METHODOLOGY REPORT:
STUDENT LOAN SURVEY

Prepared for The Pew Charitable Trusts

Prepared by:

Robyn Rapoport

Sarah Glancey

JULY 2021

ssrs
research. refined.

TABLE OF CONTENTS

OVERVIEW

The Pew Charitable Trusts (PCT) engaged SSRS to conduct the 2021 Student Loan Survey. PCT initiated this survey to better understand general attitudes and opinions about student loans and loan repayment, including the government's role vis-a-vis student loan borrowers. The PCT team was also interested in understanding how borrowers feel about their financial security now and in the future. Additionally, it hoped to learn how the coronavirus pandemic has affected how students and families pay for college and how, if at all, the pandemic affected borrowers' experiences with repayment and interactions with the repayment system.

PCT was interested in surveying three groups for this research: current students, current student loan borrowers (borrowers), and those who are neither current students nor current student loan borrowers (nonstudents, nonborrowers). Within this, PCT also wanted to maximize completes with African American and Hispanic current students and borrowers.

The 2021 Student Loan Survey was conducted online via the SSRS Opinion Panel and was supplemented with sample from our partner probability panel, IPSOS' KnowledgePanel. A total of 2,806 adults age 18 and older participated in the survey, and data collection was conducted from May 10 to June 16, 2021. Table 1, below, shows the distribution of completes by each group of interest in the final data.

Table 1: Distribution of Completed Interviews

	Total
Nonstudents, nonborrowers	1,023
Total current students	599
African American current students	169
Hispanic current students	202
Total student loan borrowers	1,507
African American student loan borrowers	380
Hispanic student loan borrowers	322
Total	**2,806**[1]

This report provides information about the sampling procedures and the methods used to collect, process, and weight data for the 2021 Student Loan 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 in more than 40 countries worldwide. The

[1] There is overlap across the current students and student loan borrower groups as some respondents were both current students and current student loan borrowers.

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.

QUESTIONNAIRE DESIGN

PCT developed the survey instrument in collaboration with SSRS. Questionnaire development occurred between March 3 and April 12, 2021, 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.

In addition, the questionnaire was translated into Spanish, so respondents were able to complete the survey in English or Spanish.

Pretest

Once the survey was programmed, SSRS completed six 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. After the pretest, adjustments were made to the questionnaire and the survey program, and it was prepared for the full launch.

SAMPLE DESIGN

The majority of interviews for the Student Loan Survey were completed using the SSRS Opinion Panel. Panelists in the SSRS Opinion Panel are recruited randomly based on a nationally representative ABS (Address Based Sample) 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.

Additionally, the SSRS Opinion Panel recruits hard-to-reach demographic groups via our Omnibus survey platform.[2] The SSRS Omnibus survey is a nationally representative (including Hawaii and Alaska) bilingual telephone survey designed to meet standards of quality associated with custom research studies. The SSRS Omnibus completes more than 50,000 surveys annually with 80% cell allocation.

The advantage of this recruiting design is that it relies on a high-quality ABS design that yields a higher response rate. Additionally, it leverages the SSRS Omnibus platform to ensure adequate representation of groups typically underrepresented in public opinion polls, such as Hispanics, African Americans, and lower-educated and lower-income populations.

Additional interviews were completed with sample from our partner probability panel, IPSOS' KnowledgePanel, to increase completes among Hispanic and African American current students. Panelists from KnowledgePanel[3] are recruited randomly based on a nationally representative ABS design (including Hawaii and Alaska) via the latest version of the U.S. Postal Service's Delivery Sequence File (DSF). The Partner Panel uses a stratified random sampling to ensure that the geodemographic composition of our panel members mimic those of the adult population in the U.S.

DATA COLLECTION

Survey Sampling

All sample members drawn for the Student Loan Survey were adult panelists, age 18 and older. Panelists were first asked if they were a current student loan borrower or a current student. Once the goal of n=1,000 nonstudents, nonborrowers was reached, panelists had to confirm they were either a current student or borrower to qualify and complete the survey.

Survey Administration Procedures

Surveys conducted using the SSRS Opinion Panel and our Partner 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 four reminder emails or text reminders.

A "soft launch" inviting a limited number of panelists to participate was conducted May 10. After soft launch data was checked to ensure that all questionnaire content and skip patterns were correct, additional sample was released to ensure that the final sample met the study goals.

Overall, the median length of the Student Loan Survey was 10 minutes.

[2] Before July 2019, the SSRS Opinion Panel was recruited entirely from the SSRS Omnibus.
[3] For additional information, please see: https://www.ipsos.com/sites/default/files/ipsosknowledgepanelmethodology.pdf.

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

Table 2: Fieldwork Schedule

Touchpoint	Date
Soft launch invitation	5/10/2021
Full launch invitation	5/11/2021
Field close	6/15/2021

COMPLETION RATE/RESPONSE RATE

Tables 3 details the completion and response rates for this study.

Table 3: Completion Rate/Response Rate[4]

Sample productivity	
Invited to participate/total sample	11,916
Completed	2,806
Removals	8
Terminates	3,000
Survey completion rate	48%
Composite response rate	2%[5]

DATA PROCESSING AND INTEGRATION

SSRS implemented several quality assurance procedures in data file preparation and processing. Before data collection was launched, extensive testing of the web survey was completed to ensure that it was working as anticipated. After the soft launch, survey data was carefully checked for accuracy, completeness, and nonresponse to specific questions so that any issues could be identified and resolved before 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.

[4] Web-panel response rates are a product of (1) response rates to the original invitation to participate as a panelist; (2) the completion rate, among panelists, with the invitation to participate in the study.
[5] Product of the SSRS Opinion Panel and Partner Panel recruitment response rates and the Student Loan Survey completion rate.

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), open-ended questions, and trap questions. For the Student Loan Survey, the trap questions included asking respondents to select a specific response when viewing a list of items. Respondents who failed the quality checks employed were not included in the final dataset. A total of n=8 cases were removed for quality control.

WEIGHTING

For this research, weighting is used 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 sample was weighted in stages. The first stage of the weighting was the application of a base weight to account for different selection probabilities and response rates across sample strata. In the second stage, sample demographics were post-stratified to match population parameters.

Base Weight

The base weight was calculated differently depending on whether the respondent was contacted through the SSRS Opinion Panel or the Partner Panel.

SSRS Opinion Panel Sample

The panel base weight (PBW) was computed differently depending on whether the panelist was recruited from the SSRS Omnibus or from ABS.

Panelists recruited via the SSRS Omnibus were assigned the original Omnibus base weight.

The Omnibus base weight, $PBW_{Omni,}$ can be expressed as a function of the size of the landline and cellphone sample frames ($F_{LL}\ and\ F_{CELL}$), the size of the landline and cellphone samples ($S_{LL}\ and\ S_{CELL}$), and the number of adults in each household (AD) as follows.[6]

$$PBW_{Omni} = \left((LL \times AD \times S_{LL}/F_{LL}) + (CP \times S_{CELL}/F_{CELL}) - (LL \times AD \times CP \times S_{LL} \times S_{CP}/(F_{LL} \times F_{CP}))\right)^{-1}$$

Where $LL = 1$ if the respondent has a landline phone and $LL = 0$ otherwise and $CP = 1$ if the respondent has a cellphone and $CP = 0$ otherwise.

The base weight for ABS recruits is the product of a sampling weight and a household size adjustment. The sampling weight accounts for selection probabilities of addresses across the 16 ABS strata and also the probability of selection of one adult in each sampled household.

[6] Buskirk, T.D., and Best, J. (2012). Venn Diagrams, Probability 101 and Sampling Weights Computed for Dual Frame Telephone RDD Designs. Journal of Statistics and Mathematics, 15, 3696-3710.

The ABS base weight, PBW_{ABS}, can be expressed as a function of the proportion of the ABS frame in stratum i, P_i, the proportion of the ABS sample that was pulled from stratum i, p_i and the number of adults in household j as follows.

$$PBW_{ABS} = (P_i/p_i) \times AD_{ij}$$

Partner Panel Sample

Interviews from our Partner Panel were assigned the base weight provided by our Partner Panel (IBW) upon completion of data collection.

Combined Base Weight

The base weight (BW) is the sampling adjustments outlined above:

$$BW_i = \begin{cases} PBW_{Omni}, & Panel\ cases\ recruited\ from\ SSRS\ Omnibus \\ PBW_{ABS}, & Panel\ cases\ recruited\ from\ ABS\ sample \\ IBW, & Partner\ Panel\ cases \end{cases}$$

Non-Internet Adjustment (NIA)

This adjustment is a propensity score adjustment to model households with internet access to be representative of all households (regardless of whether they have internet access). Propensity scores were estimated by modeling panel response mode on a range of demographic and attitudinal covariates. The model is a CART[7] (Classification and Regression Trees) decision tree built in SPSS by using its scoring wizard available with the decision tree license. Adjustments for each panel participant are then calculated as the reciprocal of the model estimated propensity to be an internet user.

Final Base Weight

The final base weight (FBW) is the product of the base weight and the non-internet adjustment.

$$FBW = BW \times NIA$$

The final standardized base weight ($FSBW$) was standardized by recruitment source (Omnibus Panel recruit, ABS Panel recruit, Partner Panel) and trimmed.

$$FSBW = \begin{cases} FBW \times n_{OMNI} / \sum_{i \in OMNI} FBW_i, & Panel\ cases\ recruited\ from\ SSRS\ Omnibus \\ FBW \times n_{ABS} / \sum_{i \in ABS} FBW_i, & Panel\ cases\ recruited\ from\ ABS\ sample \\ FBW \times n_{PARTNER} / \sum_{i \in PARTNER} FBW_i, & Partner\ Panel\ cases \end{cases}$$

[7] Practical Tools for Designing and Weighting Survey Samples (2nd ed.) by Richard Valliant, Jill A. Dever, and Frauke Kreuter. Cham, Switzerland: Springer, 2018.

Post-Stratification to Population Benchmarks

The next step in the weighting balanced the demographic profile of the sample to target population parameters.

To handle missing data among some of the self-reported demographic variables, we employ a technique called hot decking. Hot deck imputation replaces the missing values of a respondent randomly with another similar respondent without missing data. These are further determined by variables predictive of nonresponse that are present in the entire file. We use an SPSS macro detailed in "Goodbye, Listwise Deletion: Presenting Hot Deck Imputation as an Easy and Effective Tool for Handing Missing Data" (Myers, 2011).

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

The sample was weighted within race (White/other non-Hispanic; Black non-Hispanic; and Hispanic) to match population estimates. The weighting parameters for Black non-Hispanic and Hispanic were gender, age, education, census region, civic engagement, density, and respondent type (nonstudent, nonborrower; student, nonborrower; nonstudent, borrower; student, borrower). An additional parameter for White/other non-Hispanic was race/ethnicity.

The gender, age, education, race/ethnicity, and census region benchmarks were derived from 2020 Current Population Survey[8] (CPS) data. The density benchmark was derived from the 2020 ACS.[9] The civic engagement benchmark was derived from September 2019 CPS Volunteering and Civic Life Supplement data.[10] The "respondent type" benchmark was extracted from a dataset of 1,403 SSRS Probability Panel respondents weighted to be representative of the general 18-plus U.S. population.

Weights were trimmed at the second and 98th percentiles within race/ethnic group to prevent individual interviews from having too much influence on the final results.

Finally, the weights were rescaled to match the race group proportions in the population.

---

[8] Sarah Flood, Miriam King, Steven Ruggles, and J. Robert Warren. Integrated Public Use Microdata Series, Current Population Survey: Version 7.0. [dataset]. Minneapolis: University of Minnesota, 2019. https://doi.org/10.18128/D030.V5.0.

[9] Steven Ruggles, Sarah Flood, Ronald Goeken, Josiah Grover, Erin Meyer, Jose Pacas and Matthew Sobek. IPUMS USA: Version 10.0 [dataset]. Minneapolis, MN: IPUMS, 2020. https://doi.org/10.18128/D010.V10.0.

[10] Civically engaged respondents are defined as those who have volunteered in the past 12 months or who talk to their neighbors daily. https://www.census.gov/programs-surveys/cps/about/supplemental-surveys.html.

Table 4: Weighting Summary—White/Other, Non-Hispanic

	Parameter	Unweighted	Weighted
Gender			
Male	48.6%	44.6%	48.8%
Female	51.4%	55.4%	51.2%
Age			
18-29	18.4%	18.5%	17.9%
30-49	31.0%	39.7%	30.9%
50-64	25.7%	24.3%	26.0%
65+	24.9%	17.5%	25.3%
Education			
HS grad or less	32.1%	14.1%	31.2%
Some college	27.7%	26.2%	28.0%
College degree	25.2%	33.2%	25.5%
Graduate degree	15.0%	26.4%	15.2%
Race/ethnicity			
White, not Hispanic	87.9%	87.8%	88.0%
Other, not Hispanic	12.1%	12.2%	12.0%
Census region			
Northeast	18.5%	19.6%	18.4%
Midwest	24.1%	26.4%	24.2%
South	34.5%	32.8%	34.6%
West	22.9%	21.3%	22.8%
Civic engagement			
Not civically engaged	63.6%	57.8%	63.4%
Civically engaged	36.4%	42.2%	36.6%
Density			
1—Least dense	22.6%	16.7%	22.6%
2	22.4%	20.6%	22.3%
3	20.6%	22.8%	20.6%
4	19.4%	21.3%	19.3%
5—Most dense	15.0%	18.7%	15.2%
Respondent type			
Nonstudent, nonborrower	78.7%	45.4%	78.4%
Student, nonborrower	2.8%	7.1%	2.9%
Nonstudent, borrower	15.5%	41.1%	15.8%
Student, borrower	2.9%	6.5%	2.9%

Table 5: Weighting Summary—Black, Non-Hispanic

	Parameter	Unweighted	Weighted
Gender			
Male	45.7%	24.6%	39.2%
Female	54.3%	75.4%	60.8%
Age			
18-29	24.1%	20.3%	19.6%
30-49	34.9%	49.5%	37.3%
50-64	24.2%	23.4%	25.5%
65+	16.8%	6.8%	17.5%
Education			
HS grad or less	44.5%	14.6%	38.1%
Some college	30.0%	32.4%	33.7%
College degree	16.7%	30.6%	18.8%
Graduate degree	8.8%	22.4%	9.4%
Census region			
Northeast	15.7%	14.2%	15.7%
Midwest	16.9%	18.7%	18.2%
South	58.4%	58.1%	56.4%
West	9.0%	9.0%	9.7%
Civic engagement			
Not civically engaged	72.9%	63.0%	71.5%
Civically engaged	27.1%	37.0%	28.5%
Density			
1—Least dense	15.5%	9.0%	9.9%
2	14.6%	16.2%	15.1%
3	19.2%	17.2%	21.8%
4	20.8%	26.7%	20.5%
5— Most dense	29.9%	31.0%	32.7%
Respondent type			
Nonstudent, nonborrower	66.1%	15.4%	61.4%
Student, nonborrower	1.3%	10.5%	1.5%
Nonstudent, borrower	24.7%	51.5%	28.1%
Student, borrower	7.9%	22.6%	9.0%

Table 6: Weighting Summary—Hispanics

Values	Parameter	Unweighted	Weighted
Gender			
Male	49.8%	39.4%	49.5%
Female	50.2%	60.6%	50.5%
Age			
18-29	28.1%	35.2%	28.2%
30-49	40.3%	46.1%	40.6%
50-64	20.2%	14.8%	19.7%
65+	11.4%	3.9%	11.5%
Education			
HS grad or less	56.1%	17.7%	54.0%
Some college	25.2%	39.5%	26.5%
College degree	13.1%	26.0%	13.7%
Graduate degree	5.5%	16.8%	5.8%
Census region			
Northeast	13.5%	15.8%	13.9%
Midwest	8.8%	9.5%	8.8%
South	38.6%	38.4%	38.1%
West	39.1%	36.3%	39.2%
Civic engagement			
Not civically engaged	79.4%	67.4%	78.8%
Civically engaged	20.6%	32.6%	21.2%
Density			
1—Lowest	12.0%	12.9%	11.7%
2	15.5%	14.8%	15.6%
3	16.4%	22.1%	17.2%
4	25.8%	24.4%	25.2%
5—Highest	30.3%	25.8%	30.2%
Respondent type			
Nonstudent, nonborrower	74.7%	29.8%	73.4%
Student, nonborrower	5.2%	16.6%	5.5%
Nonstudent, borrower	16.2%	37.6%	17.0%
Student, borrower	3.9%	16.0%	4.1%

Table 7: Weighting Summary—Overall

Values	Parameter	Unweighted	Weighted
Gender			
Male	48.5%	39.8%	47.8%
Female	51.5%	60.2%	52.2%
Age			
18-29	20.7%	22.5%	19.8%
30-49	33.0%	42.9%	33.3%
50-64	24.6%	22.0%	24.9%
65+	21.7%	12.5%	22.0%
Education			
HS grad or less	37.6%	15.0%	35.8%
Some college	27.6%	30.3%	28.5%
College degree	22.1%	31.1%	22.8%
Graduate degree	12.7%	23.6%	13.0%
Race/ethnicity			
White, not Hispanic	62.8%	52.4%	62.8%
Black, not Hispanic	11.9%	18.3%	11.9%
Hispanic	16.7%	22.1%	16.7%
Other race, not Hispanic	8.6%	7.3%	8.6%
Census region			
Northeast	17.3%	17.8%	17.3%
Midwest	20.7%	21.2%	20.9%
South	38.0%	38.6%	37.7%
West	23.9%	22.3%	24.0%
Civic engagement			
Not civically engaged	67.3%	60.9%	66.9%
Civically engaged	32.7%	39.1%	33.1%
Density			
1—Lowest	20.0%	14.4%	19.3%
2	20.0%	18.5%	20.4%
3	20.0%	21.6%	20.2%
4	20.0%	23.0%	20.4%
5—Highest	20.0%	22.5%	19.8%
Respondent type			
Nonstudent, nonborrower		36.5%	75.6%
Student, nonborrower		9.8%	3.1%
Nonstudent, borrower		42.2%	17.4%
Student, borrower		11.5%	3.9%

Margin of Sampling Error

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 this 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 2.69.

SSRS calculates the composite design effect for a sample of size n, with each case having a weight, w, as:[11]

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

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 plus or minus 3.0 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 3.0 percentage points away from their true values in the population. Margins of error for subgroups will be larger.

It is important to remember that the sampling fluctuations captured in the margin of error 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.

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, two weighted banners, a topline, a memo summarizing key findings, and this methodology report.

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