collection is 3170-0063. It expires on 07/31/2019. The time required to complete this information collection is estimated to average 20 minutes per survey, including the time for reviewing any instructions, searching existing data sources, gathering and maintaining the data needed, and completing and reviewing the

collection of information. The obligation to respond to this collection of information is voluntary. Comments regarding this collection of information, including the estimated response time, suggestions for improving the usefulness of the information, or suggestions for reducing the burden to respond to this collection should be submitted to Bureau at the Consumer Financial Protection Bureau (Attention: PRA Office), 1700 G Street NW, Washington, DC 20552, or by email to CFPB_PRA@cfpb.gov .

The Bureau will not disclose any personally identifiable information collected except to the extent that it is required to do so by law and as provided in the Privacy Act Statement listed below. Additionally, the Bureau will treat the information collected consistent with its confidentiality regulations at 12 C.F.R. Part 1070, et seq.

Privacy Act Statement:

This collection of information is authorized by Pub. L. No. 111-203, Title X, Sections 1013 and 1022, codified at 12 U.S.C. §§ 5493 and 5512. Participation in this study is voluntary and there are no penalties for refusing to answer any question. However, your participation is extremely important to ensure the accuracy of the statistical results. The information you provide will be used to help the Consumer Financial Protection Bureau (“CFPB”) understand consumers’ financial decision-making experiences and how those decisions affect general financial well-being. You were randomly selected for this survey as part of a sample that is representative of consumers. Your responses will be combined with other data that we have collected to understand conditions in the consumer marketplace in a way that you cannot be identified. Routine uses which may be made of the collected information can be found in the CFPB’s System of Records Notice, CFPB.022 –Market and Consumer Research Records, 77 FR 67802. The CFPB may make an anonymous version of the survey data publicly available in accordance with applicable federal law.

Contact Information

If you have any questions about the study, you may contact Donna DeMarco of Abt Associates at 1-617-349-2322 or donna_demarco@abtassoc.com or Dr. Dee Warmath at the University of Wisconsin-Madison, at 608-262-2312 or warmath@wisc.edu.

If you are not satisfied with the response of the research team, have more questions, or want to talk with someone about your rights as a research participant, you may contact Katie Spearburg of Abt’s Institutional Review Board at 617-520-2499 or the Education and Social/Behavioral Science IRB Office at 608-263-2320. Please note these numbers are not toll-free.

If you consent to the survey, you may use the link below to begin. Selecting the link indicates that you have read the above and agree to take part in this research.

Do you agree to the terms described above?

1. Yes [**CONTINUE**]
2. No [**TERMINATE**]

[**TERMINATE IF REFUSED**]

MAIN QUESTIONNAIRE

[**INTRO DISPLAY**]

Please note that the following survey will show one or more questions on a single screen. Please scroll down to make sure you are able to view and respond to all questions.

Base: All respondents

SWB [GRID, SINGLE SELECT PER ROW]

Please indicate the degree to which you agree or disagree with each of the following statements.

Statements in a row [**RANDOMIZE**]:

1. I am satisfied with my life.
2. I am optimistic about my future.
3. If I work hard today, I will be more successful in the future.

Answers in a column:

7. Strongly agree
6. .
5. .
4. .
3. .
2. .
1. Strongly disagree

[**INSERT PAGE BREAK**]

Base: All respondents

FWB1 [GRID, SINGLE SELECT PER ROW] [PROMPT]

Please select the response that best indicates how well, in general, each of the following statements describes you or your situation.

This statement describes me....

Statements in a row [**RANDOMIZE**]:

1. I could handle a major unexpected expense.

2. I am securing my financial future.
3. Because of my money situation, I feel like I will never have the things I want in life.
4. I can enjoy life because of the way I'm managing my money.
5. I am just getting by financially.
6. I am concerned that the money I have or will save won't last.

Answers in a column:

5. Completely
4. Very well
3. Somewhat
2. Very little
1. Not at all

Base: All respondents

FWB2 [GRID, SINGLE SELECT PER ROW] [PROMPT]

For this next set of statements, we are interested in learning how often you would say that each statement applies to you and/or your situation.

How often would you say...?

Statements in a row **[RANDOMIZE]:**

1. Giving a gift for a wedding, birthday or other occasion would put a strain on my finances for the month.
2. I have money left over at the end of the month.
3. I am behind with my finances.
4. My finances control my life.

Answers in a column:

5. Always
4. Often
3. Sometimes
2. Rarely
1. Never

[INSERT PAGE BREAK]

Base: All respondents

FS1 [GRID, SINGLE SELECT PER ROW] [PROMPT]

Please select the response that best indicates how well, in general, each of the following statements describes you or your situation.

This statement describes me....

Statements in a row **[RANDOMIZE]:**

1. I know how to get myself to follow through on my financial intentions.

2. I know where to find the advice I need to make decisions involving money.
3. I know how to make complex financial decisions.
4. I am able to make good financial decisions that are new to me.
5. I am able to recognize a good financial investment.
6. I know how to keep myself from spending too much.
7. I know how to make myself save.

Answers in a column:

5. Completely
4. Very well
3. Somewhat
2. Very little
1. Not at all

Base: All respondents

FS2 [GRID, SINGLE SELECT PER ROW] [PROMPT]

For this next set of statements, we are interested in learning how often this statement applies to you or your situation.

How often would you say...?

Statements in a row **[RANDOMIZE]**:

1. I know when I do not have enough information to make a good decision involving my money.
2. I know when I need advice about my money.
3. I struggle to understand financial information.

Answers in a column:

5. Always
4. Often
3. Sometimes
2. Rarely
1. Never

[INSERT PAGE BREAK]

Base: All respondents

SUBKNOWL1 [SINGLE SELECT]

How would you assess your overall financial knowledge?

7. Very high
- 6.
- 5.
- 4.
- 3.
- 2.
1. Very low

Base: All respondents

ACT1 [GRID, SINGLE SELECT PER ROW]

Please select the response that best indicates how well, in general, each of the following statements describes you.

Statements in a row **[RANDOMIZE]**:

1. I follow-through on my financial commitments to others.
2. I follow-through on financial goals I set for myself.

Answers in a column:

5. Completely
4. Very well
3. Somewhat
2. Very little
1. Not at all

Base: All respondents

FINGOALS [SINGLE SELECT]

Some people set financial goals for themselves, such as planning to buy a car, becoming debt-free, paying for a vacation or preparing for college expenses. Do you have a current or recent financial goal?

1. Yes
0. No

[INSERT PAGE BREAK]

Base: All respondents

PROPPLAN [GRID, SINGLE SELECT PER ROW] [PROMPT]

We are trying to understand how people differ in how much they plan for money. Don't feel compelled to say that you engage in a lot of planning. We are just as interested in when you do not engage in much planning as when you do engage in a lot of planning. To what extent do you agree or disagree with each of the following statements?

Statements in a row **[RANDOMIZE]**:

1. I consult my budget to see how much money I have left.
2. I actively consider the steps I need to take to stick to my budget.
3. I set financial goals for what I want to achieve with my money.
4. I prepare a clear plan of action with detailed steps to achieve my financial goals.

Answers in a column:

5. Strongly agree
4. Agree
3. Neither agree nor disagree
2. Disagree
1. Strongly disagree

Base: All respondents

MANAGE1 [GRID, SINGLE SELECT PER ROW] [PROMPT]

Please indicate how often you have engaged in the following activities in the **past six months**:

Statements in a row **[RANDOMIZE]**:

1. Paid all your bills on time.
2. Stayed within your budget or spending plan.
3. Paid off credit card balance in full each month.
4. Checked your statements, bills and receipts to make sure there were no errors.

Answers in a column:

6. Always
5. Often
4. Sometimes
3. Seldom
2. Never
1. Not applicable

Base: All respondents

SAVEHABIT [GRID, SINGLE SELECT PER ROW]

To what extent do you agree or disagree with the following statements:

Statements in a row **[RANDOMIZE]**:

1. Putting money into savings is a habit for me. **[Was named SAVEHABIT_1. Renamed SAVEHABIT.]**
2. If I can re-use an item I already have, there's no sense in buying something new. **[Was SAVEHABIT_2. Renamed FRUGALITY.]**

Answers in a column:

6. Strongly agree
5. Agree
4. Agree slightly
3. Disagree slightly
2. Disagree
1. Strongly disagree

[INSERT PAGE BREAK]

Base: All respondents

AUTOMATED [GRID, SINGLE SELECT PER ROW]

Do you currently have money automatically transferred to:

Statements in a row **[RANDOMIZE]**:

1. A Retirement Savings Account
2. A Non-Retirement Savings Account

Answers in a column:

1. Yes
0. No
7. I do not have this type of account

Base: All respondents

ASK1 [GRID, SINGLE SELECT PER ROW]

Thinking about all the *large and small decisions you make involving money*, please indicate how often you do each of the following.

Statements in a row **[RANDOMIZE]**:

1. I do my own research before making decisions involving money.
2. I ask other people their opinions before making decisions involving money.

Answers in a column:

5. Always
4. Often
3. Sometimes
2. Seldom
1. Never

[INSERT PAGE BREAK]

Base: All respondents

SUBNUMERACY2 [SINGLE SELECT]

When people tell you the chance of something happening, do you prefer that they use words (“It rarely happens”) or numbers (“There’s a 1% chance”)?

1. 1 - Always prefer words
2. 2
3. 3
4. 4
5. 5
6. 6 - Always prefer numbers

Base: All respondents

SUBNUMERACY1 [SINGLE SELECT]

How good are you at working with percentages?

6. 6 – Extremely good
5. 5
4. 4
3. 3
2. 2
1. 1 – Not good at all

Base: All respondents

CHANGEABLE [SINGLE SELECT]

To what extent do you agree or disagree with the following statement: *A person's ability to manage money is something very basic about them and it cannot be changed much.*

7. Strongly agree
6. Agree
5. Somewhat agree
4. Neither agree nor disagree
3. Somewhat disagree
2. Disagree
1. Strongly disagree

Base: All respondents

GOALCONF [SINGLE SELECT]

If you were to set a financial goal for yourself today, how confident are you in your ability to achieve it?

4. Very confident
3. Somewhat confident
2. Not very confident
1. Not at all confident

[INSERT PAGE BREAK]

[DISPLAY]

Next are some questions about how financial products work. Please answer to the best of your ability.

Base: All respondents

FINKNOWL1 [SINGLE SELECT] [PROMPT]

Suppose you had \$100 in a savings account and the interest rate was 2% per year. After 5 years, how much do you think you would have in the account if you left the money to grow?

1. More than \$102
2. Exactly \$102
3. Less than \$102

Base: All respondents

FINKNOWL2 [SINGLE SELECT] [PROMPT]

Imagine that the interest rate on your savings account was 1% per year and inflation was 2% per year. After 1 year, how much would you be able to buy with the money in this account?

1. More than today
2. Exactly the same
3. Less than today

Base: All respondents

FINKNOWL3 [SINGLE SELECT] [PROMPT]

Do you think the following statement is true or false?

“Buying a single company’s stock usually provides a safer return than a stock mutual fund.”

1. True
2. False

Base: All respondents

KHKNOWL1 [SINGLE SELECT] [PROMPT]

Considering a long time period (for example 10 or 20 years), which asset described below normally gives the highest return?

1. Savings accounts
2. Bonds
3. Stocks

Base: All respondents

KHKNOWL2 [SINGLE SELECT] [PROMPT]

Normally, which asset described below displays the highest fluctuations over time?

1. Savings accounts
2. Bonds
3. Stocks

Base: All respondents

KHKNOWL3 [SINGLE SELECT]

When an investor spreads his or her money among different assets, does the risk of losing a lot of money increase, decrease or stay the same?

1. Increase
2. Decrease
3. Stay the same

Base: All respondents

KHKNOWL4 [SINGLE SELECT] [PROMPT]

Do you think the following statement is true or false?

"If you were to invest \$1,000 in a stock mutual fund, it would be possible to have less than \$1,000 when you withdraw your money."

1. True
2. False

[INSERT PAGE BREAK]

Base: All respondents

KHKNOWL5 [SINGLE SELECT] [PROMPT]

Do you think the following statement is true or false?

"'Whole life' insurance has a savings feature while 'term' insurance does not."

1. True
2. False

Base: All respondents

KHKNOWL6 [SINGLE SELECT] [PROMPT]

Do you think the following statement is true or false?

"Housing prices in the US can never go down."

1. True
2. False

Base: All respondents

KHKNOWL7 [SINGLE SELECT] [PROMPT]

Suppose you owe \$3,000 on your credit card. You pay a minimum payment of \$30 each month. At an Annual Percentage Rate of 12% (or 1% per month), how many years would it take to eliminate your credit card debt if you made no additional new charges?

1. Less than 5 years
2. Between 5 and 10 years
3. Between 10 and 15 years
4. Never, you will continue to be in debt

Base: All respondents

KHKNOWL8 [SINGLE SELECT] [PROMPT]

If interest rates rise, what will typically happen to bond prices?

1. They will rise
2. They will fall
3. They will stay the same
4. There is no relationship between bond prices and the interest rate

Base: All respondents

KHKNOWL9 [SINGLE SELECT] [PROMPT]

Do you think the following statement is true or false?

A 15-year mortgage typically requires higher monthly payments than a 30-year mortgage, but the total interest paid over the life of the loan will be less.

1. True
2. False

[INSERT PAGE BREAK]

[DISPLAY]

The next set of questions asks about your financial situation.

Base: All respondents

ENDSMEET [SINGLE SELECT]

In a **typical month**, how difficult is it for you to cover your expenses and pay all your bills?

3. Very difficult
2. Somewhat difficult
1. Not at all difficult

Base: All respondents

HOUSING [SINGLE SELECT]

Which one of the following best describes your housing situation?

1. I own my home
2. I rent
3. I do not currently own or rent

Base: All respondents

LIVINGARRANGEMENT [SINGLE SELECT]

Thinking of the adults in your household, which of the following best describes your current living arrangements?

1. I am the only adult in the household
2. I live with my spouse/partner/significant other
3. I live in my parents' home
4. I live with other family, friends, or roommates
5. Some other arrangement

[INSERT PAGE BREAK]

Base: HOUSING = 1 (I own my home)⁴⁴

HOUSERANGES [SINGLE SELECT]

About how much do you pay for your home each **month**?

1. Less than \$300
 2. \$300-499
 3. \$500-749
 4. \$750-999
 5. \$1,000-1,499
 6. \$1,500-1,999
 7. \$2,000-2,999
 8. \$3,000-4,999
 9. \$5,000 or more
- [NEED SPACE HERE]**
98. I don't know
 99. Prefer not to say

Base: HOUSING = 1 (I own my home)

VALUERANGES [SINGLE SELECT]

If you were to sell your home today, what do you think it would be worth?

1. Less than \$50,000
 2. \$50,000-99,999
 3. \$100,000-149,999
 4. \$150,000-199,999
 5. \$200,000-249,999
 6. \$250,000-299,999
 7. \$300,000-349,999
 8. \$350,000-399,999
 9. \$400,000-449,999
 10. \$450,000-499,999
 11. \$500,000-549,999
 12. \$550,000-599,999
 13. \$600,000-649,999
 14. More than \$650,000
- [NEED SPACE HERE]**
98. I don't know
 99. Prefer not to say

⁴⁴ The base was intended to be all respondents. See discussion on p. 16.

Base: HOUSING = 1 (I own my home)

MORTGAGE [SINGLE SELECT]

What do you owe on your home today?

1. Less than \$50,000
2. \$50,000-99,999
3. \$100,000-149,999
4. \$150,000-199,999
5. \$200,000-249,999
6. \$250,000-299,999
7. \$300,000-349,999
8. \$350,000-399,999
9. \$400,000-449,999
10. \$450,000-499,999
11. \$500,000-549,999
12. \$550,000-599,999
13. \$600,000-649,999
14. More than \$650,000

[NEED SPACE HERE]

98. I don't know
99. Prefer not to say

Base: All respondents

SAVINGSRANGES [SINGLE SELECT]

How much money do you have in savings today (in cash, checking, and savings account balances)?

1. \$0
2. \$1-49
3. \$50-99
4. \$100-249
5. \$250-499
6. \$500-999
7. \$1,000-1,999
8. \$2,000-4,999
9. \$5,000-9,999
10. \$10,000-19,999
11. \$20,000-49,999
12. \$50,000-74,999
13. \$75,000 or more

[NEED SPACE HERE]

98. I don't know
99. Prefer not to say

Base: All respondents

PRODHAVE [MULTIPLE SELECT] [RANDOMIZE]

Which of the following financial products and services do you currently have?

1. Checking or Savings Account at a bank or credit union
2. Life Insurance
3. Health Insurance
4. Retirement Account (such as a 401k or IRA)
5. Pension
6. Non-Retirement Investments (such as stocks, bonds or mutual funds)
7. Education Savings Account (such as 529 or Coverdale)
8. Student/Education Loan (for yourself or someone else)

Base: All respondents

PRODUSE [MULTIPLE SELECT] [RANDOMIZE]

Which of the following, if any, have you used in the **past 12 months**?

1. Payday Loan or Cash Advance Loan
2. Pawn Loan or Auto Title Loan
(An auto title loan is a small loan for a short period of time (usually 30 days) where you give the lender your auto title)
3. A reloadable card that is not linked with a checking or savings account.
(May have logos such as MasterCard, VISA, Discover or American Express. You can keep adding money onto this card and use it to make purchases and pay bills anywhere credit cards are accepted or withdraw the cash from an ATM. This does not include phone cards, gift cards for a particular store or service or cards that you cannot add more funds onto.)
4. A place other than a bank or credit union to give or send money to relatives or friends outside the U.S
5. A place other than a bank or credit union to cash a check or purchase a money order

[INSERT PAGE BREAK]

Base: All respondents

CONSPROTECT1 [SINGLE SELECT]

How often have you had experiences with financial services where you did not feel respected or where you felt mistreated?

4. Often
3. Sometimes
2. Rarely
1. Never

Base: All respondents

CONSPROTECT2 [SINGLE SELECT]

How familiar are you with any agencies or organizations that can help you resolve problems with financial services you are using?

3. Very familiar
2. Somewhat familiar
1. Not at all familiar

Base: All respondents

CONSPROTECT3 [SINGLE SELECT]

Have you ever reported a problem to one of these agencies or organizations?

1. Yes
0. No

[INSERT PAGE BREAK]

Base: All respondents

EARNERS [SINGLE SELECT]

How many people currently contribute to the income used to pay bills for your household?

1. One
2. Two
3. More than two

Base: All respondents

VOLATILITY [SINGLE SELECT]

Which of the following best describes how your household's income changes from **month to month**, if at all?

1. Roughly the same each month
2. Roughly the same most months, but some unusually high or low months during the year
3. Often varies quite a bit from one month to the next

Base: All respondents

SNAP [SINGLE SELECT]

In the **past 12 months**, did you or any member of your household receive benefits from the Food Stamp Program or SNAP (the Supplemental Nutrition Assistance Program)?

1. Yes
0. No
8. Not sure

[INSERT PAGE BREAK]

Base: All respondents

MATHARDSHIP [GRID, SINGLE SELECT PER ROW] [PROMPT]

Please indicate whether each of the following statements were often, sometimes, or never true for you in the **past 12 months**?

Statements in row **[RANDOMIZE]**:

1. I worried whether our food would run out before I got money to buy more.
2. The food that I bought just didn't last and I didn't have money to get more.
3. I couldn't afford a place to live.
4. I or someone in my household needed to see a doctor or go to the hospital but did not go because we couldn't afford it.
5. I or someone in my household stopped taking a medication or took less than directed due to the costs.
6. One or more of my utilities was shut off due to non-payment.

Answers in a column:

3. Often
2. Sometimes
1. Never

Base: All respondents

COLLECT [SINGLE SELECT]

In the past 12 months, have you been contacted by a person or company trying to collect a past-due debt from you? Include instances when you were contacted about debts that you believed you did not owe. **Do not include** instances when the person or company was trying to reach someone else.

1. Yes
0. No
8. Not sure

Base: All respondents

REJECTED [GRID, SINGLE SELECT PER ROW]

In the past 12 months, has either of the following happened to you?

Statements in row:

1. I applied for credit and was turned down.
2. I did not apply for credit because I thought would be turned down.

Answers in a column:

1. Yes
0. No

Base: All respondents

ABSORBSHOCK [SINGLE SELECT]

How confident are you that you could come up with \$2,000 in 30 days if an unexpected need arose within the next month?

4. I am certain I could come up with the full \$2,000
3. I could probably come up with \$2,000
2. I could probably not come up with \$2,000
1. I am certain I could not come up with \$2,000
8. I don't know

[INSERT PAGE BREAK]

Base: All respondents

BENEFITS [SINGLE SELECT] [PROMPT]

For each of the following benefits that employers might offer, please indicate whether you have access to the benefit through your or your spouse's/partner's/other's current or former employer?

Have Access:

Statements in row **[RANDOMIZE]**:

1. Health Insurance
2. 401(k) or Other Employer-Sponsored Retirement Savings Account
3. Defined-Benefit Pension
4. Tuition Reimbursement and/or Student Debt Repayment
5. Work/Life Benefits (such as family leave, emergency dependent care, vacation, etc.)

Answers in a column:

1. Yes
0. No

Base: All respondents

FRAUD2 [SINGLE SELECT]

In the past 5 years, has someone without your permission used or attempted to use an existing account of yours, such as a credit or debit card, checking, savings, telephone, online, or insurance account?

1. Yes
0. No

[NEED SPACE HERE]

8. I don't know

[INSERT PAGE BREAK]

Base: All respondents

COVERCOSTS [SINGLE SELECT]

If you were to find that your income did not cover your living costs, which one of the following would you be most likely to do **first** to make ends meet? **[SINGLE SELECT]**

[RANDOMIZE]:

1. Use savings or sell something you own
2. Cut back or do without
3. Earn more money
4. Borrow money

Base: All respondents

BORROW [GRID, SINGLE SELECT PER ROW]

If you found yourself needing extra money to make ends meet, would the following statements be true for you?

Statements in row **[RANDOMIZE]:**

1. My friends or family would lend me the money and expect me to repay them
2. My friends or family would give me the money with no expectation of repayment

Answers in a column:

1. Yes
0. No

Base: All respondents

SHOCKS [MULTIPLE SELECT] [RANDOMIZE]

In the **past 12 months**, did you or any members of your household experience any of the following?

1. Lost a job
2. Had work hours and/or pay reduced or a business I or someone in my household owned had financial difficulty
3. Received a foreclosure notice
4. Had a major car or home repair
5. Had a health emergency
6. Got a divorce or separation
7. Added a child to the household
8. Experienced the death of primary breadwinner
9. Received a large sum of money beyond normal income (such as inheritance, bonus or other windfall)
10. Had a child start daycare or college
11. Provided unexpected financial support to a family member or friend

Base: All respondents

INCARCERATED [GRID, SINGLE SELECT PER ROW]

Is any member of your household currently or formerly in jail or prison?

Statements in a row:

1. Member of your household **currently** in jail or prison
2. Member of your household **formerly** in jail or prison

Answers in a column:

1. Yes
0. No

[INSERT PAGE BREAK]

Base: All respondents

MANAGE2 [SINGLE SELECT]

Besides regular spending decisions, which of the following options best describes who takes care of the money matters in your household (for example, making investments, paying bills, making decisions)?

1. Someone else takes care of **all or most** money matters in my household.
2. Someone else and I take care of money matters in my household **about the same**.
3. I take care of **all or most** money matters in my household.

[INSERT PAGE BREAK]

Base: MANAGE2 = 1 or MANAGE2 = 2

PAIDHELP [SINGLE SELECT]

Do you pay the person who helps take care of money matters in your household?

1. Yes
0. No

[INSERT PAGE BREAK]

Base: All respondents

HSLOC [DROP DOWN BOX, SINGLE SELECT]

Where did you live on your 17th birthday?

[DROP DOWN LIST OF STATE CODES + DISTRICT OF COLUMBIA + “Other Territories and Possessions” + “Outside the U.S.” + “I can’t recall”]

Base: All respondents

PAREduc [SINGLE SELECT]

What is the **highest level** of education completed by the person or any of the people who raised you?

1. Less than high school
2. High school degree/GED
3. Some college
4. Associates' degree
5. Bachelor's degree
6. Graduate/professional degree

Base: All respondents

FINSOC2 [GRID, SINGLE SELECT PER ROW] [PROMPT]

While growing up at home, did your family do any of the following?

Statements in row **[RANDOMIZE]**:

1. Discussed family financial matters with me.
2. Spoke to me about the importance of saving.
3. Discussed how to establish a good credit rating.
4. Taught me how to be a smart shopper.
5. Taught me that my actions determine my success in life.
6. Provided me with a regular allowance.
7. Provided me with a savings account.

Answers in a column:

1. Yes
0. No

[INSERT PAGE BREAK]

[DISPLAY]

The next questions are about working with numbers. Please answer to the best of your ability.

Base: All respondents

OBJNUMERACY2 [NUMERIC ENTRY, RANGE 0 to 1000]

In the Bingo Lottery, the chance of winning a \$10 prize is 1%. What is your best guess about how many people will win a \$10 prize if 1,000 people each buy a single ticket for the Bingo Lottery?

_____ person(s) out of 1,000

Base: All respondents

OBJNUMERACY1 [SINGLE SELECT]

Which of the following represents the biggest risk of getting a disease?

1. 1%
2. 10%
3. 5%

Base: All respondents

MATERIALISM [GRID, SINGLE SELECT PER ROW]

Thinking about yourself in everyday life, please indicate the extent to which you agree or disagree with each of the following statements.

Statements in row **[RANDOMIZE]**:

1. I admire people who own expensive homes, cars and clothes.
2. The things I own say a lot about how well I'm doing in life.
3. I like to own things that impress people.

Answers in a column:

5. Strongly agree
4. Agree
3. Neither agree nor disagree
2. Disagree
1. Strongly disagree

Base: All respondents

CONNECT [NUMERIC ENTRY, RANGE 0 to 100]

Please think about the important characteristics that make **you** the person **you are now** – your personality, temperament, major likes and dislikes, beliefs, values, ambitions, life goals, and ideals – and please rate the degree of connectedness between the person you expect to be in **5 years** compared to the person you are now, where **0 means “I will be completely different in the future”** and **100 means “I will be exactly the same in the future.”**

My rating is _____

[INSERT PAGE BREAK]

Base: All respondents

HEALTH [SINGLE SELECT]

In general, would you say your health is . . .

5. Excellent
4. Very good
3. Good
2. Fair
1. Poor

Base: All respondents

SCFHORIZON [SINGLE SELECT]

In planning your and/or your family's saving and spending, which of the time periods is most important?

1. The next few months
2. The next year
3. The next few years
4. The next 5 to 10 years
5. Longer than 10 years

Base: All respondents

DISCOUNT [SINGLE SELECT]

If you had a choice, would you rather receive . . . ?

1. \$816 now
2. \$860 in three months

Base: All respondents

MEMLOSS [SINGLE SELECT]

During the past 12 months, have you experienced confusion or memory loss that is happening more often or is getting worse?

1. Yes
0. No

Base: All respondents

DISTRESS [SINGLE SELECT]

Please indicate your level of agreement with the following statement:

I have a lot of stress in my life.

5. Strongly agree
4. Agree
3. Neither agree nor disagree
2. Disagree
1. Strongly disagree

Base: All respondents

SELFCONTROL [GRID, SINGLE SELECT PER ROW]

Please select the response that best indicates how well, in general, each of the following statements describes you.

Statements in row **[RANDOMIZE]**:

1. I often act without thinking through all the alternatives.
2. I am good at resisting temptation.
3. I am able to work diligently toward long-term goals.

Answers in a column:

4. Completely well
3. Very well
2. Not very well
1. Not at all

[INSERT PAGE BREAK]

Base: All respondents

OUTLOOK [GRID, SINGLE SELECT PER ROW]

To what extent do you agree or disagree with each of the following statements.

Statements in row **[RANDOMIZE]**:

1. There are services in this area to help me (such as transportation, child care, health care, in-home services, etc.).
2. There are good work opportunities for me, if I choose to take them.

Answers in a column:

5. Strongly agree
4. Somewhat agree
3. Neither agree nor disagree
2. Somewhat disagree
1. Strongly disagree

Base: All respondents

INTERCONNECTIONS [MULTIPLE SELECT] [RANDOMIZE]

Do you seek advice on matters involving money from any of the following types of people or organizations?

1. Parent
2. Spouse /Partner
3. Extended family (e.g., uncle, aunt, cousins, grandparents)
4. Employer
5. Friends/Co-workers
6. Community or faith-based organizations
7. Financial institution
8. Professional advisor, planner or counselor/coach
9. Government

Base: All respondents

PEM [SINGLE SELECT]

To what extent do you agree or disagree with the following statement?

Everyone has a fair chance at moving up the economic ladder.

7. Strongly agree
6. Agree
5. Somewhat agree
4. Neither agree nor disagree
3. Somewhat disagree
2. Disagree
1. Strongly disagree

Base: All respondents

HOUSESAT [SINGLE SELECT]

How satisfied are you with the place you live currently?

4. Very satisfied
3. Somewhat satisfied
2. Not very satisfied
1. Not at all satisfied

[INSERT PAGE BREAK]

Base: IF PPAGE >= 62

SOCSEC1 [SINGLE SELECT]

Have you started receiving Social Security retirement benefits?

1. Yes
0. No

[INSERT PAGE BREAK]

Base: SOCSEC1 = 1

SOCSEC2 [NUMERIC ENTRY, RANGE 18 to 99]

At what age did you begin receiving benefits?

ENTER NUMBER _____

[INSERT PAGE BREAK]

Base: SOCSEC1 = 2 OR Missing OR PPAGE < 62

SOCSEC3 [NUMERIC ENTRY, RANGE 18 to 99]

[IF PPAGE < 62, INSERT: “People can choose to receive Social Security retirement benefits at different ages. At what age do you think you will be likely to start receiving Social Security retirement benefits?”][IF PPAGE >= 62, INSERT “At what age do you plan to start receiving Social Security retirement benefits?”]

ENTER NUMBER _____

[INSERT PAGE BREAK]

[DISPLAY]

The final questions are about you and others in your household.

Base: PPAGE <= 75

LIFEEXPECT [NUMERIC ENTRY, RANGE 0 to 100]

How likely do you believe it is that you will live beyond age 75? Use the scale from zero to 100 to indicate your response. 0 would mean not at all likely and 100 would mean that it is certain. You can choose any number between 0 and 100.

ENTER NUMBER _____

[INSERT PAGE BREAK]

Base: All respondents

HHEDUC [SINGLE SELECT]

What is the highest level of education **anyone in your household** (including yourself) has completed?

1. Less than high school
2. High school degree/GED
3. Some college
4. Associates' degree
5. Bachelors' degree
6. Graduate/professional degree

Base: All respondents

KIDS [NUMERIC ENTRY, RANGE 0 to 20]

How many children in each of the following age ranges do you currently support financially either living with you or not?

[PROGRAMMER: RECORD ACTUAL NUMBER FOR EACH AGE RANGE BELOW WITH A DEFAULT OF 0 DISPLAYED.]

I have no children that I support financially **[SINGLE SELECT]**

[SPACE HERE]

Less than 7 years old _____
7 to 12 years old _____
13 to 17 years old _____
More than 18 years old _____

Base: All respondents

EMPLOY1 [MULTIPLE SELECT]

Which of the following describe(s) your current employment or work status? (Check all that apply)

1. Self-employed
2. Work full-time for an employer or the military
3. Work part-time for an employer or the military
4. Homemaker
5. Full-time student
6. Permanently sick, disabled or unable to work
7. Unemployed or temporarily laid off
8. Retired

[INSERT PAGE BREAK]

Base: IF MORE THAN ONE SELECTIONS IN EMPLOY1

EMPLOY2 [MULTIPLE SELECT]

Which **one of** the following would you say is your **primary** employment or work status?

[PIPE IN RESPONSES SELECTED IN EMPLOY1]

[INSERT PAGE BREAK]

Base: EMPLOY1 = 8 (Retired)

RETIRE [SINGLE SELECT]

Did you retire earlier than you planned, later than you planned, or about when you planned?

1. Earlier than planned
2. About when planned
3. Later than planned

[INSERT PAGE BREAK]

Base: All respondents

MILITARY [SINGLE SELECT]

Are you a current or former member of the United States Armed Forces, or the spouse or dependent of a service member or veteran?

1. I am a service member or veteran
2. I am a dependent or spouse of a service member or veteran

3. Both
4. Neither

[INSERT PAGE BREAK]

Base: MILITARY = 1 or MILITARY = 3

MILITARYSTATUS [MULTIPLE SELECT]

What is your status? (Check all that apply)

1. Active
2. Reserve
3. National Guard
4. Retired
5. Veteran

[INSERT PAGE BREAK]

Base: MILITARY = 2 or MILITARY = 3

SPOUSEMILSTATUS [MULTIPLE SELECT]

What is your service member's status? (Check all that apply)

1. Active
2. Reserve
3. National Guard
4. Retired
5. Veteran

[INSERT PAGE BREAK]

THANK YOU FOR COMPLETING OUR SURVEY!