Memorandum **United States Department of Education**
Institute of Education Sciences
National Center for Education Statistics

DATE: August 23, 2024

TO: Beverly Pratt, OMB

THROUGH: Carrie Clarady, OMB Liaison, IES

FROM: Tracy Hunt-White, NCES

SUBJECT: 2023–24 National Postsecondary Student Aid Study (NPSAS:24) Incentive Boost for Key Sample Groups in Later Data Collection Waves Change Request (OMB# 1850-0666 v.39)

The 2023-24 National Postsecondary Student Aid Study (NPSAS:24) is a nationally representative cross-sectional study of how students and their families finance education beyond high school in a given academic year. NPSAS is conducted by the National Center for Education Statistics (NCES) and was first implemented by NCES during the 1986–87 academic year and has been fielded every 3 to 4 years since. This request pertains to the 12th cycle in the NPSAS series being conducted during the 2023–24 academic year. NPSAS:24 will consist of a nationally representative sample of undergraduate and graduate students, and a nationally representative sample of bachelor's degree completers. Subsets of questions in the NPSAS:24 student survey focus on describing aspects of the experience of bachelor's completers in their last year of postsecondary education.

The request is to conduct all activities related to NPSAS:24, including materials and procedures related to the NPSAS:24 student data collection, consisting of abstraction of student data from institutions and a student survey was approved by OMB in December 2023, with updates approved in January 2024 (OMB#1859-0666 v. 36 and 37, respectively) and August 2024 ((OMB#1859-0666 v. 38), and carried over respondent burden, procedures, and materials related to the NPSAS:24 institution sampling, enrollment list collection, and matching to administrative data files as approved by OMB in September 2023 (OMB#1859-0666 v. 35). The NPSAS:24 enrollment list collection from institutions takes place from October 2023 to October 2024, the student records and student survey data collections take place from February 2024 through November 2024.

This request is to share the results of a second logistic regression modeling to determine whether or not a $10 boost offered to historically underrepresented student subgroups will increase their likelihood of participation in the survey. OMB approved a $10 boost incentive for the earlier data collection waves (1-3) on August 13, 2024 (OMB#1859-0666 v. 38). Our findings of this second logistic regression lead us to recommend offering a $10 promised incentive boost during the remaining waves of data collection for cases belonging to the following groups identified in our analyses:

1) *institution level* is less-than-2 years, 2) *institution level* is 2-years but less-than-4 years and *gender* is male, 3) *undergraduates* whose *control* is private for-profit and *gender* is male, 4) *undergraduates* whose *control* is private for-profit and are *veterans*, 5) *undergraduates* whose *control* is private for-profit and *age group* is 40 years of age or older, and 6) *undergraduates* whose *control* is private for-profit and who identify as *Hispanic*. This request would target later data collection waves (4-9). This request does not introduce significant changes to the estimated respondent burden or the costs to the federal government.

Part B: Boost Incentives for Key Sample Member Groups

To identify subgroups for the incentive boost, we estimated a binary logistic regression model predicting the probability of a NPSAS:24 survey response for respondents, partial respondents, and nonrespondents in Waves 1-7 of data collection. This model included nine sample member characteristics as substantive predictors: gender (coded as male, female, other), age (25 or younger, 26 to 39, 40 and older), race (White, races other than White), ethnicity (Hispanic/Latino, not Hispanic/Latino), veteran status (veteran, not a veteran), control of sample member's institution (public, private non-profit, private for-profit), level of sample member's institution (less than 2 year, 2 to less than 4 year, 4 year or higher non-doctoral, 4 year or higher doctoral), whether sample member's course of study is STEM (STEM, not STEM), and undergraduate status (undergraduate, not an undergraduate). These variables were obtained from enrollment lists. In cases where data from the enrollment lists were missing, we replaced missing values with sample members' substantive answers to the NPSAS:24 survey, where available.

The model also included three variables controlling for design features of the survey: the sample member's data collection wave, whether the sample member was assigned for CATI calling, and time of day that reminder emails were sent to the sample member. The pseudo R-square for the final model was 0.22.

We used this model to estimate predicted probabilities of NPSAS:24 survey response for each category of each of our nine sample member characteristics, holding all other variables at their means. Table 1 below displays these predicted probabilities, along with response rates for each subgroup as of August 11, 2024[1].

We identified four subgroups that had lower response rates and/or propensities – *institution level* of less-than-2 years (low response propensity of 0.43 and low response rate of 36.6 percent), *institution level* of 2-years but less-than-4 year institutions (low response propensity of 0.45 and low response rate of 38.6 percent), *control* of private for-profit (low response rate of 38.0 percent), and *undergraduates* (low response propensity of 0.46 and low response rate of 43.2 percent). These groups have historically responded at lower rates across NPSAS surveys and may benefit from a design change in line with the leverage saliency theory, stipulating that "one-fits all" incentive amount is not a good solution to nonresponse error (Groves, Singer and Corning, 2000).[2]

As the *institution level* of 2-year but less-than-4 year institutions, *control* of private for-profit, and *undergraduates* groups have large sample sizes, we conducted further analyses to investigate differences in nonresponse across our other substantive predictors. We began these additional analyses by first estimating a binary logistic regression model predicting the probability of a NPSAS:24 survey response for respondents, partial respondents, and nonrespondents in Waves 1-7 of data collection, subset to sample members whose *institution level* is 2-years but less-than-4[3]. We used

[1] Notably, this date was prior to implementing two data collection changes approved in OMB# 1850-0666 v.38: 1) sample members in Waves 1-3 of data collection receiving a boost offer if identified as eligible for boost, and 2) the start of a contacting materials experiment.
[2] Ibid.
[3] As all sample members in the *institutional level* of two years but less than four years are undergraduates, it was not necessary to explicitly include undergraduate status as a subset.

the same substantive predictors and controls as the previous model, excluding the institution level, the variable driving the subset. The pseudo R-square for the final model was 0.17. We then used this model to estimate predicted probabilities of NPSAS:24 survey response for each category of each of our seven sample member characteristics, holding all other variables at their means. Table 2 below displays these predicted probabilities, along with response rates for each subgroup as of August 11, 2024. This new analysis identified one additional subgroup of interest for sample members whose institutional level is 2-years but less-than-4 year institutions: sample members whose *gender* is male (low response rate of 35.6 percent; 9 percentage points lower than females).

Similarly, as the *control* of private for-profit and *undergraduates* groups have large sample sizes, we again conducted further analyses to investigate differences in nonresponse across our other substantive predictors. We estimated a binary logistic regression model predicting the probability of a NPSAS:24 survey response for respondents, partial respondents, and nonrespondents in Waves 1-7 of data collection, subset to sample members who are *undergraduates* and whose *control* is private for-profit. We used the same substantive predictors and controls as the first model, excluding the variables driving the subset: undergraduate status and control. The pseudo R-square for the final model was 0.15. We then used this model to estimate predicted probabilities of NPSAS:24 survey response for each category of each of our seven sample member characteristics, holding all other variables at their means. Table 3 below displays these predicted probabilities, along with response rates for each subgroup as of August 11, 2024.

This final analysis identified four additional subgroups of interest within undergraduates in the private for-profit control: 1) sample members whose *gender* is male (low response propensity of 0.34 and low response rate of 33.9 percent), 2) sample members who are *veterans* (low response propensity of 0.35 and low response rate of 32.5 percent), 3) sample members whose *age group* is 40 years of age or older (low response propensity of 0.36 and low response rate of 33.2 percent), and 4) sample members who identify as *Hispanic* (low response propensity of 0.36 and low response rate of 37.6 percent).

Differential incentives have been proven successful in bringing in groups of focal importance who were otherwise underrepresented (e.g., Groves, Singer and Corning, 2000; Groves and Heeringa, 2006; Peytcheva, Kirchner and Cooney, 2018).[4,5] Such a strategy was successfully employed in NPSAS:20 when an additional $10 were offered to nonrespondents in three key analyses groups during the last 8 waves of data collection, resulting in an average response rate increase of 17.53 percent across waves relative to the projected response rate under the original design. We therefore recommend offering a $10 promised incentive boost during the remaining waves of data collection for cases belonging to the following groups identified in our analyses: 1) *institution level* is less-than-2 years, 2) *institution level* is 2-years but less-than-4 years and *gender* is male, 3) *undergraduates* whose *control* is private for-profit and *gender* is male, 4) *undergraduates* whose *control* is private for-profit and are *veterans*, 5) *undergraduates* whose *control* is private for-profit and *age group* is 40 years of age

[4] Groves RM, Heeringa SG. Responsive design for household surveys: tools for actively controlling survey errors and costs. *Journal of the Royal Statistical Society Series a-Statistics in Society.*2006;169(3):439-457. doi: DOI 10.1111/j.1467-985X.2006.00423.x.

[5] Peytcheva, E, Kirchner, A., and Cooney, J. 2018. Experimental Comparison of Two Data Collection Protocols for Previous Wave Nonrespondents. Paper presented at the Methodology of Longitudinal Surveys II conference, Essex, U.K.

or older, and 6) *undergraduates* whose *control* is private for-profit and who identify as *Hispanic*. The purpose of this boost is to encourage participation and reduce the potential for nonresponse bias. This would result in an incentive boost for approximately 22,000[6] nonresponding sample members from Waves 4 – 9. For the rest of the sample member characteristics, response rates and propensities were generally similar across subgroups.

Table 1. Response Rates and Mean Predicted Propensities for Selected Subgroups

Sample member characteristic	Predicted Probability of Survey Response	Standard Error	Response Rate as of 8/11/2024
Gender			
Male	0.55	0.00	44.1%
Female	0.45	0.00	50.1%
Other	0.64	0.02	64.3%
Age			
25 or younger	0.50	0.00	44.4%
26 - 39	0.46	0.00	47.9%
40 or older	0.53	0.01	47.4%
Race			
Races other than White	0.50	0.00	46.6%
White	0.48	0.00	47.9%
Ethnicity			
Not Hispanic	0.49	0.00	48.0%
Hispanic	0.49	0.00	45.4%
Veteran Status			
Not a Veteran	0.49	0.00	45.7%
Veteran	0.47	0.01	44.6%
Control of Institution			
Public Institution	0.48	0.00	44.8%
Private non-profit Institution	0.53	0.00	50.4%
Private for-profit Institution	0.49	0.01	38.0%
Institution Level			
Less-than-2-year	0.43	0.02	36.6%
2-year but less-than-4 year	0.45	0.00	38.6%
4-year or higher non-doctorate granting	0.50	0.00	47.8%
4-year or higher doctorate granting	0.52	0.01	50.3%
STEM Status			
Not in a STEM Program	0.49	0.00	44.9%
In a STEM Program	0.51	0.01	50.8%
Undergraduate Status			
Not an Undergraduate	0.58	0.00	53.4%
Undergraduate	0.46	0.00	43.2%

[6] As Wave 9 of data collection has not been fielded, this estimate reflects our assumption that 30 percent of the sample members in Wave 9 will respond to the NPSAS:24 survey before the boost is offered.

Table 2. Response Rates and Mean Predicted Propensities for Selected Subgroups for Sample Members whose Institutional Level is Two Years but Less than Four Years

Sample member characteristic	Predicted Probability of Survey Response	Standard Error	Response Rate as of 8/11/2024
Gender			
Male	0.44	0.01	35.6%
Female	0.37	0.01	44.5%
Other	0.59	0.07	56.3%
Age			
25 or younger	0.41	0.01	38.0%
26 - 39	0.38	0.01	41.4%
40 or older	0.38	0.01	35.9%
Race			
Races other than White	0.42	0.01	41.3%
White	0.36	0.01	38.2%
Ethnicity			
Not Hispanic	0.41	0.01	40.5%
Hispanic	0.38	0.01	39.7%
Veteran Status			
Not a Veteran	0.40	0.00	38.7%
Veteran	0.40	0.02	36.5%
Control of Institution			
Public Institution	0.39	0.00	38.4%
Private non-profit Institution	0.47	0.03	50.9%
Private for-profit Institution	0.42	0.01	39.0%
STEM Status			
Not in a STEM Program	0.40	0.00	38.6%
In a STEM Program	0.62	0.16	71.4%*

*Note that the response rate for this group is high due to a very small sample size in this cell.

Table 3. Response Rates and Mean Predicted Propensities for Selected Subgroups for Undergraduates whose Control is Private For-Profit

Sample member characteristic	Predicted Probability of Survey Response	Standard Error	Response Rate as of 8/11/2024
Gender			
Male	0.34	0.01	33.9%
Female	0.39	0.01	39.7%
Other	0.62	0.16	61.5%
Age			
25 or younger	0.37	0.01	37.1%
26 - 39	0.37	0.01	38.0%
40 or older	0.36	0.02	33.2%
Race			
Races other than White	0.38	0.01	38.1%
White	0.35	0.01	39.1%
Ethnicity			
Not Hispanic	0.38	0.01	39.8%
Hispanic	0.36	0.01	37.6%
Veteran Status			
Not a Veteran	0.37	0.01	37.4%
Veteran	0.35	0.02	32.5%
Institution Level			
Less-than-2-year	0.35	0.02	36.7%
2-year but less-than-4 year	0.38	0.01	39.0%
4-year or higher non-doctorate granting	0.38	0.01	35.6%
4-year or higher doctorate granting	0.12	0.13	16.7%
STEM Status			
Not in a STEM Program	0.37	0.01	36.9%
In a STEM Program	0.41	0.05	38.9%