National Survey of Public Education's Response to COVID-19

METHODOLOGY BRIEF

Technical Supplement

Authors: Cong Ye | Jordan Rickles | Sarah Hodgman | Mike Garet

JUNE 2021

Introduction

In an effort to better understand how public and charter school governing bodies responded to the COVID-19 pandemic in the 2019–20 school year, the American Institutes for Research (AIR) and our partner NORC at the University of Chicago conducted the National Survey of Public Education's Response to COVID-19. This web survey of public school districts focused on a nationally representative sample of 2,536 public school districts across the United States. The same 2,536 public school districts were sent a second round of surveys to gather their responses about the COVID-19 pandemic in the 2020–21 school year. Due to low response rates, the study team decided to survey all 13,277 eligible public school districts 2 weeks before the close of the survey to yield a reasonable sample size for the analysis. The data collection period was January 26 through April 7, 2021.

This methodology brief describes the sampling and weighting methods for this study, along with the final response rate.

Sample Design

The original sample consisted of 2,536 public school districts from a sampling frame constructed using the 2018–19 Common Core of Data Local Education Agency Universe File (see the sample details in this methodology brief). The study team also decided to reach out to all other districts in the population 2 weeks before the close of the survey to yield a reasonable sample size for the analysis. This made the survey a census, where every unit had a selection probability of 1.

Survey Response Rate

A completed survey was defined as having responses to a set of critical survey items. In total, there were 565 completed surveys from the total sample of 13,277 eligible public school districts (4.3%), with 223 from the original sample (8.8%) and 342 from the remaining sample (3.2%). The final survey response rates were calculated using Response Rate 1 from the American Association for Public Opinion Research (2016), where the number of completes is divided by the number of all eligible sample cases. (Note that the denominators might include some potentially ineligible districts for which we did not have information.)

Weighting Adjustments

The purpose of weighting was to enable unbiased estimates from the survey of public school districts. Without using weights in analyses, the estimates would be biased because there were nonresponses.

Base weights. The base weight for the districts was 1 because all districts were included in the sample.

Nonresponse adjustments. Nonresponse bias may occur if the outcome variables correlate with response propensity (i.e., the likelihood of response). However, because information on the outcome variables was not available for nonrespondents, other information on the sampling frame was used to assess and adjust for the nonresponse bias. Available variables included whether the district was in the original sample, district enrollment, percentage of students with limited English proficiency, percentage of students eligible for free or reduced-price lunch, census division, district locale, and whether White students made up more than half of the student population in the district.[1] We built a logistic regression model with selected variables to predict each case's response propensities. The estimated response propensities for all districts were divided into five weighting classes using quintiles, sorting the cases by the predicted response propensity and classifying them into five categories. This process produces weights that are smoother but still effective (Cochran, 1968). For each weighting class, a nonresponse adjustment factor was calculated so that the sum of adjusted weights for the respondents was equal to the sum of the base weights of all eligible cases, as follows:

$$nrw_{ci} = NRF_c * w_{ci}$$

where w_{ci} is the base weight for case *i* in class *c*, and NRF_c is the nonresponse adjusting factor for the case in class *c*.

Weights calibration. After the nonresponse adjustment, we performed raking adjustments to adjust the nonresponse adjusted weights so that the adjusted weights summed up to the marginal totals of district locale and census division. Raking—a poststratification method usually employed to avoid the problem of small cell sizes—is an iterative process in which some weights are adjusted up and some are adjusted down to match the sums of the weights in each category of each variable for the respondents to the marginal totals of each variable in the population. The raked weights were calculated as follows:

$$Rnrw_{ci} = RF_i * nrw_{ci}$$

where nrw_{ci} is the nonresponse adjusted weight in the previous step, and RF_i is the raking factor for each case.

Evaluation of the final weights. As previously discussed, the purpose of weighting is to enable unbiased estimates using the final weights. Conducting analyses without applying the final weights may result in biased estimates. To evaluate the effectiveness of the final weights, we would need to compare the weighted estimates of outcome variables with population values; however, population values are not available for those variables. Instead, we compared the weighted and unweighted estimates for variables available on the sampling frame with the population (frame) values.

---

[1] Some of the variables have missing data. Those missing cases were imputed using the hot deck imputation procedure.

As shown in Table 1, a few statistically significant differences existed between the responding sample and the population but there were no remaining statistically significant differences after applying the weights in the analysis. The comparison results suggest that the weight adjustments were effective in reducing the differences in the frame variables between the population values and the estimates from the sample.

Table 1. Comparison of Weighted and Unweighted Estimates of Variables Available on the Sampling Frame With the Population (Frame) Values

Variable	Population Value[a] (N= 13,277)	Unweighted (N= 565)		Weighted (N= 565)	
		Estimate	Standard Error	Estimate	Standard Error
In the original sample (%)	19.1	39.5*	2.1	21.8	1.6
Mean enrollment	3539.8	6017.2*	889.9	3918.4	475.9
English learners (%)	6.5	7.1	0.5	6.8	0.5
Students eligible for free or reduced-price lunch (%)	48.3	49.5	1.0	48.1	1.1
Minority students (%)	30.8	33.1	1.2	31.0	1.3
Poverty (%)	16.3	16.4	0.4	16.1	0.5
Historical average achievement[b] (x 100)	2.5	2.2	1.4	2.7	1.7
District locale (%)					
City	5.9	7.4	1.1	5.9	1.1
Suburban	23.2	23.9	1.8	23.2	2.1
Town	18.0	18.9	1.6	18.0	1.9
Rural	52.9	49.7	2.1	52.9	2.5
Census division[c] (%)					
New England	7.8	4.1*	0.8	7.8	1.7
Middle Atlantic	13.0	6.0*	1.0	13.0	2.2
East North Central	20.8	30.8*	1.9	20.8	1.6
West North Central	15.3	14.7	1.5	15.3	1.7
South Atlantic	5.1	9.2*	1.2	5.1	0.8
East South Central	4.5	3.7	0.8	4.5	1.2
West South Central	13.9	12.0	1.4	13.9	1.7
Mountain	8.3	8.5	1.2	8.3	1.3
Pacific	11.4	15.0*	0.6	12.4	0.9

[a] The population value includes some potentially ineligible districts. However, the percentage of ineligible districts that remain in the sampling frame is expected to be low and should not significantly affect the comparisons.

[b] Historical average achievement data were retrieved from the Stanford Education Data Archive (Reardon et al., 2019). The historical achievement measure is based on the district's average performance on English language arts and mathematics state tests for Grades 3–8 across the years 2008–09 through 2015–16. Scores were placed on a common standardized scale so that they were in standard deviations of the national grade- and subject-specific performance of the spring 2009 fourth-grade cohort. Average scores were multiplied by 100 for reporting in the table.

[c] Regional divisions were retrieved from the U.S. Census Bureau (n.d.).

* The difference from the population value is statistically significant at the 5% level.

References

American Association for Public Opinion Research. (2016). *Standard definitions: Final dispositions of case codes and outcome rates for surveys* (9th ed.). https://www.aapor.org/AAPOR_Main/media/publications/Standard-Definitions20169theditionfinal.pdf

Cochran, W. G. (1968). Removing bias in observational studies. *Biometrics, 24*, 295–313.

Reardon, S. F., Ho, A. D., Shear, B. R., Fahle, E. M., Kalogrides, D., Jang, H., Chavez, B., Buontempo, J., & DiSalvo, R. (2019). Stanford Education Data Archive (Version 3.0). http://purl.stanford.edu/db586ns4974.

U.S. Census Bureau. (n.d.). *Census regions and divisions of the United States*. https://www2.census.gov/geo/pdfs/maps-data/maps/reference/us_regdiv.pdf

AIR®

1400 Crystal Drive, 10th Floor
Arlington, VA 22202-3289
+1.202.403.5000 | **AIR.ORG**

15197_06/21