Vote.org 2018 SMS Voter Mobilization Program: Mobilization for In-person Voting from Any Treatment

Prepared by Christopher B. Mann, Ph.D.
and Katherine Haenschen, Ph.D.

Executive Summary

For the 2018 general election, Vote.org conducted SMS voter mobilization programs covering 12,681,951 people of color and unmarried women in 33 states. These programs used "cold" text messaging to registered voters who have no prior relationship to Vote.org.

This memo evaluates the overall impact of Vote.org's SMS voter mobilization program for in-person voting, including heterogeneous treatment effects across subgroups. This memo pools the treatments used in four message tests, plus two late additions to the in-person voter mobilization program. Details of these 6 elements of the in-person mobilization program can be found in the following Vote.org 2018 SMS Voter Mobilization Program memos:

- Message Test of Adopt-a-Voter vs. Calendar Reminder vs. Standard Practice
- Message Test of Candidate Name vs. Standard Practice
- Message Test of Digital Sticker vs. Standard Practice
- Message Test of Social Pressure vs. Political Efficacy vs. Standard Practice
- Mobilizing Voters For Election Day With One SMS Message
- Extending Mobilization to High Propensity Voters in States with In-Person Voting

This experiment covered 9 million low propensity and/or newly registered voters in 30 states: AZ, CA, CT, FL, GA, IA, IL, IN, KS, MD, ME, MI, MN, MO, MS, MT, NC, ND, NM, NV, NY, OH, PA, TN, TX, VA, VT, WI, WV.

Vote.org's SMS treatments generated an average increase in turnout of 0.26 percentage points at a cost per net vote of $85.69 (11.7 net votes/$1000). The program caused 17,586 people to vote who otherwise would have missed participating in the 2018 General Election. If the weak confidence cell phone matches and unsuccessful attempt to mobilize people for early in person voting only are excluded, the average effect is 0.42 percentage points at a cost per net vote of $53.14 (18.8 net votes/$1000).

Vote.org's "cold" SMS voter mobilization program increased turnout in the 2018 election and identified several important lesson about targeting (discussed in this memo) and messages (discussed in other memos) to improve future effectiveness.

Objectives and Context

For the 2018 general election, Vote.org conducted SMS voter mobilization programs covering 12,681,951 people of color and unmarried women in 33 states. Despite widespread use, SMS messages have received little attention from researchers as a medium for campaign communication. In 2016, Vote.org established that "cold" SMS messages could increase turnout with a randomized experiment design covering 1.2 million young people of color and unmarried women in 7 states. Vote.org's 2016 "cold" SMS voter mobilization program increased turnout by 0.2 percentage points.

In 2017, Vote.org replicated and expanded testing of "cold" SMS voter mobilization with a randomized experiment covering 714k young people of color and unmarried women for the Virginia gubernatorial and legislative elections. Vote.org's 2017 test of "cold" SMS voter mobilization increased turnout by 0.6 percentage points and identified Standard Practices regarding timing and message framing.

The 2018 programs build on Vote.org's successful SMS voter mobilization programs in 2016 and 2017. This memo evaluates the overall impact of Vote.org's SMS voter mobilization program for in-person voting, including heterogeneous treatment effects across subgroups. This memo pools the treatments used in four message tests, plus two late additions to the in-person voter mobilization program. Details of these 6 elements of the in-person mobilization program can be found in the following Vote.org 2018 SMS Voter Mobilization Program memos:

- Message Test of Adopt-a-Voter vs. Calendar Reminder vs. Standard Practice
- Message Test of Candidate Name vs. Standard Practice
- Message Test of Digital Sticker vs. Standard Practice
- Message Test of Social Pressure vs. Political Efficacy vs. Standard Practice
- Mobilizing Voters For Election Day With One SMS Message
- Extending Mobilization to High Propensity Voters in States with In-Person Voting

This test was conducted across 9 million low propensity and/or newly registered voters in 30 states: AZ, CA, CT, FL, GA, IA, IL, IN, KS, MD, ME, MI, MN, MO, MS, MT, NC, ND, NM, NV, NY, OH, PA, TN, TX, VA, VT, WI, WV. These states cover a range of electoral contexts (defined by competitiveness, voting procedures, and other characteristics). In states with extensive EIPV use, the treatments were targeted at both EIPV and EDay voting: Arizona, Florida, Georgia, Illinois, Indiana, Kansas, Nevada, New Mexico, North Carolina, Tennessee, Texas, Wisconsin, West Virginia. In the remaining states, the treatments targeted only EDay voting: California, Connecticut, Iowa, Maryland, Maine, Michigan, Minnesota, Missouri, Mississippi, Montana, North Dakota, New Jersey, New York, Ohio, Pennsylvania, Virginia, Vermont.

In the states with EIPV and EDay voting, the 2018 SMS voter mobilization program addresses a secondary research question about mobilization for these two types of in-person voting: is it more effective to mobilize voters to vote early, to vote on Election Day, or to mobilize for early voting and then for Election Day voting? To the best of our knowledge, this question has received almost no attention despite the widespread availability of early in person voting. This research question is evaluated in a separate memo, "Vote.org 2018 SMS Voter Mobilization Program: Timing of Encouraging In Person Voting for Early Voting or Election Day".

Selected Universe

The data for the experiment was selected by Vote.org from the voter file maintained by TargetSmart, a firm providing voter data.

The 9,028,556 registered voters included in the experiment met the following criteria:

1) A cell number available in the TargetSmart database
    - TargetSmart provided the best single record for each available cell phone number (i.e. no duplicate numbers)
2) Registered to vote in the following states: AZ, CA, CT, FL, GA, IA, IL, IN, KS, MD, ME, MI, MN, MO, MS, MT, NC, ND, NM, NV, NY, OH, PA, TN, TX, VA, VT, WI, WV
3) Low propensity voter or new registrant:
    - 10-70 Vote propensity OR
    - Voted in Gen 2016 and registered between Dec 2014-Nov 2016 OR
    - Registered December 2016-present)
4) People of color or unmarried women:
    - People of color: individuals coded as non-white by TargetSmart or individuals residing in areas with a Census population of at least 67% non-white.
        - The latter criteria is intended to capture false negatives for non-white in the individual coding data. The race coding is based on state voter file information about race (where available) and proprietary models of race maintained by TargetSmart.
    - Females under age 30 who were not included using the criteria above
        - Arizona, Illinois, Michigan, Minnesota, Nevada, New Mexico, Vermont, Wisconsin only.
5) Exclusions:
    - Request mail ballot for Gen 2018 -OR- permanent mail ballot status
    - Age under 18 years old or over 100 years old

Treatments:

The experiment compares an uncontacted control group to the treatment group. Details of the treatments are discussed in the memos of the different tests in the in-person mobilization program.

In states without EIPV, each treatment consisted of a series of three SMS messages. In states with EIPV, treatment consisted of a series of three to five SMS messages. The three message series were identical to the non-EIPV states. In the five message series, the first two messages were repeated (1st & 3rd; 2nd & 4th) for EIPV and then EDay voting. In this memo, the three and five message treatments in EIPV states are pooled. As noted above, the differences between assignment to three messages for EIPV, three messages for EDay, and five messages for both EIPV and EDay is evaluated in a separate memo “Encouragement to use different methods of in-person voting”.

Prior to each round of text messages, anyone who "opted out" of receiving text messages was removed from the contact list. Also, anyone who who cast a ballot (EIPV or mail ballots) according to public records acquired by TargetSmart LLC were removed from the contact list upon Vote.org’s receipt of this information.

Intended Effects[i]

- Each treatment was intended to increase turnout in the November 2018 election.
- Each treatment was expected to have different effects on turnout.
- Different treatment effects were expected across the following groups:
    - States
    - Voters under and over age 30
    - Cell phone match confidence
    - Competitive vs. non-competitive areas[ii]
    - Gender
    - Age
    - Vote propensity score
    - Drop-off voters (voted in 2016 but not 2014)
    - New registrants (since 2016)
    - Race / ethnicity
    - Households with single vs. multiple targets

Evaluation Design

The evaluation is based on a randomized trial design (or field experiment) that is considered best practice by academic researchers and the Analyst Institute. Each treatment group received SMS messages from Vote.org; the control group was sent none of the SMS messages.

The randomization is conducted at the household level to reduce the risk of contaminating behavior of co-habitants. For this experiment, households were defined as people with the same mailing address. The randomization uses an automated re-randomization procedure to ensure good balance in characteristics available from the voter file prior to delivery of treatment (see Technical Appendix).

Random Assignment to Treatment & Control

	Individuals
Treatment	6,763,685
Control	2,264,871

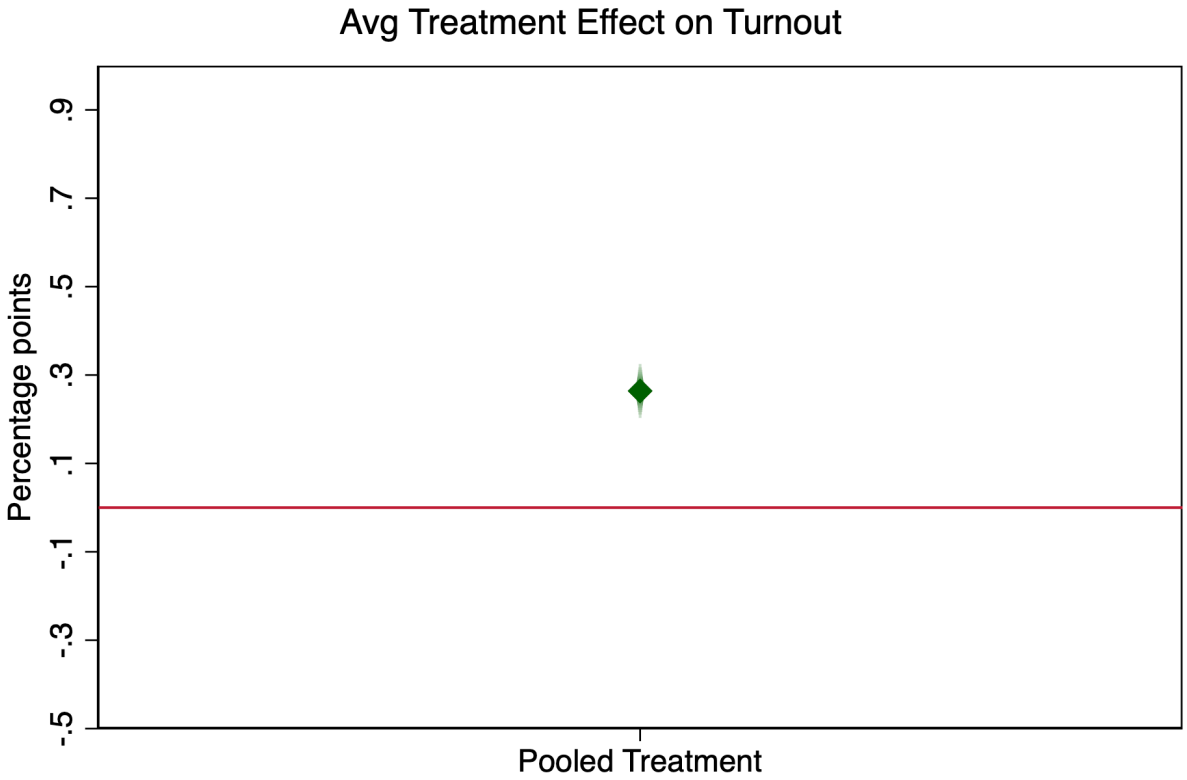
Since each of the six tests have different probabilities of assignment, the overall treatment effects are estimated with weights of the inverse probability of assignment and dichotomous variables for each experiment.

Results

Vote.org's SMS treatments generated an average increase in turnout of 0.26 percentage points.[iii] The program caused 17,586 people to vote who otherwise would have missed participating in the 2018 General Election.

Note on reading the graphs in this memo: The estimated treatment effect is represented by the diamond shape in the middle of each bar. The gradient error bars display the statistical uncertainty of this estimate. Like traditional error bars, the ends of the gradient error bars indicate the 95% confidence range. If these bars cross the red horizontal line at zero, the difference from the control group is not statistically significant. The width and intensity (darkness) of the bar indicate the

statistical likelihood that the treatment effect falls in this range, so the bars are wider and darker close to the diamonds, thinning and fading farther away. When comparing to treatment effects, the likelihood of being different can be seen by the width and intensity of the overlapping gradient bars.[iv]



Notes: Treatment effects estimated from regression with covariates for precision. Gradient confidence intervals by line width and intensity (max=95% c.i.). If confidence intervals cross line at zero, then effect is not statistically significant.

Subgroups

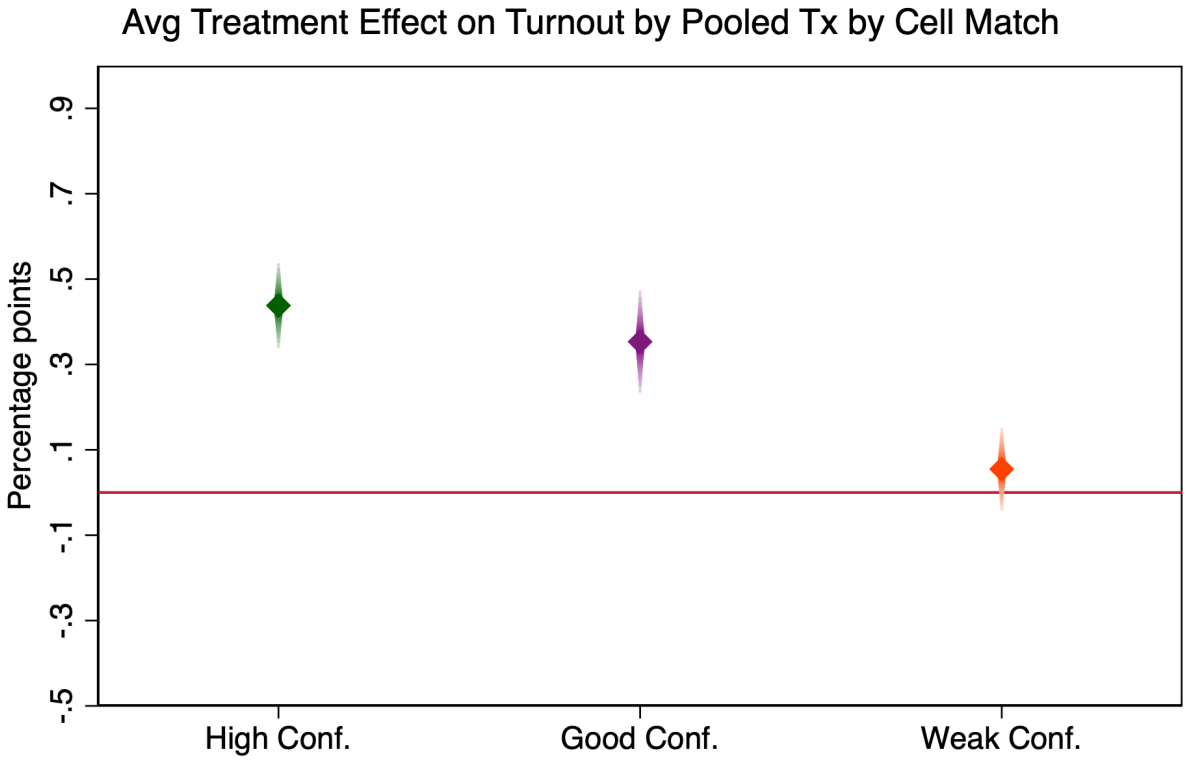
Noteworthy differences appear for several subgroups for the in-person voter mobilization program.

Cell Phone Match Confidence

Vote.org uses cell phone numbers obtained when voter registration records are matched to cell phone records (by the voter file firm). In 2018, Vote.org requested the individual with the highest confidence mobile phone match for each unique phone number. The process of matching voter registration records to mobile phone records uses an iterative process of different combinations of information (first name, last name, initials, address elements). In the data from TargetSmart, matches using more information are considered higher confidence than matches using less information on a scale of 1 (low) to 11 (high). Many civic and political organizations exclude records with lower confidence matches from SMS contact programs, but in 2018 Vote.org included the best confidence individual for each unique phone number for two reasons: 1) determining the appropriate level of confidence for future exclusions and 2) belief that lower likelihood of contact by other organizations would make Vote.org's program more important (normatively and empirically). Based on the

distribution of confidence match scores, records were stratified into three approximately equal size categories for blocked random assignment: high confidence, good confidence, weak confidence.

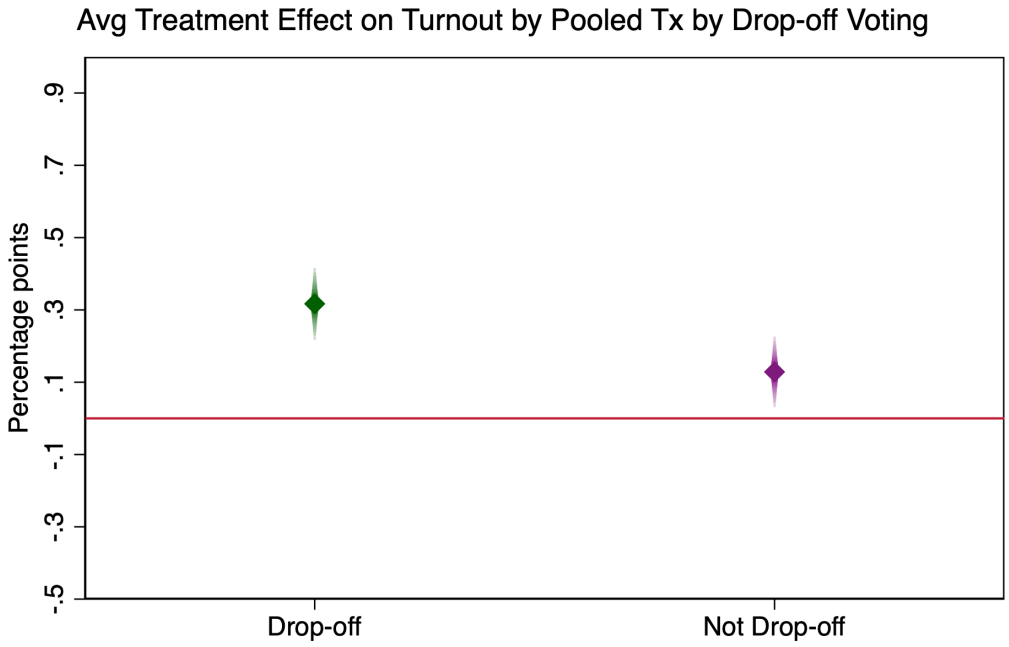
Vote.org's treatment generated significant increases in turnout among records with high confidence (0.44 percentage points[v]) and good confidence (0.35 percentage points[vi]). However, among the weak confidence records, there was no statistically significant effect (0.05 percentage points[vii]). The difference across these subgroups is statistically significant.[viii]



Notes: Treatment effects estimated from regression with covariates for precision. Gradient confidence intervals by line width and intensity (max=95% c.i.). If confidence intervals cross line at zero, then effect is not statistically significant. Difference in effects across cell match confidence is statistically significant (p=0.000).

Drop-off Voting

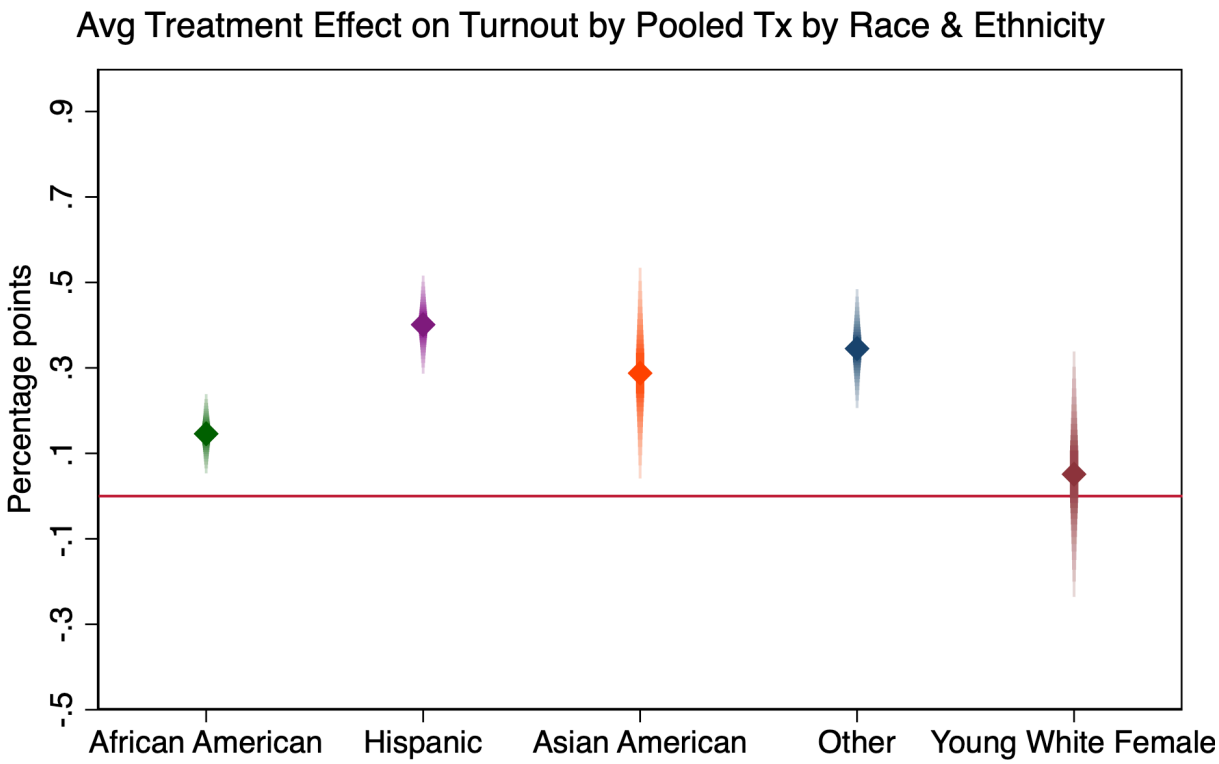
Drop-off voters, who were registered in 2014 and voted in 2016 but not in 2014, are a classic voter mobilization targets in mid-term elections. Vote.org's treatments increased turnout in both subgroups. As expected, the treatments are somewhat more effective among voters who voted in 2016 but had not voted in 2014 (0.32 percentage points[ix]) compared to people who voted in both elections (0.13 percentage points[x]). The difference across these subgroups is statistically significant.[xi]



Notes: Treatment effects estimated from regression with covariates for precision. Gradient confidence intervals by line width and intensity (max=95% c.i.). If confidence intervals cross line at zero, then effect is not statistically significant. Difference in effects between Drop-off and Not Drop-off is statistically significant (p=0.025).

Race & Ethnicity

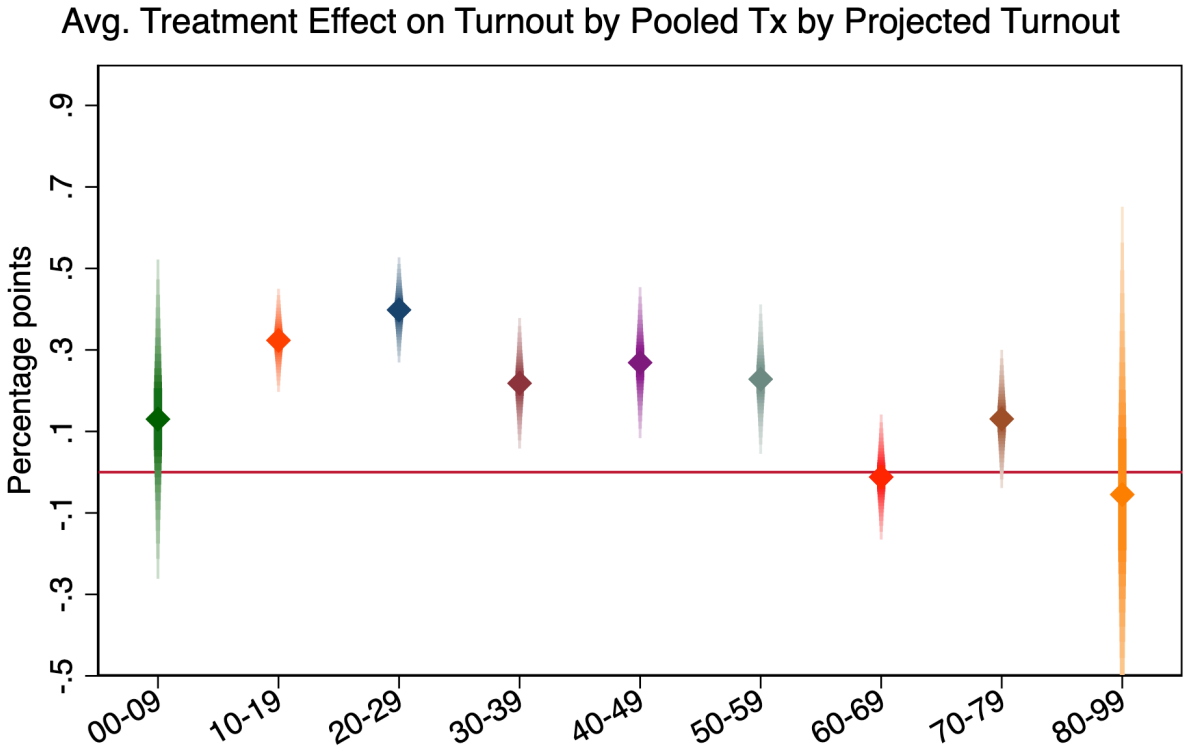
Vote.org's treatments significantly increased turnout in all subgroups defined using TargetSmart's codes for race and ethnicity, except young white females. The effect among African Americans is somewhat smaller in 2018 than among Hispanic, Asian American, and Other non-white. The difference across these subgroups is statistically significant.[xii]



Notes: Treatment effects estimated from regression with covariates for precision. Gradient confidence intervals by line width and intensity (max=95% c.i.). If confidence intervals cross line at zero, then effect is not statistically significant. Difference in effects across race/ethnicity is statistically significant (p=0.030).

Projected Turnout

Vote.org's targeting used TargetSmart's vote propensity score for the 2014 mid-term general election, plus newly registered voters (accounting for a small number of records below and above the score selection parameters). Based on past research by academics and practitioners, the largest effects should be expected for registered voters in the middle of the vote propensity distribution. Vote.org's treatments significantly increased turnout in subgroups from 10-59, as well as a positive but not significant effect in the lowest group. The effects attenuate to null for the high propensity voters. This pattern is marginally statistically significant.[xiii]

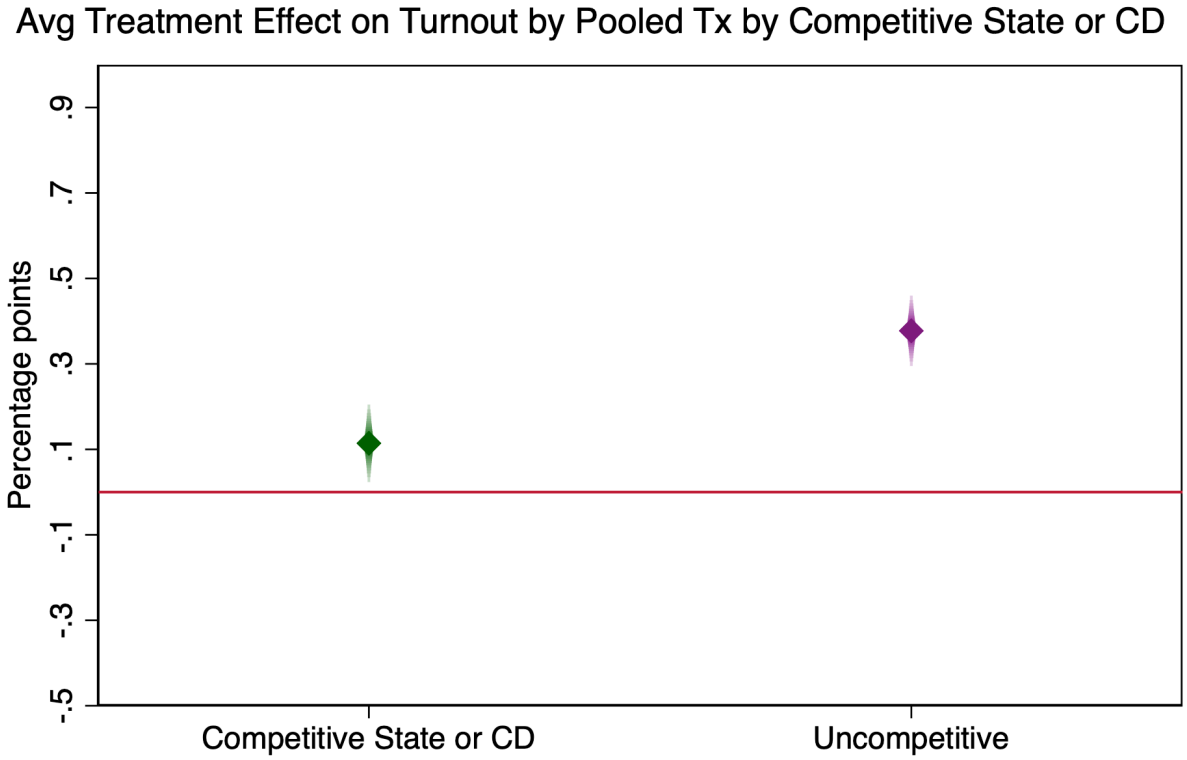


Notes: Treatment effects estimated from regression with covariates for precision. Gradient confidence intervals by line width and intensity (max=95% c.i.). If confidence intervals cross line at zero, then effect is not statistically significant. Difference in effects between age groups is *marginally* statistically significant (p=0.070).

Competitiveness

Higher levels of electoral competition for big-budget top-of-the-ticket contests often attenuate the observed effects of voter mobilization treatments. Due to the amount of money spent in big-budget top-of-the-ticket contests, higher levels of advertising, media coverage and other mobilization efforts make it more difficult to have a net effect. To estimate the influence of these big-budget contests on magnitude of treatment effect, we use the final RealClearPolitics rankings of contests for US House, US Senate, and Governor (the top-of-the-ticket in the mid-term election). The treatments increased turnout in both competitive (0.11 percentage points[xiv]) and uncompetitive areas (0.38 percentage points[xv]). This difference is statistically significant.[xvi]

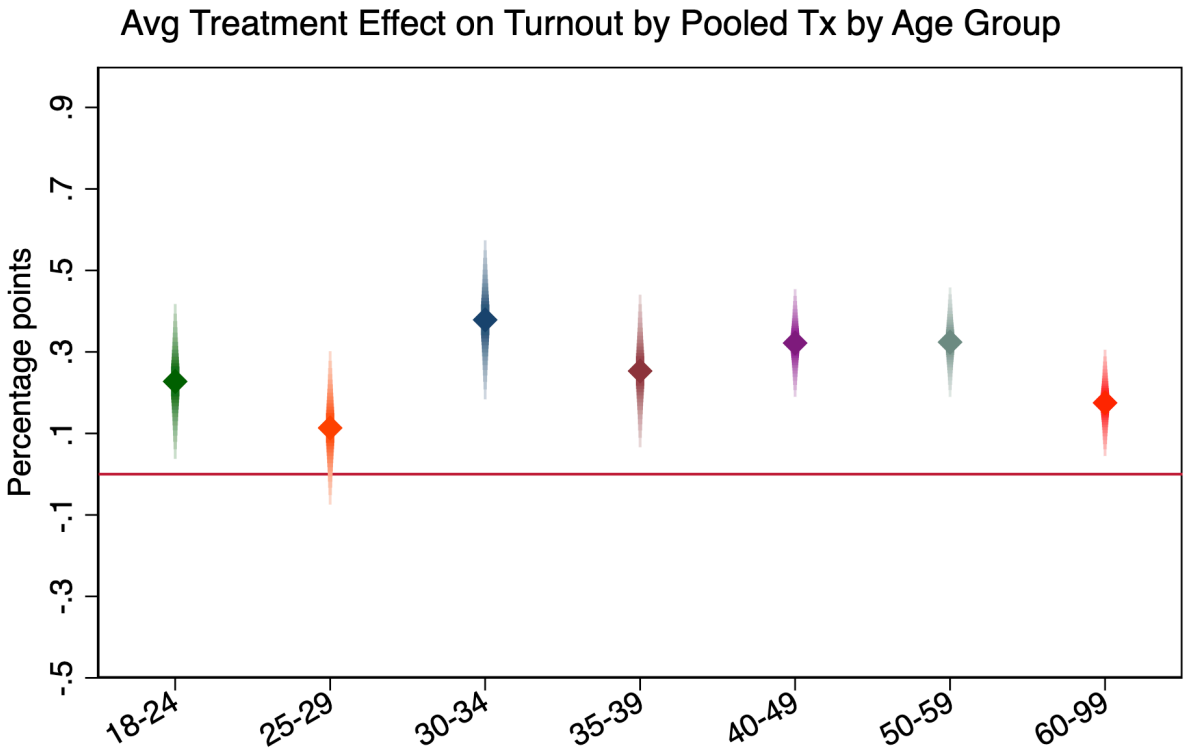
It is important to note that there are many competitive races for other offices in areas that do not have a competitive big-budget contest, including several surprising outcomes in races for top-of-the-ticket offices and many state and local offices that make important policy decisions. Competitive and uncompetitive are proxies for other campaign spending, not for importance of mobilizing under-represented voices in the electorate.



Notes: Treatment effects estimated from regression with covariates for precision. Gradient confidence intervals by line width and intensity (max=95% c.i.). If confidence intervals cross line at zero, then effect is not statistically significant. Difference in effects by state competitiveness is statistically significant (p=0.000).

Age

For many people, SMS may seem like a tactic for mobilizing young people but not older people. Consistent with this expectation, the population of people who can be matched to cell phone numbers skews young. However, among the population for whom TargetSmart could obtain phone numbers there was no significant difference across age groups.[xvii]



Notes: Treatment effects estimated from regression with covariates for precision. Gradient confidence intervals by line width and intensity (max=95% c.i.). If confidence intervals cross line at zero, then effect is not statistically significant. Difference in effects between age groups is *not* statistically significant (p=0.525).

Additional Subgroups

There was no evidence of significant heterogeneity across any of the other subgroups listed in the Intended Effects Section.

Net Votes

The cost per net vote (and net votes/$1000) calculation includes all costs of design, delivering, and managing the treatment delivery process.

Treatment	Effect	Net Votes	Votes/$1000	CPV	Treatment Cost
Any Treatment	0.26 pp	17,586	11.7	$85.69	[$0.2228/individual]
Excluding Weak Cell Matches & EIPV Only	0.42 pp	14,839	18.8	$53.14	[$0.2232/individual][xviii]

Notes: Treatment cost reflects average cost for the series of SMS messages in each treatment. Net votes is the number of people who voted in response to the treatment(s), and would not have otherwise voted in the November 2018 election.

Lessons Learned

Vote.org's strategy of using "cold" SMS messages for voter mobilization continues to generate significant and cost-effective increases in voter turnout in mid-term elections.

Vote.org should select good and strong match confidence cell phone records for future programs.

Future Steps

Vote.org should continue to invest in "cold" SMS voter mobilization programs to increase voter turnout.

Cautions

The effect of any voter mobilization communication is conditional on the execution of the program, the jurisdiction, the type of election, the level of interest in the election, and the activities of other organizations. Repeating these treatments in other election contexts or with variations of the treatments could produce different results.

Technical Appendix

Randomization Procedure:

Randomization was conducted at the household level. The random assignment was conducted in Stata using the "re-randomize" procedure developed by Kennedy and Mann (2015) to ensure balance across observable covariates.[xix]

This procedure rejects any instance of randomization outside of pre-determined parameters: minimum of 10 iterations and maximum of 25 iterations. Iterations stopped between 10 and 25 when iteration had p>0.8 based on Malahanobis distance test. This procedure produced equal sized groups, and each group was designated as an experimental condition. Blocked randomization used equal probabilities of assignment in all blocks.

Blocked randomization using the following variables: State, Young (under 30 years old), Quality of cell phone match to individual (three strata based on TargetSmart cell phone match confidence code)

Balance checked using age, female, individual-level race codes (Hispanic, African American, white), past voting history (dummies for voting in the 2010, 2012, 2014, and 2016 general elections), and three-digit zip-code (geography).

Statistical Methods for Analysis:

The analysis is based on matching the pre-election experimental population to post-election vote history from TargetSmart. The matching used the unique TargetSmart record identification number. Analysis was conducted using standard regression techniques for evaluating experimental results.

Hypothesis testing uses robust standard errors clustered by unique address to account for potential correlation between the behaviors of co-habitants.

All reported estimates are calculated using models that include the covariates used to check balance in the random assignment procedure. As expected from a well-balanced experiment, the estimates are essentially identical when estimated without these covariates.

Technical Endnotes

[i] Following best practice in academic research, the intended treatment effects and plans for analysis were pre-registered with the Evidence in Governance and Politics program at the University of California at Berkeley (egap.org).
[ii] Competitive districts defined by RealClearPolitics.com as Toss-up, Leans Dem or Leans GOP.
[iii] Avg. treatment effect compared to the control group is statistically significant at $p<0.001$. SE = 0.004
[iv] Research by Isabelle Fischer (2018) finds that people are much more likely to correctly interpret data displayed with gradient error bars than other more commonly used data visualizations.
[v] Avg. treatment effect among high confidence records compared to the control group is statistically significant at $p<0.001$. SE = 0.006

[vi] Avg. treatment effect among good confidence records compared to the control group is statistically significant at $p<0.001$. SE = 0.007
[vii] Avg. treatment effect among weak confidence records compared to the control group is not statistically significant at $p=0.347$. SE = 0.006
[viii] The difference in the avg. treatment effect across confidence levels is statistically significant at $p<0.001$. 7
[ix] Avg. treatment effect among drop-off voters compared to the control group is statistically significant at $p<0.001$. SE = 0.006
[x] Avg. treatment effect among people who voted in 2016 & 2014 compared to the control group is statistically significant at $p=0.030$. SE = 0.006
[xi] The difference in the avg. treatment effect voting history is statistically significant at $p=0.030$.
[xii] The difference in the avg. treatment effect across race & ethnicity is statistically significant at $p=0.025$.
[xiii] The difference in the avg. treatment effect across projected turnout is marginally statistically significant at $p=0.070$.
[xiv] Avg. treatment effect among voters in areas with big-budget top-of-the-ticket contests compared to the control group is statistically significant at $p=0.038$. SE = 0.006
[xv] Avg. treatment effect among voter in in areas with big-budget top-of-the-ticket contests compared to the control group is statistically significant at $p<0.001$. SE = 0.005
[xvi] The difference in the avg. treatment effect across competitive and uncompetitive areas is statistically significant at $p<0.000$.
[xvii] The difference in the avg. treatment effect across age groups is not statistically significant at $p=0.525$.
[xviii] The average cost per record assigned to treatment is slightly higher when excluding the EIPV Only treatment condition with three SMS messages (i.e. the average increases because a larger proportion of records were sent 5 SMS messages).
[xix] Kennedy, Chris, and Christopher B. Mann. 2015. *RANDOMIZE: Stata Module to Create Random Assignments for Experimental Trials, Including Blocking, Balance Checking, and Automated Rerandomization*. Boston College Department of Economics. https://ideas.repec.org/c/boc/bocode/s458028.html (May 16, 2017).