METHODOLOGY REPORT:

STUDENT LOAN RECONTACT SURVEY

Prepared for The Pew Charitable Trusts

Prepared by:

Robyn Rapoport

Elizabeth Sciupac

FEBRUARY 2023

TABLE OF CONTENTS

OVERVIEW

The Pew Charitable Trusts (PCT) engaged SSRS to conduct the Student Loan Recontact Survey. PCT initiated this survey to better understand how much, if anything, student loan borrowers know about repayment options and potential student loan forgiveness. The PCT team was also interested in understanding how borrowers feel about their financial security now and in the future. This survey recontacted qualified student loan borrowers from the 2021 Student Loan Survey, which was conducted by SSRS on behalf of PCT.

For this recontact survey, PCT was primarily interested in surveying student loan borrowers who are not current students ("borrowers, not current students" in Table 1). This group, also referred to as non-student borrowers, qualified to complete the entire survey. If a borrower indicated they *are* a current student ("borrowers, current students" in Table 1), they qualified for the financial security questions but not the questions about student loan repayment.

The Student Loan Recontact Survey was conducted online via the SSRS Opinion Panel. A total of 909 adults 18 and older participated in the survey, and data collection was conducted from Nov. 17-Nov. 30, 2022. Table 1, below, shows the distribution of completes by each group of interest in the final data.

Table 1: Distribution of Completed Interviews

	Total
Borrowers, not current students	760
Borrowers, current students	149
Total	**909**

This report provides information about the sampling procedures and the methods used to collect, process, and weight data for the Student Loan Recontact Survey.

SSRS PROFILE

SSRS is a full-service survey and market research firm managed by a core of dedicated professionals with advanced degrees in the social sciences. SSRS designs and implements research solutions for complex strategic, tactical, public opinion, and policy issues in the U.S. and more than 40 countries worldwide. The SSRS team specializes in creative problem-solving and informed analysis to meet its clients' research goals. SSRS provides the complete set of analytical, administrative, and management capabilities needed for successful project execution. We partner with clients interested in conducting high-quality research. In the industry, SSRS is renowned for its sophisticated sample designs and its experience with all facets of data collection, including those involving multimodal formats. SSRS also has extensive statistical and analytical capabilities for extracting important insights from the survey data and suggesting strategies based on those insights.

SAMPLE DESIGN

Interviews for the Student Loan Recontact Survey were completed using the SSRS Opinion Panel. A total of 1,227 panelists who qualified as student loan borrowers and completed the 2021 Student Loan Survey in English, regardless of their current student status, were invited to participate in the survey. SSRS Opinion Panel members are recruited randomly based on a nationally representative Address Based Sample (ABS) design (including Hawaii and Alaska). Addresses are randomly sampled by our sister company, Marketing Systems Group (MSG), through the U.S. Postal Service's Computerized Delivery Sequence (CDS), a regularly updated listing of all known addresses in the U.S. For the Opinion Panel, known business addresses are excluded from the sample frame.

The SSRS Opinion Panel is a multi-mode panel. Internet households participate via web while all non-internet households (including those who have internet but are unwilling to take surveys online) participate via phone. For the Student Loan Recontact Survey, all panelists participated via web by self-administering the survey. See the Weighting section of this report for details on the non-internet adjustment that was applied in order for the weighted data to be reflective of the broader population, rather than only internet users.

QUESTIONNAIRE DESIGN

PCT developed the survey instrument in collaboration with SSRS. Questionnaire development occurred between Sept. 29-Oct. 10, 2022, with PCT providing an initial draft and SSRS supplying survey feedback. The SSRS team provided feedback regarding question wording, order, clarity, and other issues pertaining to questionnaire quality. Together, SSRS and the PCT team worked to finalize the questionnaire for pretesting.

Upon final approval, SSRS formatted and programmed the survey for completion online. Additional steps were employed to ensure a quality experience in survey administration regardless of the device utilized by respondents, whether a desktop computer, tablet, or phone.

Pretest

Once the survey was programmed, SSRS completed five cognitive pretest interviews to help identify questions that were confusing or not understood as intended, and to evaluate the usability of the online survey instrument. Upon completion of the pretest interviews, SSRS provided recordings and a detailed memo to PCT that included feedback and suggested revisions to the overall instrument. Following the pretest, adjustments were made to the questionnaire and the survey program and it was prepared for the full launch.

DATA COLLECTION

Survey Sampling

All sample drawn for the Student Loan Recontact Survey were adult web panelists, ages 18 and older, who completed the 2021 Student Loan Survey in English. Panelists were first asked if they borrowed student

loans for themselves or for someone else. Then, they were asked about their current student status. If the panelist borrowed student loans for themselves or someone else and was not a current student, they qualified to complete the entire survey. If a borrower indicated they are a current student, they qualified for the questions about financial security but not the questions about student loan repayment.

Survey Administration Procedures

Surveys conducted using the SSRS Opinion Panel are self-administered web surveys. Panelists were emailed an invitation, which included a unique passcode-embedded link, to complete the survey online. In appreciation for their participation, panelists received a modest incentive (in the form of an electronic gift card). All respondents who did not respond to their first invitation received up to three reminder emails or text reminders.

A "soft launch" inviting a limited number of panelists to participate was conducted Nov. 15-Nov. 17, 2022. After checking soft launch data to ensure that all questionnaire content and skip patterns were correct, an additional sample was released to ensure the final sample met the study goals.

Overall, the median length of the Student Loan Recontact Survey for non-student borrowers was 6.9 minutes; the median length for student loan borrowers who are current students was 2.5 minutes.

For the Student Loan Recontact Survey, the survey administration schedule for panelists was as follows:

Table 2: Fieldwork Schedule

Touchpoint	Date
Soft launch invitation	11/15/2022
Full launch invitation	11/17/2022
Field close	11/30/2022

DATA PROCESSING AND INTEGRATION

SSRS implemented several quality assurance procedures in data file preparation and processing. Prior to launching data collection, extensive testing of the web survey was completed to ensure it was working as anticipated. After the soft launch, survey data were carefully checked for accuracy, completeness, and non-response to specific questions so that any issues could be identified and resolved prior to the full launch.

The data file programmer implemented a "data cleaning" procedure in which web survey skip patterns were created in order to ensure that all questions had the appropriate numbers of cases. This procedure involved a check of raw data by a program that consisted of instructions derived from the skip patterns designated on the questionnaire. The program confirmed that data were 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.

As a standard practice, quality checks were incorporated into the survey. Quality control checks for this study included a review of "speeders," reviewing the internal response rate (number of questions answered divided by the number of questions asked) and any verbatim responses. No cases were eliminated after quality control checks.

WEIGHTING

Weighting is generally used in survey analysis to compensate for sample designs and patterns of nonresponse that might bias results. The weighting ensures that the demographic profile of the sample matches the profile of the target population. The data were weighted by applying a base weight and balancing the demographic profile of the sample to target population parameters.

Base Weight

Recontact interviews are susceptible to nonresponse bias stemming from systematic differences between respondents willing to complete a second interview, and those who are not. Table 3 compares the unweighted demographic distributions of those who were invited to take the survey with those who completed the 2022 survey.

Table 3: Demographic Summary of Invited Sample and Completed Interviews

	Eligible Invited Sample n=1227 Unweighted	Completed Interviews n=909 Unweighted
Gender		
Male	35.1%	36.1%
Female	64.9%	63.9%
Age		
18-29	22.6%	22.1%
30-49	53.6%	54.9%
50-64	18.9%	18.8%
65+	4.8%	4.2%
Education		
HS grad or less	6.4%	5.9%
Some college	28.5%	27.3%
College degree	36.0%	37.1%
Graduate degree	29.1%	29.7%
Race/Ethnicity		
White, not Hispanic	50.0%	51.4%
Black, not Hispanic	21.8%	20.8%
Hispanic	20.7%	20.5%
Other race, not Hispanic	7.4%	7.4%
Census region		
Northeast	17.8%	17.9%
Midwest	22.7%	23.4%
South	39.2%	39.1%
West	20.2%	19.6%
Civic engagement		
Not civically engaged	61.3%	50.2%
Civically engaged	38.7%	49.8%
Density		
1 - Least dense	13.0%	13.6%
2	18.1%	17.8%
3	22.3%	21.3%
4	23.5%	22.4%
5 - Most dense	23.1%	24.8%

We use a propensity adjustment to correct for this attrition. Characteristics of the respondents as measured in the initial survey were used to model their probability of responding to the second survey. Propensity is modeled through logistic regression in which the dependent variable is whether or not sampled

respondents completed the second survey. The predictive values are calculated as the probability of a person completing the survey.

The base weight was calculated as product of the final weight from the original survey and the inverse of the predicted probability of completing the second survey calculated by the logistic regression model. Variables used in this model include demographics (gender, marital status, employment status, age, educational attainment, population density in county of residence, number of adults in the household) and behavioral items such as ever having contacted a politician.

Raking

With the base weight applied, the data were weighted to balance the demographic profile of the sample to the target population parameters.

Missing data in the raking variables were imputed using hot decking. Hot deck imputation replaces the missing values of a respondent randomly with another similar respondent without missing data. Hot decking was done using an SPSS macro detailed in "Goodbye, Listwise Deletion: Presenting Hot Deck Imputation as an Easy and Effective Tool for Handling Missing Data."[1]

Weighting was accomplished using SPSSINC RAKE, an SPSS extension module that simultaneously balances the distributions of all variables using the GENLOG procedure.[2]

Data were weighted to distributions of sex, age, education, race/ethnicity, census region, civic engagement, and population density. The calibration totals were obtained from the full, weighted sample of borrowers from the initial study.

Weights were trimmed at the second and 98th percentiles to prevent individual interviews from having too much influence on survey-derived estimates. The table below compares unweighted and weighted sample distributions to target population benchmarks.

[1] Myers, T. (2011). "Goodbye, Listwise Deletion: Presenting Hot Deck Imputation as an Easy and Effective Tool for Handling Missing Data." *Communication Methods and Measures*. Vol. 5, No. 4, 2011, pp. 297-310.
[2] Peck, J. (2011). "Extension Commands and Rim Weighting with IBM SPSS Statistics: Theory and Practice."

Table 4: Weighting Summary

	Parameter	Unweighted	Weighted
Gender			
Male	39.0%	36.1%	38.5%
Female	61.0%	63.9%	61.5%
Age			
18-29	29.9%	22.1%	29.3%
30-49	48.1%	54.9%	48.7%
50-64	16.5%	18.8%	16.7%
65+	5.4%	4.2%	5.4%
Education			
HS grad or less	15.1%	5.9%	13.6%
Some college	34.4%	27.3%	35.0%
College degree	32.1%	37.1%	32.8%
Graduate degree	18.3%	29.7%	18.6%
Race/Ethnicity			
White, not Hispanic	53.7%	51.4%	54.5%
Black, not Hispanic	20.7%	20.8%	19.9%
Hispanic	16.6%	20.5%	16.6%
Other race, not Hispanic	9.0%	7.4%	8.9%
Census region			
Northeast	16.7%	17.9%	16.9%
Midwest	22.4%	23.4%	22.0%
South	40.3%	39.1%	40.8%
West	20.7%	19.6%	20.2%
Civic engagement			
Not civically engaged	65.0%	50.2%	64.2%
Civically engaged	35.0%	49.8%	35.8%
Density			
1 - Least dense	16.1%	13.6%	16.4%
2	18.6%	17.8%	18.7%
3	22.6%	21.3%	22.3%
4	22.3%	22.4%	22.0%
5 - Most dense	20.5%	24.8%	20.7%

Effects of Sample Design on Statistical Inference

Post-data collection statistical adjustments require analysis procedures that reflect departures from simple random sampling. SSRS calculates the effects of these design features so that an appropriate adjustment can be incorporated into tests of statistical significance when using these data. The so-called "design effect" or *deff* represents the loss in statistical efficiency that results from a disproportionate sample design and systematic nonresponse. The total sample design effect for this survey is 1.90.

SSRS calculates the composite design effect for a sample of size n, with each case having a weight, w, as:[3]

$$deff = \frac{n\sum w^2}{(\sum w)^2}$$

The survey's margin of error is the largest 95% confidence interval for any estimated proportion based on the total sample — the one around 50%. For example, the margin of error for the entire sample is ±4.5 percentage points. This means that in 95 out of every 100 samples drawn using the same methodology, estimated proportions based on the entire sample will be no more than 4.5 percentage points away from their true values in the population. Margins of error for subgroups will be larger. It is important to remember that sampling fluctuations are only one possible source of error in a survey estimate. Other sources, such as respondent selection bias, questionnaire wording, and reporting inaccuracy, may contribute additional error of greater or lesser magnitude.

COOPERATION RATE/RESPONSE RATE

Panel response rates are a product of (1) response rates to the original invitation to participate as a panelist and (2) the response rates, among panelists, to the invitation to participate in the study. The table below details the cooperation and response rates for this study.

Table 5: Cooperation Rate/Response Rate[4]

Sample Productivity	
Invited to participate/total sample	1,227
Completed	909
Removals	0
Terminates	84
Survey cooperation rate	74%

[3] Kish, L. (1992). "Weighting for Unequal Pi." *Journal of Official Statistics*, Vol. 8, No. 2, 1992, pp. 183-200.

[4] The cumulative response rate is calculated to be 4%, using AAPOR's Response Rate 3 formula, which accounts for response rates to initial recruitment.

DELIVERABLES

Final deliverables for this study included a final formatted questionnaire, audio recordings of the cognitive pretest interviews, a memo of the pretest findings, a final weighted SPSS file, three weighted banners (including one comparing 2022 and 2021 data), a topline showing the 2022 results, a topline showing comparisons between 2022 and 2021, and this methodology report.