



To Vote Rev Action Fund
From One Minus Beta Analytics
Date April 21, 2023

2022 High-Traffic On-The-Spot (HTOTS) Vote Tripling Program Evaluation

Executive Summary

Vote Rev ran a relational GOTV program to encourage people to vote in the 2022 Georgia runoff election. Vote Rev partnered with GRSG, a paid canvassing vendor, to station canvassers in high-traffic locations such as college campuses and shopping centers. Canvassers recruited people who were willing to send a message to five friends to remind them about the upcoming election. People who were recruited by canvassers (“triplers”) were asked to text their friends (“triplees”) then and there (“on the spot”).¹

We collaborated with Vote Rev to embed an experiment into this program. To conduct this experiment, Vote Rev randomized all registered voters in Georgia between a treatment group and a control group. Vote Rev uploaded the experiment assignment for each registered voter to a voter file lookup app. Canvassers asked triplers to provide their friends’ names, used the voter file lookup app to find each friend, and then asked triplers *either* to text that friend (if the friend was in the treatment group) *or* not to text that friend (if the friend was in the control group). About 75% of triplers in the experiment named at least two friends, and about 40% of triplers in the experiment named five friends.

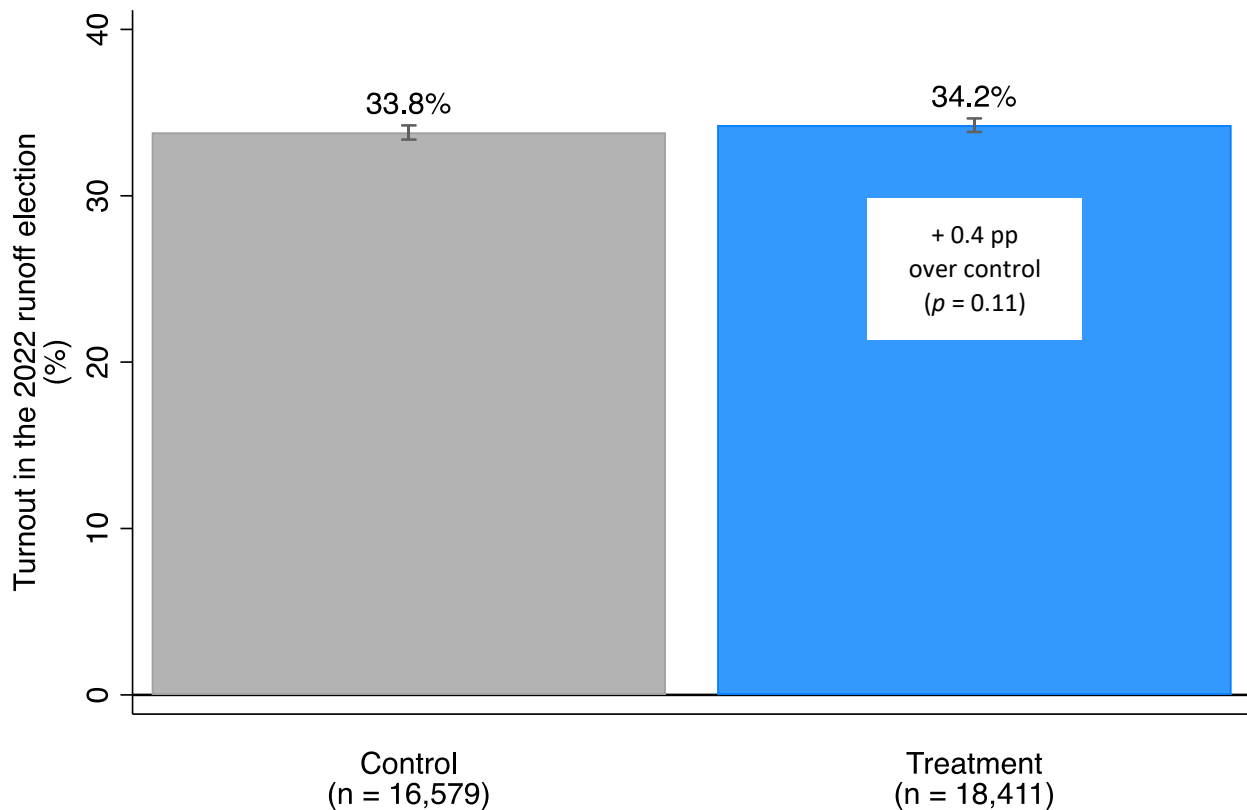
This experiment evaluates the effect of the program on triplees. Triplees in this experiment were young and diverse. Over half of triplees were under 35 years old, and nearly half of triplees were identified on the voter file as Black. The implementation of this experiment appears to have been very good. Canvassers recorded that they observed triplers sending a text message to 98% of triplees in the treatment group and recorded that they did not observe triplers sending a text message to 95% of triplees in the control group.

We are pleased to report that the HTOTS program probably increased turnout in the 2022 runoff election (Figure 2). About 33.8% of triplees in the control group voted in the runoff election, and about 34.2% of triplees in the treatment group voted in the runoff election. While the difference between experiment conditions does not meet the threshold for statistical significance, it is probably not due to chance. Additionally, we ran several robustness checks, all of which indicate that the program probably increased turnout by 0.4-0.5 percentage points. We estimate that the HTOTS program mobilized 81 voters in the 2022 runoff election.

We provide several recommendations for future HTOTS programs in the Discussion section. We hope that these results are informative as Vote Rev begins to make plans for the 2023-2024 election cycle. We thank Vote Rev for conducting this project with us, and we look forward to collaborating with you again.

¹ According to Vote Rev, the terms “tripler” and “triplee” were coined in an earlier version of this intervention when people were asked to name three friends. We use the legacy terms “tripler” and “triplee” in this report for consistency.

Figure 2: The HTOTS program probably increased turnout in the 2022 Georgia runoff election



Notes: The p-value is calculated using a one-tailed test. The black lines represent 90% confidence intervals. The raw difference between conditions may not equal the estimated treatment effect due to rounding.

Experiment Design and Implementation

Research Questions

This experiment explores the following research question:

- How effective was the high-traffic on-the-spot (HTOTS) vote tripling program at increasing turnout in the 2022 Georgia runoff election?

Experiment Universe

The program included nearly 46,000 tripler-triplee relationships across 15,000 triplers and 44,000 triplees.² We excluded tripler-triplee relationships if they met one of the following criteria: (1) 4,065 tripler-triplee relationships where the tripler was canvassed during the first two days of the program, which Vote Rev specified would be a pilot, (2) 414 tripler-triplee relationships where the canvass data was identified by the Vote Rev field team as highly likely to be fraudulent based on unreasonable submission times, unreasonable speed, or suspicious GPS coordinates, (3) 527 tripler-triplee

² These counts exclude 227 tripler-triplee relationships that met one of the following criteria: the same triplee was named more than once by the same tripler or the triplee did not match to the voter file (either because the canvasser attempted to manually enter a triplee into the voter file or because the canvassing software did not accurately record the interaction).

relationships where the triplee was named by different triplers and was in different experiment conditions for each tripler, and (4) 4,791 tripler-triplee relationships where the triplee voted in the runoff election prior to being named by the tripler.³

After applying these exclusions, the experiment universe contains 36,235 tripler-triplee relationships across 12,753 triplers and 34,990 triplees. The demographics of triplees are provided in Table 1. Over half of triplees are under 35 years old, and nearly half of triplees are identified on the voter file as Black. About half of triplees voted in the 2022 general election, and slightly more than half of triplees voted in the 2020 runoff election.

Table 1: Over half of triplees are under 35 years old, and nearly half of triplees are identified on the voter file as Black

	Experiment Universe (n = 34,990)
Female	54%
Voter file race: Black	47%
Voter file race: White	40%
Voter file race: Asian	5%
Voter file race: Hispanic	4%
Voter file race: Native American	1%
Voter file race: Other or Unknown	3%
Age: 18-23	25%
Age: 24-29	16%
Age: 30-39	19%
Age: 40-49	15%
Age: 50-59	13%
Age: 60+	13%
Average runoff turnout score	55
Average ideology score	65
Voted in 2022 general election	42%
Voted in 2020 runoff election	52%

Notes: The turnout rate in the 2020 runoff election is calculated among people who were old enough to vote in that election. Average turnout and ideology scores are calculated among people who are not missing those scores.

³ These cases were evenly balanced between experiment conditions, representing about 12% of each condition.

Randomization

Vote Rev randomized all households in Georgia with at least one registered voter between the following experiment conditions:

- (1) *Treatment group*: If a tripler named a triplee assigned to the treatment group, the tripler was asked to text the triplee to remind them to vote in the runoff election. (n = 50% of households with at least one registered voter in Georgia)
- (2) *Control group*: If a tripler named a triplee assigned to the control group, the tripler was not asked to text the triplee. (n = 50% of households with at least one registered voter in Georgia)

Prior to launching the program, Vote Rev developed a research protocol. This research protocol stipulated that Vote Rev would stratify on household size in the randomization. When we reviewed the research protocol, we recommended that Vote Rev stratify on additional variables including the number of registered voters in the household who voted in the 2020 runoff election, the number of registered voters in the household identified as Black on the voter file, the number of registered voters in the household identified as white on the voter file, and an indicator variable for households in the greater Atlanta area (because most canvassing was going to take place in the greater Atlanta area). Vote Rev had already randomized the universe but agreed to re-randomize the universe to incorporate our suggestions. However, due to communication and technical issues between Vote Rev and GRSG, the canvass program launched using the experiment assignments from the first randomization. After Vote Rev re-randomized the universe, GRSG refreshed the program to reflect the experiment assignment from the second randomization.

As a result, most people who met the targeting criteria were randomized twice, with several exceptions. Importantly, Vote Rev did not re-randomize triplees who had already been named by a tripler. Additionally, because the voter file updated between the first and second randomizations, people who were dropped from the voter file in the update were included in the first randomization but not in the second randomization. Similarly, people who were added to the voter file in the update were included in the second randomization but not in the first randomization.

The switch from the first randomization to the second randomization would not have affected the experiences of canvassers, triplers, or triplees. Unfortunately, there were also technical issues when Vote Rev switched from the first randomization to the second randomization, which resulted in 242 triplees being assigned to different experiment conditions in each randomization and being named by a tripler when each randomization was live. Because these triplees were assigned to both experiment conditions when they were named, we exclude these triplees from the analysis.

Vote Rev was able to identify whether each triplee was named when the first randomization was live or when the second randomization was live. To account for the multiple randomizations, we control for which randomization was live when triplees were contacted in the analysis.

Program Implementation

Vote Rev partnered with GSRG to run the canvass. The program launched on Friday 11/25, but Vote Rev specified that the first two days of the program would be a pilot. This experiment evaluates the program between Sunday 11/27 and Tuesday 12/6 (the date of the runoff election).

Overall, the implementation of this experiment appears to have been very good (Table 2). Canvassers recorded that they observed triplers sending a text message to 98% of triplees in the treatment group. Similarly, canvassers recorded that they did not observe triplers sending a text message to 95% of triplees in the control group. Most triplees in the treatment group were named by one tripler and sent exactly one message. About 4% of triplees in the treatment group were named by multiple triplers and sent more than one message.

About 41% of tripler-triplee relationships were provided by triplers on or near college campuses, while about 74% of tripler-triplee relationships were provided by triplers within one week of Election Day. Unfortunately, there are statistically-significant differences between experiment conditions on these program characteristics. We discuss this finding further in the following section.

Table 2: 98% of triplees in the treatment group were recorded as having been sent at least one message

	Treatment group	Control group
By contacts		
Sent at least one message	98%	5%
Sent 2+ messages	4%	0%
By location		
Campus	40%	42%
Non-campus	60%	58%
By date		
Sunday 11/27	6%	6%
Monday 11/28	7%	7%
Tuesday 11/29	13%	12%
Wednesday 11/30	11%	11%
Thursday 12/1	11%	10%
Friday 12/2	12%	12%
Saturday 12/3	9%	9%
Sunday 12/4	11%	11%
Monday 12/5	11%	11%
Tuesday 12/6	10%	10%

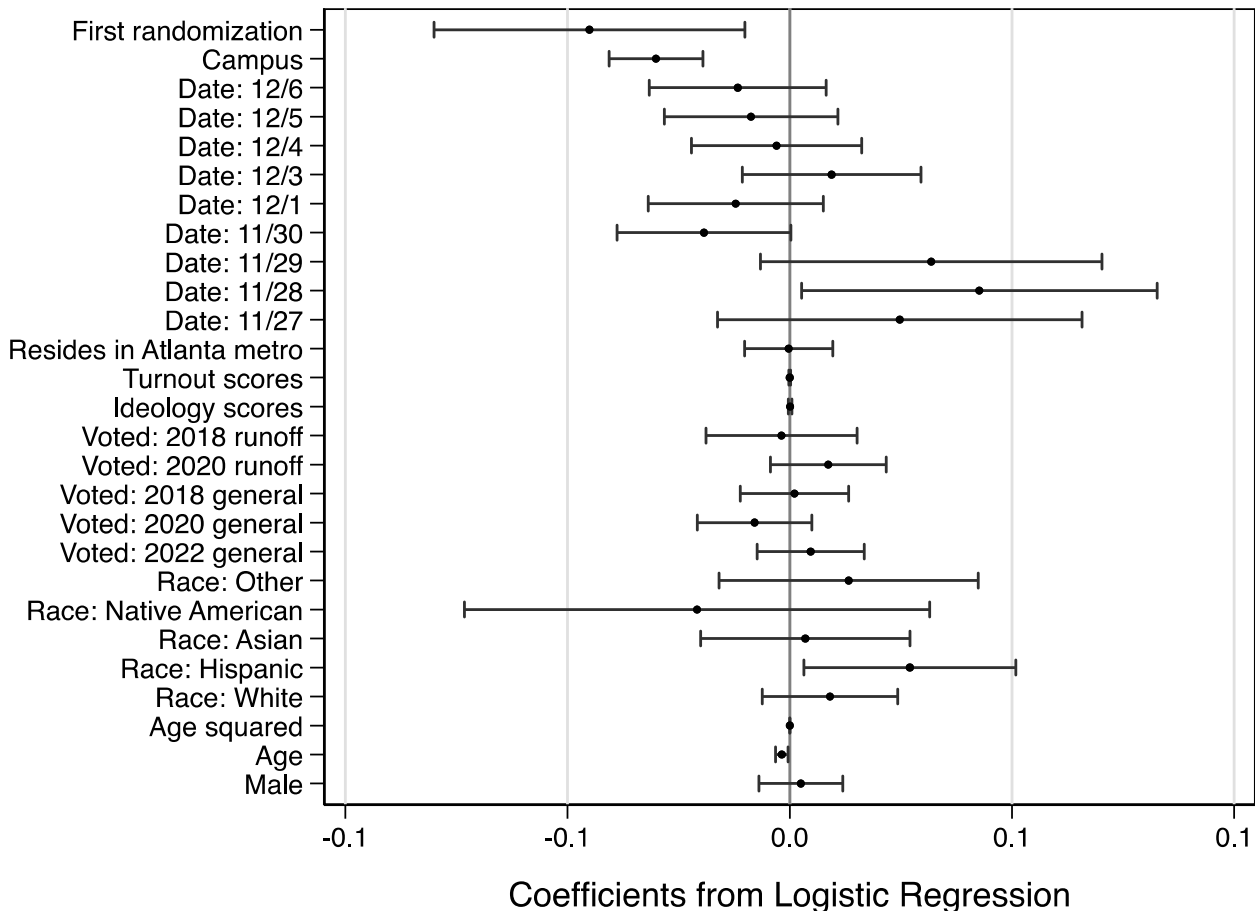
Notes: Contacts are expressed as a percent of triplees. Location and date are expressed as a percent of tripler-triplee relationships because the same triplee may have been named in different locations and on different dates.

Balance Checks

As noted previously, Vote Rev randomized all registered voters in Georgia. With about [7 million registered voters](#) in Georgia, only about 0.5% of people who were randomized were ultimately included in the experiment universe. We were concerned that the experiment conditions would not be comparable across demographic and program characteristics, and we brainstormed alternative experiment designs with Vote Rev prior to launching this experiment. However, Vote Rev was limited by the capabilities of the voter file lookup app and by the very short timeline in which to launch the program (which eliminated the possibility of building new capabilities into the app or using a different app).

The experiment universe contains 18,411 people (52.6%) in the treatment group and 16,579 people (47.4%) in the control group. As feared, people in each experiment condition varied across several demographic and program characteristics (Figure 1). Triplees in the treatment group were significantly less likely to have been named by a tripler recruited on a college campus than triplees in the control group. Meanwhile, triplees in the treatment group were significantly more likely to have been named closer to the election than triplees in the control group. Additionally, triplees in the treatment group were significantly younger and more likely to be identified on the voter file as Hispanic than triplees in the control group.

Figure 1: The experiment conditions varied across several demographic and program characteristics



After observing these imbalances, we explored the overlap in these covariates between experiment conditions using normalized differences. All normalized differences are less than 0.10, indicating that there is still sufficient overlap for a fully-interacted regression to adjust for underlying imbalances in the covariate distribution between experiment conditions. All analyses interact treatment assignment with all covariates with a normalized difference of 0.01 or greater to adjust for this covariate imbalance.⁴

Outcome Measurement

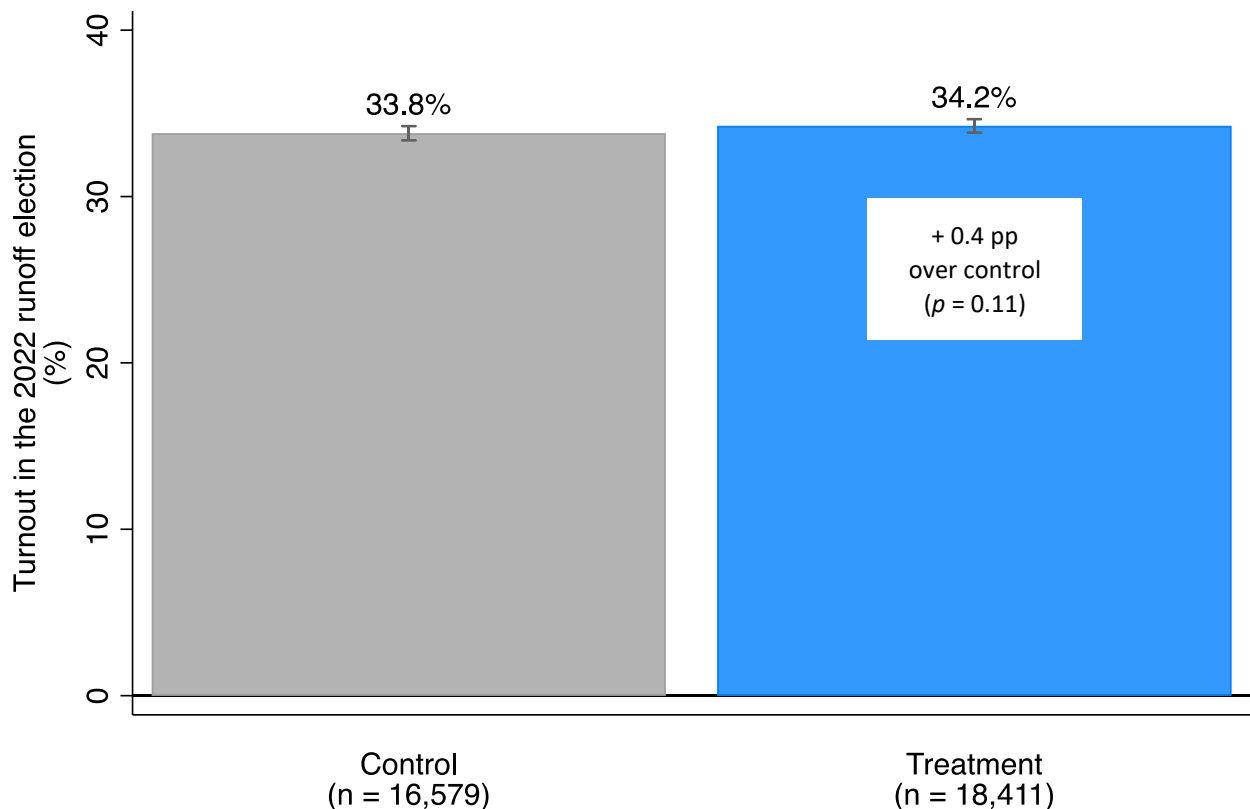
The outcome for this experiment is whether triplees voted in the 2022 runoff election on Tuesday 12/6. This outcome was measured using the official voter file provided by TargetSmart in March 2023.

Results

Main Results

We are pleased to report that the HTOTS program probably increased turnout in the 2022 runoff election (Figure 2). About 33.8% of triplees in the control group voted in the runoff election, and about 34.2% of triplees in the treatment group voted in the runoff election. While the difference between experiment conditions does not meet the threshold for statistical significance, it is probably not due to chance.⁵

Figure 2: The HTOTS program probably increased turnout in the 2022 Georgia runoff election



Notes: The p -value is calculated using a one-tailed test. The black lines represent 90% confidence intervals. The raw difference between conditions may not equal the estimated treatment effect due to rounding.

⁴ These covariates include age, age squared, an indicator variable for people who are identified as Hispanic on the voter file, an indicator variable for the first randomization, an indicator variable for campus, and days before the election.

⁵ In the research protocol, VoteRev specified that the effect of the program would be statistically significant if $p < 0.10$.

Robustness Checks

We conducted several robustness checks to vary two aspects of the analysis. In our first set of robustness checks, we replace canvasser fixed effects with tripler fixed effects. The main results include canvasser fixed effects to be consistent with the research protocol, but we believe that it is more appropriate to include tripler fixed effects (as we would if we were able to block on tripler in the randomization). The results using canvasser fixed effects are provided in Table 3, column 1, and the results using tripler fixed effects are provided in Table 3, column 2. While the point estimate is similar, the standard error increases due to a smaller sample size. (Triplers who only provide one triplee are dropped from the analysis using this approach.)

In our second set of robustness checks, we exclude 1,575 tripler-triplee relationships where the canvass data was identified by the Vote Rev field team as suspicious. Most of these cases were identified as potentially but not demonstrably fraudulent. For example, some canvassers submitted data that Vote Rev identified as highly likely to be fraudulent on some days but not on other days. The former is omitted in all analyses; the latter is included in the main analysis but omitted in this robustness check. A small number of tripler-triplee relationships were also excluded in this set of robustness checks because Vote Rev suspects that they may have been used in canvasser training (for example, cases where the triplee was named “John Doe” or shared a name with a trainer) or may have been inaccurately recorded by the app (for example, cases where canvass data appears in the wrong fields). The results excluding these relationships are reported in columns 3 and 4, with column 3 using canvasser fixed effects and column 4 using tripler fixed effects.

Across specifications, we see evidence that the HTOTS program was probably effective at increasing turnout by 0.4-0.5 percentage points. Our preferred specifications use tripler fixed effects, and we recommend that Vote Rev use tripler fixed effects in future HTOTS evaluations.

Table 3: The results are similar across specifications and suggest an effect of about 0.4-0.5 percentage points

	(1)	(2)	(3)	(4)
Effect size	0.44	0.39	0.52	0.44
Standard error	0.36	0.45	0.37	0.46
<i>p</i> -value	0.11	0.20	0.08	0.17
Sample size	34,990	31,274	33,585	29,896
Fixed effects	Canvasser	Tripler	Canvasser	Tripler

Notes: p-values are calculated using one-tailed tests.

Discussion

We are pleased to report that the HTOTS program was probably effective at increasing turnout in the 2022 runoff election in Georgia. Across specifications, the program appears to have increased turnout by 0.4-0.5

percentage points. While these effects are generally not statistically significant, they are also unlikely to be due to chance. We estimate that the HTOTS program mobilized 81 voters in the 2022 runoff election.⁶

We have several recommendations for future HTOTS programs:

- **Identify the locations that generated the highest number of triplees.** We encourage Vote Rev to review the canvass data to identify the characteristics of sites that generated the highest number of triplees per canvasser hour. For example, did the program generate more triplees per canvasser hour on college campuses or at other locations? Did the program generate more triplees per canvasser hour outside grocery stores or outside retail stores?
- **Identify the times that generated the highest number of triplees.** Similarly, we encourage Vote Rev to review the canvass data to identify the days of the week and the hours of the day that generated the highest number of triplees per canvasser hour. Did the program generate more triplees per canvasser hour on weekday evenings or on weekend afternoons?
- **Explore treatment-responsive subgroups.** Just as certain locations and times may produce more triplees, certain locations and times may also produce triplees for whom the program is more effective. We encourage Vote Rev to explore whether the program was more effective at mobilizing some triplees than other triplees. If so, Vote Rev could target geographies with larger concentrations of those triplees in future programs.
- **Consider asking triplers to provide the date of the election in their text message.** Vote Rev did not provide us with a canvass script, so we are not sure what canvassers told triplers. However, it is likely that a text message from a friend saying “Go vote by 12/6” will be more effective than a message from a friend saying “Go vote.” Triplers could also be encouraged to share an extremely short link where triplees could find more information about voting.
- **Determine whether this tactic is cost effective.** Vote Rev did not provide us with the cost of the program, so we cannot calculate the cost-per-vote. Regardless, the cost-per-vote of this experiment would not be representative of the cost-per-vote of a future HTOTS program because this experiment required canvassers to match triplees to the voter file, which slowed down the program. When Vote Rev piloted a HTOTS program without matching triplees to the voter file, it produced 2.6 times more triplees per hour than this program. That said, canvass programs are very expensive, and we caution that this tactic may not be cost effective even in the absence of an experiment.
- **Explore alternative experiment designs in future evaluations.** We brainstormed alternative experiment designs with Vote Rev prior to launching this experiment, and we encourage Vote Rev to work towards implementing an alternative design in future election cycles. We recommend one of the following approaches: (1) modifying the app so that it conducts the randomization in real time, *or* (2) modifying the app so that it accepts a list randomized by interaction rather than by person. In the latter, Vote Rev would upload a randomized list of five triplees for each of 15,000+ triplers – tripler 1, triplex 1, treatment; tripler 1, triplex 2, control,

⁶ This estimate is calculated as follows: 18,411 people in the treatment group * 0.0044 treatment effect = 81 voters.

etc. In both cases, we strongly recommend stratifying by tripler to increase the comparability of experiment conditions.

We hope that these results are informative as Vote Rev begins to make plans for the 2023-2024 election cycle. We thank Vote Rev for conducting this project with us, and we look forward to collaborating with you again.