

# Grace Hopper 2020 SMS analysis

## Data preparation technical note

Different from the original analysis, about half respondents not receiving the follow-up phone interview attempt in the SMS data were also merged, because they can be included for the analysis on the binary voting behavior (1/0).

169 duplicated observations were removed after merging. Race, marital status and age are binned to be more interpretable in the regression models.

In the experiment, 1169 respondents (4.6%) did not receive text message correctly corresponding with their assignment (e.g. 355 people in the placebo group did not receive any text). I filtered out these observations to maintain a good treatment/control distinction.

There are around 3% missing values (both hand-coded and NA) in some of RAND's demographic data. The missing pattern is largely random and the portion is small, so I deleted missing observations list-wise before analysis.

Because we want to predict against support for Hopper. Compared with ordinal models, OLS would be a good alternative since it has 5 levels. I transform the survey response to a 0-10 scale, hoping it will make the understanding of models easier later.

After recoding, a data consists of complete demographic information has 21974 observations, with 4293 valid response on Hopper support and 21770 on voting behavior. From a logistic regression on the final data, there's only small difference with no significance in distribution of confounding variables between treatment and placebo group, therefore it's safe to regard it as a result from randomized experiment.

## High level analytical findings

At an aggregated level, the SMS program showed a weak significance in increasing the likelihood of people to participate voting: receiving the treatment will increase the odds of voting by 0.05, with a about 0.05 p-value indicating moderate level of confidence. However, by simply looking at ATE for subgroups in gender and race, such effect turns out to be actually very heterogeneous. The treatment effect for women adds 0.24 to the odds of voting significantly, yet there's barely any signs of encouragement for men. Another subgroup analysis on race shows that the program can even have a negative treatment effect on latino race voters, discouraging them from voting.

Within those we were able to interview, good news is that we see a significant positive impact of the SMS program on the support for Hopper. In a scale of 10 (0=strongly oppose, 10=strongly support), the treatment would increase around 0.4 degree of support. Moreover, such treatment effect seems to be universal among subgroups: after roughly checking regression tables for subgroup analysis, overall we didn't find significant differences in ATE across subgroups.

Going forward as to recommendation, when encouraging respondents to vote, it's the best to target the efforts to subgroups where the ATE is larger and positive (e.g. women). If the purpose is to increase general support for Hopper, the impact is largely universal thus the campaign can be directed to a wider audience.