

Using incentives to reduce nonresponse bias in the American Housing Survey



Targeting incentives increased the response rates but did not decrease nonresponse bias

Key findings

Targeting unconditional cash incentives to those most at risk of nonresponse to a statistical survey significantly increased the response rate compared to allocating the same amount of incentives completely at random. However, respondents in the targeted and non-targeted groups had very similar survey responses, suggesting targeting did not meaningfully decrease nonresponse bias.

Agency priority

The American Housing Survey (AHS) is a biannual, longitudinal survey of housing units designed by the U.S. Department of Housing and Urban Development (HUD) and administered by the U.S. Census Bureau. As with many statistical surveys, the AHS has experienced declining response rates¹ requiring increasing time and effort to reach the 80% response rate recommended by the Office of Management and Budget.² Declining response rates raise concerns about *nonresponse bias* – divergence between a sample estimate and its true value in the population created by systematic differences between those who do and do not respond to a survey. For example, if nonrespondents are more likely to live in inadequate housing than respondents, an uncorrected sample estimate of housing inadequacy would underestimate the true proportion of adults living in inadequate housing. The AHS adjusts estimates to account for nonresponse bias.³ However, there is no guarantee that bias-adjustment models are correct and reweighting data can increase variance in estimates.⁴ We focused on the comparison between targeted versus random incentives, rather than incentives versus no incentives, because the goals of the evaluation were to increase response rates and decrease nonresponse bias while minimizing survey costs.

Program change description

Existing evidence from federal and other surveys shows that financial incentives can increase response rates,⁵ but leaves an important question unresolved: how should agencies with a finite incentive budget decide which potential respondents to offer them to? We investigated whether targeting incentives to potential respondents estimated to have the highest risk of nonresponse improved the response rate and decreased estimated nonresponse bias relative to a comparison group who received incentives totally-at-random.

Evaluation design

We evaluated the effectiveness of targeting incentives during the 2021 wave of the AHS. Among the 86,000 potential respondents in the 2021 wave of the AHS who were also in the 2015, 2017, and 2019 waves, we estimated the risk of 2021 nonresponse using supervised machine learning models trained on data from those three past waves.⁶ The model was able to correctly predict the majority (55%) of nonrespondents, showing 44% better performance than simply using nonresponse in preceding waves as a proxy for nonresponse in the present one.

We grouped potential respondents with similar estimated risks of nonresponse into pairs and assigned one member of the pair to each of two different groups: 1. a “targeted” group (n = 43,000) of potential respondents who received incentives if

⁵ Financial incentives: Mercer, Andrew, Andrew Caporaso, David Cantor, and Reanne Townsend. 2015. “How Much Gets You How Much? Monetary Incentives and Response Rates in Household Surveys,” *Public Opinion Quarterly* 79, no. 1 (January): 105–129. Targeting: Jackson, Michael T., Cameron B. McPhee, and Paul J. Lavrakas. 2020. “Using Response Propensity Modeling to Allocate Noncontingent Incentives in an Address-Based Sample: Evidence from a National Experiment” *Journal of Survey Statistics and Methodology* 8, no. 2 (April): 385–411.

⁶ The U.S. Census Bureau reviewed this data product for unauthorized disclosure of confidential information and approved the disclosure avoidance practices applied to this release. (Data Management System (DMS) number: P-1234567, Disclosure Review Board (DRB) approval number: CBDRB-FY22-349). The counts are rounded according to Census Bureau entity count rounding rules.

¹ 85% in 2015, 80.4% in 2017, and 74.3% in 2019.

² [Standards and Guidelines for Statistical Surveys](#)

³ See [here](#) for noninterview adjustment factor (NAF) and raking methods.

⁴ See the full report for an example.

and only if they fell into the 30% with the highest estimated risk of nonresponse; 2. a “non-targeted” group (n = 43,000) of potential respondents, each of whom had a 30% probability of receiving an incentive irrespective of their estimated nonresponse risk. Those receiving an incentive via either method were randomly assigned an unconditional cash payment of \$2, \$5, or \$10, delivered in a letter sent to everyone in the sample in the weeks preceding the survey.⁷

Analysis of existing data

The response rate outcome was measured using data from the 2021 fielding of the AHS.⁸ To estimate changes in nonresponse bias, we evaluated⁹ whether the targeted and non-targeted respondents looked different across a range of key demographic and housing characteristics.¹⁰

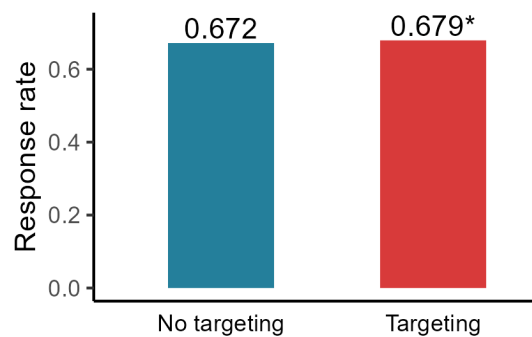
Results

The response rate among the group who received incentives targeted to those with the highest risk of nonresponse was 0.7 percentage points higher than the response rate of 67.2% in the group who received them completely at random. This result is statistically significant (p = 0.018).¹¹

In other words, targeting incentives to those predicted most at risk of nonresponse was more effective at increasing response rates than simply allocating incentives at random. Holding the total incentive budget constant, simply targeting incentives to the potential respondents with the

30% highest predicted risk of nonresponse induced an additional 300 people to respond compared to allocating incentives totally at random.

Figure 1. Targeting incentives increased the response rate relative to allocating incentives totally at random



Note: Asterisk indicates p < 0.05

Regarding nonresponse bias, a test of the joint statistical significance of differences across the ten key attributes produced a p-value of 0.42. Any small differences we observed could thus have been produced by random chance. Targeting for predicted nonresponse did not measurably improve or worsen nonresponse bias.

Implications

While our findings on the response rate increase are statistically significant, the measured improvement is small in magnitude and is not accompanied by discernable reductions in nonresponse bias. Though targeting was cost neutral compared to allocating incentives at random, the level of implementation effort was high, making the return on investment relatively low. Notably, the data collection overlapped with the onset of the COVID-19 pandemic and associated lockdowns, which may have muted the effectiveness of financial incentives.

Future evaluations could explore alternative targeting strategies such as whether it's possible to concentrate higher incentive amounts among those who not only have a higher estimated risk of nonresponse bias, but who are also predicted to contribute more towards minimizing nonresponse bias if they complete the survey.

⁷ See full report for details on the letter, study design, and methods.

⁸ Unless noted otherwise, all of the analysis reported in this abstract was prespecified in an [analysis plan](#).

⁹ Specifically, we employed a two-step procedure. First, we estimated the joint statistical significance of differences between respondents in the targeted and non-targeted groups across ten key attributes measuring demographics and housing quality issues. Conditional on finding the differences were jointly significant in an F-test, the second planned step involved measuring whether the attributes in the targeting group moved closer to population benchmarks in the Decennial Census and American Community Survey.

¹⁰ The full list of attributes is: own house (no; yes with mortgage/loan; yes with no mortgage/loan); average household size; white alone (householder); age (householder); presence of rodents; presence of mold; Census division; HUD-assisted unit as of 2013; metropolitan area as of 2013; type of housing unit.

¹¹ The p-value is derived using the pre-registered randomization inference procedure. See [analysis plan](#) and full report.