

Making Ends Meet Survey Public Use File User Guide

1700 G Street NW
Washington, D.C. 20552
(855) 411-2372



Consumer Financial
Protection Bureau

Table of contents

Table of contents.....	1
1. Introduction.....	2
2. Protocol/methodology.....	3
2.1 Sampling.....	3
2.2 Contacting respondents.....	4
2.3 Response Rates & Sample Sizes	5
2.4 Weighting.....	7
3. Public Use File Data Structure.....	9
3.1 Survey Variables	9
3.2 Derived Credit Variables	9
3.3 Missing Data	10
4. Privacy Protections	11
4.1 Protections during data collection	11
4.2 Protections implemented prior to publication	11
5. Acceptable Use.....	14

1. Introduction

The Consumer Financial Protection Bureau’s (CFPB’s) Making Ends Meet (MEM) survey program is a general financial survey conducted regularly by the Office of Research at the CFPB. The surveys cover financial decision-making, perceptions, expectations, well-being, marketing and usage of different consumer finance products. This public use file (PUF) consists of surveys from the initial Making Ends Meet sample in 2019 (Sample 1) as well as surveys from all samples from 2022 forward (Sample 3 and later).¹

The CFPB has released information and research based on these surveys and the related credit panel data for the last several years and is now making the data from the surveys public.² These public use files also include select credit panel variables for the respondents to the 2019 survey and its follow-up surveys in 2020 and 2021. (Later surveys do not have credit variables.) These data will allow researchers to compare consumers’ financial status along several dimensions, including financial wellbeing, difficulty paying bills and expenses, and ability to cover expenses following income loss.

¹ The inaugural Making Ends Meet survey occurred as a standalone study and was converted into a yearly rolling panel survey beginning in 2022. Sample 2 was a small, specialized sample and is omitted from this release.

² The first research brief presenting the results from the Making Ends Meet survey was published in July 2020 for the data collected in 2019, referred to as “Wave 1”, prior to the COVID-19 pandemic. For more information on the insights from this wave of the survey, refer to <https://www.consumerfinance.gov/data-research/research-reports/insights-making-ends-meet-survey/>. The next iteration of the data, Wave 2, was released in April 2021 and focused on the financial status of consumers during the early stages of the pandemic. The report can be found here: <https://www.consumerfinance.gov/data-research/research-reports/changes-in-consumer-financial-status-during-early-months-pandemic/>. The third wave of the survey was published later in 2021 and expanded research on consumers’ finances throughout the pandemic including the effects of pandemic-related policies such as the student loan flexibility and eviction moratoria. For additional information, please refer to <https://www.consumerfinance.gov/data-research/research-reports/consumer-finances-during-pandemic-insights-making-ends-meet-survey/>.

2. Protocol/methodology

2.1 Sampling

Samples 1, 3, and 4 were drawn from the CFPB's Consumer Credit Panel (CCP),³ a comprehensive, national 1-in-48 sample of credit records maintained by one of the three nationwide consumer reporting agencies (NCRAs).⁴ The CCP is generally representative of U.S. adults aged 18 and older who have a credit record with a NCRA. However, it does not include consumers without a credit record, which limits its coverage.

Samples 5 and later were drawn from the Consumer Credit Information Panel (CCIP), a 1-in-50 sample of credit records maintained by the same NCRA and the successor to the CCP. These samples are generally subject to the same limitations.

The MEM surveys oversample certain target populations. Their association with the CCP/CCIP strengthens the survey by allowing this oversampling. Table 1 shows the characteristics of each oversampled group. As discussed below, the final survey weights are adjusted to make sure they still reflect the overall population accurately.

Note that Sample 2 was a small, specialized sample fielded in 2021 and is not included in this public release.

³ The CCP has since been superseded by the new Consumer Credit Information Panel (CCIP). More recent Making Ends Meet Surveys are drawn from the CCIP.

⁴ The CCP/CCIP excludes any information that might reveal consumers' identities, such as names, addresses, and Social Security numbers. For more information on the privacy protections associated with this survey, see the Consumer Experience Research Privacy Impact Assessment (available: http://files.consumerfinance.gov/f/201406_cfpb_consumer-experience-research_pia.pdf; and System of Records Notice CFPB.022, Market and Consumer Research Records (available: <http://www.consumerfinance.gov/privacy/system-records-notices/market-and-consumer-research-records-2/>).

Table 1: Oversampling*

Sample Number	Oversample description
Sample 1	Individuals with low credit scores, with recent credit delinquencies, those living in rural areas or on Native American trusts, tribal lands, or reservations.
Sample 3	Individuals with recent collections, 60-day delinquencies, or very poor or poor credit scores; individuals living in majority African American or Hispanic areas; individuals living in below median-income areas; individuals who were likely to be African American or Hispanic as predicted by the NCRA; and individuals who were likely to have low incomes, as predicted by the NCRA.
Sample 4	Same as Sample 3, with the addition of individuals with recent <i>medical</i> collections
Sample 5	Individuals living in majority African American or Hispanic areas, individuals living in below median-income areas, individuals with recent medical collections, individuals with any listed medical collection, individuals with 60-day delinquencies, individuals with recent collections, individuals with very poor or poor credit scores.
Sample 6	Individuals living in majority African American or Hispanic areas, individuals living in below median-income areas, individuals with recent inquiries, individuals with recent disputes, individuals with 60-day delinquencies, individuals with very poor, poor, or fair credit scores.

* Sample 2 was a small, specialized sample and is omitted from the public use data.

2.2 Contacting respondents

The Making Ends Meet Survey is administered through the NCRA that manages the CCP/CCIP. Once a sample is drawn from the CCP/CCIP, the NCRA associates the sample to its data through a process that preserves the confidentiality of consumers in the survey sample, survey responses, and credit record information.

The NCRA partners with a subcontractor who specializes in government survey administration. It is this subcontractor who mails the survey packages to sample members. Table 2 shows when each survey was sent and when processing closed.

The surveys return to the same consumers over time – referred to as a different ‘wave’ of each sample – to see how the same consumer adjusts to evolving economic conditions. The follow-up survey and the new survey are in the field at the same time each year.

Table 2: Survey field dates*

Sample Number	Survey Wave	First mailing date	Survey close date
Sample 1	Wave 1	May 22, 2019	August 16, 2019
	Wave 2	May 7, 2020	August 7, 2020
	Wave 3	February 8, 2021	May 7, 2021
Sample 3	Wave 1	January 10, 2022	April 8, 2022
	Wave 2	January 17, 2023	April 14, 2023
Sample 4	Wave 1	January 17, 2023	April 14, 2023
	Wave 2	January 8, 2024	April 5, 2024
Sample 5	Wave 1	January 8, 2024	April 5, 2024
	Wave 2	January 13, 2025	March 28, 2025
Sample 6	Wave 1	January 13, 2025	March 28, 2025

* Beginning in 2022, the survey was fielded at the same time each year. Sample 2 was a small, specialized sample and is omitted from the public use data.

2.3 Response Rates & Sample Sizes

The Making Ends Meet Survey is voluntary, and participants are provided incentives to participate according to best practices in survey methodology and the results of internal experiments. Prior to Sample 4, survey recipients were only provided a pre-completion

incentive. Following an experiment in Sample 4, respondents were provided both a pre- and post-completion incentive that was paid after returning the survey.⁵

Table 3 shows sample sizes and American Association for Public Opinion Research (AAPOR) response rates for each sample and wave.

Table 3: Samples sizes and response rates*

Sample Number	Survey Wave	No. of complete responses	AAPOR Response Rate #1 (%)	AAPOR Response Rate #2 (%)	AAPOR Response Rate #3 (%)
Sample 1	Wave 1	2,986	19.9	20.4	22.0
	Wave 2	1,871	63.7	65.2	64.3
	Wave 3 ⁶	1,749	59.2	60.6	61.1
Sample 3	Wave 1	2,125	12.6	13.2	14.2
	Wave 2	1,062	50.5	53.8	52.2
Sample 4	Wave 1	2,136	15.3	15.8	17.0
	Wave 2	1,373	67.0	67.8	69.3
Sample 5	Wave 1	3,113	18.9	20.0	20.8
	Wave 2	1,952	67.0	67.1	68.8
Sample 6	Wave 1	2,630	16.4	18.2	18.2

* Beginning in 2022, the survey was fielded at the same time each year. Sample 2 was a small, specialized sample and is omitted from the public use data

⁵ See the annual reports for details on these incentives and the experiments conducted to improve efficiency: <https://www.consumerfinance.gov/data-research/research-reports/making-ends-meet-in-2023-insights-from-the-making-ends-meet-survey/>.

⁶ The consumers that responded to the third wave of the survey are not necessarily the same as in Wave 2. In total, 1,399 consumers responded to at least the first questions in all three waves.

2.4 Weighting

2.4.1 General weighting procedure

Each file of survey data includes weights to account for non-response and oversampling. The association with the CCP/CCIP also allows the surveys to adjust for non-response far more comprehensively and exactly than is possible in most other surveys. Most surveys observe almost nothing about non-respondents; accordingly, non-response adjustment reweights the survey so that demographic characteristics of respondents roughly match some external source.

Because of the CCP/CCIP association, we observe the credit characteristics of both respondents and non-respondents, such that we can adjust for non-response at the individual level. For example, consumers with lower credit scores are less likely to respond than consumers with higher credit scores, and the survey weights adjust for this individually. For follow up surveys, the ability to adjust for attrition between waves using not just Wave 1 variables, but also observable changes in the CCP/CCIP between waves, is another key advantage of the survey and makes the survey results generally reflect the range of consumers' experiences since Wave 1.

The final survey weights combine sampling weights (to account for the fact that certain sets of credit records were sampled at higher rates than others) and nonresponse adjustment weights (to account for systematic differences in response rates). We model non-response based on: (1) observables used to draw the initial sample which typically include credit score, recent delinquencies, and recent collections; and (2) predictive variables from the CCP/CCIP, including change in credit score over the previous year and age; and (3) an indicator for whether any of the variables used for weighting was missing and imputed. Finally, for some surveys we applied a raking step to better match the American Community Survey for adults 18 and over. After calculating initial weights, we applied a weight smoothing method to reduce the influence of very large weights. Excessive weight variation can lead to instability of estimates and large estimate variances because some individual records receive far more weight than others. We employed a weight smoothing method to average weights within quintiles of cells defined by our sampling cells.⁷

2.4.2 Available weights

Different weights are appropriate for different analyses. For example, analyses of financial conditions in 2019 might use Wave 1 weights and the larger Wave 1 sample, but a panel

⁷ The method we use is a generalization of: Jean-Francois Beaumont, "A new approach to weighting and inference in sample surveys" *Biometrika*, 2008, 95(3): 539–553; doi: 10.1093/biomet/asn028.

comparison over the three waves should use Wave 3 weights which account for attrition between Wave 1 and Wave 3. Not adjusting for attrition would mean that the results would no longer be representative of the CCP/CCIP and so generally of people with a credit report.

In addition, data files starting with Sample 3 Wave 2 have a ‘combined’ weight that allows researchers to improve statistical power by combining the follow-up survey data with the new sample fielded at the same time (See Section 2.2).

For Sample 5 Wave 2 and Sample 6 Wave 1, several questions were included in an online-only version of the survey. Analyses of these questions should use weights with the ‘_online’ suffix. These questions are marked in the codebook. Questions not marked as online questions should use the weights with the ‘_paper’ suffix (note that these questions were *also* included for respondents who took the survey online).

3. Public Use File Data Structure

The PUF consists of select survey responses and derived variables from credit report information for all complete responses to Sample 1, along with survey responses to Samples 3-6. The credit-derived variables include approximations of credit card debt, limits, VantageScore and whether a consumer has a mortgage, auto or student loan.

For each sample the PUF survey data for all waves are combined into a single file. Sample 1 also contains the credit-derived variables. Each file contains a unique respondent identification number that allows researchers to trace the same respondent's survey responses year to year and combine it with their credit information.

Users can access these data both as a CSV file or as a DTA file. The DTA file includes variable and value labels that correspond with the associated survey question. A detailed list of data fields, along with variable and value labels is provided in the codebooks.

The CSV file, which is compatible with most statistical analysis software, does not include the labels. Users should refer to the original survey instruments and/or the codebooks to understand which variable is associated with each survey question.

3.1 Survey Variables

The survey fields included in the public use files consist of questions related to household finances and demographics, collected over different time periods from the same respondents. Some survey fields are suppressed to reduce reidentification risk as described in Table 4.

3.2 Derived Credit Variables

Sample 1 includes a selection of variables derived from respondent's credit records. As discussed in Section 2.1, the credit panel file is derived from the CCP.

There are four binary indicator variables that indicate whether a consumer has a mortgage, student loan, auto loan and/or credit card on their credit report in a given quarter. Variables starting with an "I" indicate whether that respondent had a tradeline for that debt in that period: Imortgagedebt_, Istudentdebt_, and Iautodebt_ indicate having positive mortgage debt,

student loan debt, and auto debt, while IopenCC_ indicates the respondent has at least one open credit card account.

In addition, there are three continuous credit-derived variables including: a consumer's dollar amount of credit card debt; a consumer's credit card utilization; and a consumer's VantageScore. We have added statistical noise to these latter three variables to reduce the risk of reidentification and protect the privacy of survey respondents.

3.3 Missing Data

These data include missing data for different questions. If a respondent has a missing value for a given question, it can mean any of the following:

- The respondent was not asked this question due to skip logic.
- Respondent's answer was recoded because they followed skip logic incorrectly.
- Respondent was asked this question but did not respond to it.

Users should note that select-all-that-apply questions have no missing values. Each response is marked as either "Selected" or "Not Selected" unless the question itself is embedded in skip logic. If a response is coded as "Not Selected" for all options of a select-all-that-apply question, it is possible that the respondent skipped the question *or* that they believed that none of the options applied to them.

4. Privacy Protections

4.1 Protections during data collection

To limit data sharing only to the minimum required to conduct the survey, the NCRA provides only names and mailing addresses for sample members to the survey subcontractor so that they can send sample members the survey. The survey subcontractor is required to delete this file after all mailings are complete.

4.2 Protections implemented prior to publication

To protect respondents' privacy, The CFPB's contractors remove any direct Personally Identifiable Information (PII) from survey data before they are delivered to the CFPB.

To further mitigate reidentification risk, CFPB researchers have removed, combined, or masked several sensitive fields from the survey and associated CCP/CCIP data. Table 4 shows which survey fields were removed. Numeric fields may have been perturbed with random noise, topcoded, swapped, or otherwise subjected to privacy-preserving manipulation. The PUF does not include geographic identifiers.

To ensure that these techniques sufficiently reduce reidentification risk while maintaining usability of the data, the CFPB iteratively tested the Public Use File with increasing levels of privacy protection and selected mitigation methods that balanced reidentification risk based on this testing with the data's public value.

Table 4: Privacy protections applied to survey variables.*

Description of privacy preservation	Sample number	Wave number	Question number*
Removed open-ended text responses from 'Other' option; combined certain response options with 'other' categories.	Sample 1	Wave 1	8, 51, 64, 67, 68, 69
		Wave 2	18, 21, 22, 23
	Sample 3	Wave 1	61
		Wave 2	44

Description of privacy preservation	Sample number	Wave number	Question number*
	Sample 4	Wave 1	55
	Sample 5	Wave 2	30, 44, 47, 49, 63, 66, 69, 70
	Sample 6	Wave 1	14, 18, 43, 45, 62, 68, 91, 92, 102, 103
Topcoded	Sample 1	Wave 1	65
		Wave 2	19
Topcoded and rounded to nearest 100	Sample 3	Wave 2	5
	Sample 4	Wave 1	25
		Wave 2	57, 65
	Sample 5	Wave 1	21, 26, 34
		Wave 2	11, 13, 17
	Sample 6	Wave 1	24, 27, 31
Dropped response options for marriage, death, moving to a new residence, eviction, natural disasters, and incarceration.	Sample 1	Wave 1	71
		Wave 2	11
		Wave 3	15
	Sample 3	Wave 1	23, 24
		Wave 2	29
	Sample 4	Wave 1	32, 34
		Wave 2	37, 39, 63
	Sample 5	Wave 1	32, 39
		Wave 2	15, 21, 22
	Sample 6	Wave 1	29, 36
Recoded/combined variables into a smaller number of groups or new variable	Sample 1	Wave 1	87, 88, 93, 95
		Wave 2	18, 21
		Wave 3	5, 31
	Sample 3	Wave 1	105, 106, 109, 111

Description of privacy preservation	Sample number	Wave number	Question number*
	Sample 4	Wave 1	96, 97, 102, 103
	Sample 5	Wave 1	91, 92, 97, 98
	Sample 6	Wave 1	109, 114, 115
Dropped questions from dataset	Sample 1	Wave 2	28
	Sample 3	Wave 1	31, 112
	Sample 4	Wave 1	98, 104
	Sample 5	Wave 1	93, 99
	Sample 6	Wave 1	110, 115

* Question numbers for Sample 6 and Sample 5 Wave 2 correspond to the questions as they appear in the codebook, as some questions in those surveys were only included in the online version of the survey.

5. Acceptable Use

The Terms and Conditions of the MEM Public Use File contain a full list of prohibited uses of the data, as well as instructions on proper citation.

All users must abide by these Terms and Conditions.