THE PEW CHARITABLE TRUSTS STUDENT LOAN OMNIBUS SURVEY
METHODOLOGY REPORT

Prepared by:

Robyn Rapoport

Sarah Glancey

SEPTEMBER 2020

Table of Contents

OVERVIEW

This methodology report addresses The Pew Charitable Trusts' (PCT) Student Loan Omnibus survey. This survey was conducted via landline and mobile telephone with a nationally representative sample of respondents through the SSRS Omnibus. For this insert, PCT was interested in surveying a sample of student loan borrowers and nonborrowers (i.e., those who do not have anyone living in their household with student loans).

During the four weeks PCT's survey was included in the SSRS Omnibus, all respondents 18 or older were asked if they currently have student loans, or if anyone living in the household currently has student loans (LOAN1). A total of 4,022 interviews were asked LOAN1 on the SSRS Omnibus Aug. 4-23, 2020, and Sept. 1-6, 2020. Approximately 70% of the interviews asked LOAN1 were conducted via mobile phone (n=2,812) and 30% (n=1,210) asked LOAN1 were completed via landline. If a respondent said either they or someone in their household has student loans, they were classified as a "borrower." If no one in the household currently has student loans, the respondent was considered a "nonborrower." Throughout the entire field period, each borrower was asked the entire insert (i.e., LOAN1 through LOAN17). For the first three weeks in field, we completed interviews with a random 4 out of every 10 respondents 18 or older who were classified as nonborrowers. During the fourth and final week of data collection, only borrowers screened in and were asked the entire survey insert (i.e., no nonborrowers were selected during this time).

In total, N=4,022 respondents were asked LOAN1. Among those 4,022 respondents, almost one-quarter (21%) were borrowers (N=875) and more than three-quarters (78%) were nonborrowers (N=3,147). Among the nonborrowers, N=956 were randomly chosen to complete the entire survey insert. The final sample size total is n = 1,831. Seventy-two percent (n=1,331) completed the survey via mobile phones and 27% (n=500) completed it via landline.

This report is organized into five sections. The first section provides an overview of the project. The next two sections describe the sample design and data collection and fielding. The final two sections address weighting procedures and project deliverables.

SAMPLING METHODS

Overview

Data was collected through the SSRS Omnibus. Each Omnibus survey consists of a minimum of 1,000 interviews, of which 700 interviews are completed with respondents on their cellphones and at least 35 are conducted in Spanish. Completes are representative of the United States population of adults 18 and older. The SSRS Omnibus uses a fully replicated, stratified, single-stage, random-digit-dialing (RDD) sample of telephone households and randomly generated cellphones. Sample telephone numbers are computer-generated and loaded into online sample files accessed directly by SSRS' computer assisted telephone interviewing (CATI) system. Within each sample household, a single respondent is randomly selected.

Household and Respondent Selection

Within each landline household, a single respondent is selected through the following selection process: First, interviewers ask to speak with the youngest adult male/female at home. The term "male" appears first for a randomly selected 30% of the cases and "female" for the other randomly selected 70%. If there are no men/women at home during that time, interviewers ask to speak with the youngest female/male at home.

Cellphones are treated as individual devices, and the interview may take place outside the respondent's home; therefore, cellphone interviews are conducted with the person answering the phone.

DATA COLLECTION

Overview

The Pew Charitable Trusts developed the survey instrument in collaboration with SSRS. Questionnaire development occurred between July 16 and Aug. 3, 2020, with PCT providing an initial draft and SSRS providing survey feedback. Specifically, when reviewing the instrument, SSRS provided feedback regarding question wording, order, clarity, and other issues pertaining to questionnaire quality. Prior to the field period, SSRS formatted the questionnaire and translated the survey instrument into Spanish. The study was then programmed into SSRS' CATI system. Interviews were conducted Aug. 4-23, 2020, and Sept. 1-6, 2020.

Pretest

SSRS conducted a CATI pretest on July 28, 2020, and July 30, 2020, and achieved a total of 39 pretest completes through the SSRS Omnibus. On July 28, 2020, a random 34 respondents were selected to receive the insert. On July 30, 2020, only borrowers were selected to continue past LOAN1.

Following the pretest, SSRS provided PCT with recordings of completed interviews and a detailed memo including feedback on the overall instrument, average length, and interviewing quality. Based on the pretest feedback, PCT made adjustments to the questionnaire, and the survey was finalized on Aug. 3, 2020.

Programming and Testing

Once the questionnaire was finalized, the study was programmed into SSRS' CATI system. Prior to the field period, extensive checking of the program was conducted to assure skip patterns followed the design of the questionnaire. Also, thousands of cases of random data were run through the program to provide another check of the accuracy of the CATI program.

Completed Interviews

Field Period

Interviews were conducted Aug. 4-23, 2020, and Sept. 1-6, 2020.

Field Procedures

Interviewing for each SSRS Omnibus survey is conducted over a six-day period, typically from Tuesday through Sunday, thus including weekdays and weekends. Sample telephone numbers are displayed directly on each interviewer's screen. Based on the result of the call, the interviewer either begins the questionnaire introduction or dispositions that particular sample number by entering an appropriate call result code. The result of every attempt is stored for each sample telephone number.

Each wave of the SSRS Omnibus is composed of two distinct parts. The first is a series of inserts contracted for by various clients; these inserts may range from a single, closed-ended question to a 20-minute battery of open- and closed-ended questions. The second part of the SSRS Omnibus questionnaire includes standard demographic/classification questions.

The CATI system allows for computer control of questionnaire administration, automatic handling of skip pattern response editing, and range checks. Closed-ended responses are ready for tabulation following completion of the last interview. Each unit in the sample receives as many calls as necessary in order to survey qualified respondents and to fulfill the required number of interviews within each substrata of the samples. Additional callback attempts follow a differential callback schedule (a.m./p.m., alternate days, weekdays/weekends) to ensure the highest completion rate possible.

In an effort to maximize the response rate in the interview phase, respondents are given every opportunity to complete the interview at their convenience. For instance, those refusing to continue at the initiation of or during the course of the interview will be offered the opportunity to be re-contacted at a more convenient time to complete the interview. Nonresponsive numbers, such as no answers, answering machines, and busy signals, receive three call attempts.

A key way to increase responses rates is through the use of refusal conversions. Though all of SSRS' interviewers regularly go through "refusal aversion" training, refusals are still a regular part of survey research. SSRS has a core group of specially trained and highly experienced refusal converters who will redial all initial refusals on this project to attempt to convert them to final completed interviews.

Spanish-Language Interviewing

SSRS utilizes a staff of Spanish-speaking interviewers who offer respondents the option of completing the survey in Spanish or in English. Bilingual interviewers call all respondents identified as Spanish-speaking and ask them to complete the SSRS Omnibus.

Interviewer Briefing/Training

Prior to the start of every Omnibus, interviewers are personally briefed and trained on the issues specific to the inserts. The inserts' overall objectives, specific procedures, and questionnaire content are conveyed to the interviewers. Each question is reviewed, and mock interviews are conducted to ensure that all procedures are being followed correctly. In addition, written "job decisions" are created to serve as a manual and record of how to handle out-of-the-ordinary responses to the questionnaires, thus promoting interviewer consistency over time.

Throughout the week, strict control procedures are maintained. Field personnel and project directors continually monitor the interviewers. Each time an interview is monitored, a supervisor reviews the responses as they are entered into the CATI system to verify the accuracy of the responses recorded by the interviewer.

In addition, evaluation forms are completed to rate interviewers on diction, refusal conversion, verbatim recording of open ends, probing, etc. Interviewers are counseled on their weaknesses and praised on improvements.

Data Processing and Integration

Data was checked using multiple methods. Data was checked after the first night of interviewing and throughout the field period to confirm that skip patterns were correctly followed. In addition, the back-end programmer created a program consisting of instructions derived from the skip patterns designated on the questionnaire to check the data. The program confirmed that data was consistent with the definitions of codes and ranges and matched the appropriate bases of all questions. The SSRS team also reviewed preliminary SPSS files and conducted an independent checking of all created variables to ensure that all variables were accurately constructed.

Data file preparation began soon after the study entered the field. The programmer developed syntax for the created variables requested by Pew. Independently, the SSRS project team checked all SPSS variables to confirm that they were created correctly, had the correct number of cases, and were coded according to specifications. Frequencies and marginals were also run against clean data and reviewed to further verify valid codes and skip patterns.

WEIGHTING

Each SSRS Omnibus insert was weighted to provide nationally representative and projectable estimates of the adult population 18 years of age and older. The weighting process took into account the disproportionate probabilities of household and respondent selection due to the number of separate telephone landlines and cellphones answered by respondents and their households, as well as the probability associated with the random selection of an individual household member. Following application of the above weights, the sample was post-stratified and balanced by key demographics such as age, race, sex, region, and education. The sample was also weighted to reflect the distribution of phone usage in the general population, meaning the proportion of those who are cellphone only, landline only, and mixed users.

- **Probability of Selection (P_{phone}):** A phone number's probability of selection depends on the number of phone numbers selected out of the total sample frame. So for each respondent whose household had a landline phone number, this was calculated as total landline numbers dialed divided by total numbers in the landline frame, and for respondents answering at least one cellphone number, this was conversely calculated as total cellphone numbers divided by total numbers in the cellphone frame.
- **Probability of Respondent Selection (P_{select}):** In households reached by landline, a single respondent was selected. Thus, the probability of selection within a household was inversely related to the number of adults in the household.

- **Total Probability of Selection:** This was calculated as the phone number's probability of selection (by frame), and for landlines, divided by the number of adults in the household. Thus, for each respondent, a probability was calculated for being reached via landline (LL_{prob}) and for being reached via cellphone ($Cell_{prob}$). These calculations are:

$$LL_{prob} = P_{phone}* / P_{select}$$

$$Cell_{prob} = P_{phone}$$

The sample weights derived at this stage are calculated as the inverse of the combined probability of selection, or:

$$1/(LL_{prob} + Cell_{prob} - LL_{prob}*Cell_{prob})$$

- **Post Stratification Iterative Proportional Fitting ("raking"):** With the sample weights applied, the sample underwent the process of iterative proportional fitting (IPF), in which the sample was balanced to match known adult-population parameters based on the most recent March supplement of the U.S. Census Bureau's 2019 Current Population Survey (CPS)[1]. This process of weighting was repeated until the root mean square error for the differences between the sample and the population parameters was zero or near-zero.

The population parameters used for post-stratification are: age (18-29; 30-49; 50-64; 65+) by gender, census region (Northeast, North-Central, South, West), education (less than high school, high school graduate, some college, four-year college, or more) by age, education (same breaks) by gender, race/ethnicity (White non-Hispanic, Black non-Hispanic, Hispanic and born in the U.S., Hispanic and born outside of the U.S.[2], other non-Hispanic), marital status (married/not married), population density (divided into quintiles), and phone usage (cellphone only, landline only, both).

- **Weight truncation ("trimming"):** To ensure the consistency of the population estimates produced week to week by Excel, the weights underwent truncation (or "trimming") so that they did not exceed 4 or fall below 0.25.

Because all respondents 18 and older in the Omnibus were screened using LOAN1, the sum of the weights equals the sample N of 4,022.

Margin of Sampling Error

Weighting procedures increase the variance in the data, with larger weights causing greater variance. Complex survey designs and post data-collection statistical adjustments increase variance estimates and, as a result, the error terms applied in statistical testing. The design effect on the sample as a whole for both weights is 1.28. The margin of error, similarly, is +/- 1.8 percentage points.

---

[1] University of Minnesota Institute for Social Research and Data Innovation, Integrated Public Use Microdata Series, Current Population Survey, accessed Oct. 15, 2019, https://doi.org/10.18128/D030.V6.0.

[2] Because this is meant to address the percent of Spanish speakers in the weighted sample, respondents born in Puerto Rico are included with those born outside of the U.S.

DELIVERABLES

SSRS delivered to The Pew Charitable Trusts (1) final weighted SPSS dataset, (2) two final weighted banners in Microsoft Word format, (3) topline results, (4) final version of the English questionnaire, and (5) final methodology report.

CONTACT

Contact Robyn Rapoport for additional information.

rrapoport@ssrs.com | 484.840.4354 | 🐦 @RobynRapoport

1 Braxton Way
Suite 125
Glen Mills, PA 19342