**Appendices: Large-Scale Evidence for the Effectiveness of Partisan GOTV Robo Calls**

# Appendix A: Detailed Subject Characteristics and state-by-state results

Table 5 lists the number of subjects in each group by their state of residence.[[1]](#footnote-1) Table 6 lists the number of subjects in each group by the number of registered voters in their household.

Table : Number of subjects by state and treatment group

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| State | Control | T1 | T3 | T6 | Total |
| Georgia | 21,571 | 21,937 | 21,807 | 21,399 | 86,714 |
| Nebraska | 22,464 | 22,202 | 22,663 | 22,749 | 90,078 |
| New Mexico | 23,284 | 22,985 | 22,111 | 22,166 | 90,546 |
| Ohio | 23,030 | 22,804 | 22,681 | 22,882 | 91,397 |
| Pennsylvania | 21,356 | 21,534 | 21,541 | 20,844 | 85,275 |
| Virginia | 23,626 | 24,458 | 24,023 | 23,450 | 95,557 |
| Total | 135,331 | 135,920 | 134,826 | 133,490 | 539,567 |

Note: Each cell includes the number of subjects in the six states used in our study. The first column shows the number of subjects in the control group by state. The column labeled T1 shows the number of subjects in households that were assigned to be called one time. Column T3 shows the number of subjects in houseolds assigned to be called three times. Column T6 shows the number of subjects in households assigned to be called six times.

Table : Number of subjects by household size

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| Registered voters per household | Control | T1 | T3 | T6 | Total |
| 1 | 13,740 | 13,470 | 13,399 | 14,240 | 54,849 |
| 2 | 61,434 | 61,616 | 62,598 | 61,294 | 246,942 |
| 3 | 40,029 | 39,207 | 39,369 | 39,405 | 158,010 |
| 4 | 15,096 | 16,004 | 14,676 | 14,064 | 59,840 |
| 5 | 4,030 | 4,385 | 3,900 | 3,565 | 15,880 |
| 6 | 780 | 948 | 684 | 696 | 3,108 |
| 7 | 147 | 231 | 126 | 210 | 714 |
| 8 | 48 | 40 | 56 | 16 | 160 |
| 9 | 27 | 9 | 18 | 0 | 54 |
| 10 | 0 | 10 | 0 | 0 | 10 |
| Total | 135,331 | 135,920 | 134,826 | 133,490 | 539,567 |

Note: The table shows how many subjects assigned to each treatment group are members of households by household size.

Table 7 compares summary statistics for subject characteristics separately for the control and combined treatment groups. Although subjects were assigned randomly, some subject characteristics are statistically different between both groups, though the magnitude of the difference, relative to the mean of the corresponding variables, is small. Table 8 shows the summary statistics for the control and each of the three treatment groups separately. When comparing the control group to the treatment groups, there are statistically significant differences in mean age, income, and single-voter household status, though again, the magnitudes of the differences are small. Table 9 shows summary statistics for subject characteristics separately for the control and combined treatment groups in each of the six states in the experiment. Table 10 shows the intent-to-treat effects for the subjects in each state.

Table : Balance table: Subject characteristics comparing all treatment groups to control

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
|  | Control | All treatments | Difference | Prob > T |
| Mean Age | 49.9 | 50.0 | -0.1 | 0.165 |
|  | (0.05) | (0.03) | (0.06) |  |
| Mean Income | 74190.0 | 75865.7 | -1675.7 | < 0.001 |
|  | (194.9) | (120.5) | (229.1) |  |
| Percent Male | 49.3% | 49.4% | -0.2% | 0.117 |
|  | (0.1) | (0.05) | (0.1) |  |
| Percent in SVH | 10.2% | 10.2% | 0.0% | 0.871 |
|  | (0.1) | (0.1) | (0.1) |  |
| Percent voted in 2010 | 49.5% | 49.2% | 0.3% | 0.136 |
|  | (0.2) | (0.1) | (0.2) |  |
| Percent voted in 2012 | 76.7% | 77.2% | -0.5% | 0.001 |
|  | (0.1) | (0.1) | (0.1) |  |
| Observations | 135,331 | 404,236 | 539,567 |  |

Note: The table shows the average subject characteristics for the control and combined treatment groups. Percent in SVH is the percent of subjects in single-voter households. Column 3 shows the difference between the control and treatment averages for a given characteristic. Column 4 shows the p-value from the t-test for the difference in means. Standard errors are clustered by household and displayed in parentheses below each average.

Table : Balance table: Subject characteristics by treatment group and differences from the control

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
|  | T1 | T1 vs. control difference | T3 | T3 vs. control difference | T6 | T6 vs. control difference |
| Mean Age | 49.7 | 0.2 | 49.9 | 0.0 | 50.3 | -0.4 |
|  | (0.1) | (0.1) | (0.1) | (0.1) | (0.1) | (0.1) |
| Mean Income | 81804.2 | -7614.2 | 76540.6 | -2350.6 | 69085.3 | 5104.7 |
|  | (227.1) | (299.3) | (211.8) | (287.7) | (178.5) | (264.3) |
| Percent Male | 49.7% | -0.4% | 49.3% | 0.0% | 49.3% | -0.1% |
|  | (0.1) | (0.1) | (0.1) | (0.1) | (0.1) | (0.1) |
| Percent in SVH | 9.9% | 0.2% | 9.9% | 0.2% | 10.7% | -0.5% |
|  | (0.1) | (0.1) | (0.1) | (0.1) | (0.1) | (0.1) |
| Percent voted in 2010 | 49.1% | 0.3% | 49.3% | 0.2% | 49.2% | 0.3% |
|  | (0.2) | (0.2) | (0.2) | (0.2) | (0.2) | (0.2) |
| Percent voted in 2012 | 77.1% | -0.5% | 77.1% | -0.4% | 77.2% | -0.6% |
|  | (0.1) | (0.2) | (0.1) | (0.2) | (0.1) | (0.2) |
| Observations | 135,920 |  | 134,826 |  | 133,490 |  |

Note: The table shows average subject characteristics for the control and each treatment group. Percent in SVH is the percent of subjects in single-voter households. Columns 2, 4, and 6 show the treatment group averages subtracted from the control group average. Standard errors are clustered by household and displayed in parentheses under each average. Thus, the ratio of the difference in characeristics between treatment and controls and the standard error, provides the t-test for the difference in means.

Table : State-by-state balance table: Subject characteristics comparing all treatment groups to control

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| Georgia | Control | All treatments | Difference | Prob > T |
| Mean Age | 49.4 | 49.3 | 0.2 | 0.209 |
|  | (0.1) | (0.1) | (0.1) |  |
| Mean Income | 74,688.2 | 76,969.8 | -2,281.6 | 0.001 |
|  | (483.6) | (313.0) | (576.0) |  |
| Percent Male | 48.7% | 49.2% | -0.6% | 0.022 |
|  | (0.2) | (0.1) | (0.2) |  |
| Percent in SVH | 12.0% | 11.6% | 0.4% | 0.157 |
|  | (0.2) | (0.1) | (0.3) |  |
| Percent voted in 2010 | 47.6% | 47.2% | 0.4% | 0.348 |
|  | (0.4) | (0.2) | (0.4) |  |
| Percent voted in 2012 | 78.8% | 77.9% | 1.0% | 0.006 |
|  | (0.3) | (0.2) | (0.3) |  |
| Observations | 21,571.0 | 65,143.0 | 86,714.0 |  |

(continued on next page)

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| Nebraska | Control | All treatments | Difference | Prob > T |
| Mean Age | 48.2 | 47.4 | 0.8 | < 0.001 |
|  | (0.1) | (0.1) | (0.1) |  |
| Mean Income | 67,344.5 | 73,159.6 | -5,815.2 | < 0.001 |
|  | (427.4) | (269.6) | (505.3) |  |
| Percent Male | 49.3% | 50.0% | -0.7% | 0.004 |
|  | (0.2) | (0.1) | (0.2) |  |
| Percent in SVH | 11.2% | 10.6% | 0.6% | 0.025 |
|  | (0.2) | (0.1) | (0.3) |  |
| Percent voted in 2010 | 36.9% | 34.8% | 2.1% | < 0.001 |
|  | (0.4) | (0.2) | (0.4) |  |
| Percent voted in 2012 | 68.0% | 70.4% | -2.4% | < 0.001 |
|  | (0.3) | (0.2) | (0.4) |  |
| Observations | 22,464.0 | 67,614.0 | 90,078.0 |  |

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| New Mexico | Control | All treatments | Difference | Prob > T |
| Mean Age | 49.5 | 51.3 | -1.8 | < 0.001 |
|  | (0.1) | (0.1) | (0.1) |  |
| Mean Income | 70,468.3 | 64,963.6 | 5,504.7 | < 0.001 |
|  | (439.1) | (245.5) | (503.1) |  |
| Percent Male | 49.0% | 48.8% | 0.2% | 0.41 |
|  | (0.2) | (0.1) | (0.3) |  |
| Percent in SVH | 11.1% | 12.2% | -1.1% | < 0.001 |
|  | (0.2) | (0.1) | (0.3) |  |
| Percent voted in 2010 | 52.4% | 56.0% | -3.6% | < 0.001 |
|  | (0.4) | (0.2) | (0.4) |  |
| Percent voted in 2012 | 69.1% | 71.7% | -2.6% | < 0.001 |
|  | (0.3) | (0.2) | (0.4) |  |
| Observations | 23,284.0 | 67,262.0 | 90,546.0 |  |

(continued on next page)

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| Ohio | Control | All treatments | Difference | Prob > T |
| Mean Age | 49.6 | 49.4 | 0.2 | 0.178 |
|  | (0.1) | (0.1) | (0.1) |  |
| Mean Income | 69,897.6 | 72,656.5 | -2,758.9 | < 0.001 |
|  | (446.2) | (279.9) | (526.7) |  |
| Percent Male | 49.4% | 49.3% | 0.1% | 0.721 |
|  | (0.2) | (0.1) | (0.2) |  |
| Percent in SVH | 8.6% | 8.8% | -0.2% | 0.362 |
|  | (0.2) | (0.1) | (0.2) |  |
| Percent voted in 2010 | 54.7% | 53.4% | 1.3% | 0.002 |
|  | (0.4) | (0.2) | (0.4) |  |
| Percent voted in 2012 | 81.0% | 81.2% | -0.3% | 0.426 |
|  | (0.3) | (0.2) | (0.3) |  |
| Observations | 23,030.0 | 68,367.0 | 91,397.0 |  |

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| Pennsylvania | Control | All treatments | Difference | Prob > T |
| Mean Age | 50.3 | 50.3 | 0.1 | 0.691 |
|  | (0.1) | (0.1) | (0.2) |  |
| Mean Income | 78,847.2 | 76,817.2 | 2,030.0 | < 0.001 |
|  | (503.9) | (300.4) | (586.7) |  |
| Percent Male | 49.3% | 49.4% | -0.1% | 0.795 |
|  | (0.2) | (0.1) | (0.2) |  |
| Percent in SVH | 12.4% | 12.4% | 0.0% | 0.88 |
|  | (0.2) | (0.1) | (0.3) |  |
| Percent voted in 2010 | 51.4% | 51.7% | -0.3% | 0.444 |
|  | (0.4) | (0.2) | (0.5) |  |
| Percent voted in 2012 | 79.9% | 79.3% | 0.6% | 0.071 |
|  | (0.3) | (0.2) | (0.4) |  |
| Observations | 21,356.0 | 63,919.0 | 85,275.0 |  |

(continued on next page)

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| Virginia | Control | All treatments | Difference | Prob > T |
| Mean Age | 52.3 | 52.2 | 0.1 | 0.483 |
|  | (0.1) | (0.1) | (0.1) |  |
| Mean Income | 84,136.4 | 89,808.2 | -5,671.7 | < 0.001 |
|  | (533.4) | (325.6) | (624.9) |  |
| Percent Male | 50.0% | 49.9% | 0.1% | 0.697 |
|  | (0.2) | (0.1) | (0.2) |  |
| Percent in SVH | 6.1% | 5.9% | 0.2% | 0.358 |
|  | (0.2) | (0.1) | (0.2) |  |
| Percent voted in 2010 | 53.4% | 52.0% | 1.4% | 0.001 |
|  | (0.4) | (0.2) | (0.4) |  |
| Percent voted in 2012 | 83.2% | 82.2% | 1.0% | 0.001 |
|  | (0.3) | (0.2) | (0.3) |  |
| Observations | 23,626.0 | 71,931.0 | 95,557.0 |  |

Note: The six panels of this table show the average subject characteristics for the control and combined treatment groups for each state in the experiment. Percent in SVH is the percent of subjects in single-voter households. Column 3 shows the difference between the control and treatment averages for a given characteristic. Column 4 shows the p-value from the t-test for the difference in means. Standard errors are clustered by household and displayed in parentheses below each average.

Table : State-by-State intent-to-treat effects

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
|  | (1) | (2) | (3) | (4) |
| Variables | voted | voted | voted | voted |
|  |  |  |  |  |
| Population | Georgia | Georgia | Nebraska | Nebraska |
|  |  |  |  |  |
| T1 |  | 0.0113\*\* |  | -0.00282 |
|  |  | (0.00482) |  | (0.00458) |
| T3 |  | -0.00214 |  | -0.00321 |
|  |  | (0.00484) |  | (0.00456) |
| T6 |  | -0.00873\* |  | -0.0126\*\*\* |
|  |  | (0.00490) |  | (0.00456) |
| All Treatments | 0.000230 |  | -0.00624\* |  |
|  | (0.00395) |  | (0.00372) |  |
|  |  |  |  |  |
| Subject controls | Yes | Yes | Yes | Yes |
| Observations | 86,714 | 86,714 | 90,078 | 90,078 |
| R-squared | 0.245 | 0.246 | 0.335 | 0.335 |
| F-test | 1376 | 1279 | 2265 | 2104 |
| Prob > F | < 0.001 | < 0.001 | < 0.001 | < 0.001 |

(continued on next page)

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
|  | (1) | (2) | (3) | (4) |
| Variables | voted | voted | voted | voted |
|  |  |  |  |  |
| Population | New Mexico | New Mexico | Ohio | Ohio |
|  |  |  |  |  |
| T1 |  | 0.00192 |  | -0.0102\*\* |
|  |  | (0.00436) |  | (0.00495) |
| T3 |  | 0.0200\*\*\* |  | 0.00438 |
|  |  | (0.00447) |  | (0.00497) |
| T6 |  | 0.0446\*\*\* |  | -0.0286\*\*\* |
|  |  | (0.00450) |  | (0.00498) |
| All Treatments | 0.0216\*\*\* |  | -0.0115\*\*\* |  |
|  | (0.00361) |  | (0.00406) |  |
|  |  |  |  |  |
| Subject controls | Yes | Yes | Yes | Yes |
| Observations | 90,546 | 90,546 | 91,397 | 91,397 |
| R-squared | 0.354 | 0.355 | 0.230 | 0.230 |
| F-test | 2246 | 2108 | 1437 | 1340 |
| Prob > F | < 0.001 | < 0.001 | < 0.001 | < 0.001 |

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
|  | (1) | (2) | (3) | (4) |
| Variables | voted | voted | voted | voted |
|  |  |  |  |  |
| Population | Pennsylvania | Pennsylvania | Virginia | Virginia |
|  |  |  |  |  |
| T1 |  | -0.000757 |  | 0.00359 |
|  |  | (0.00501) |  | (0.00448) |
| T3 |  | 0.0148\*\*\* |  | 0.00761\* |
|  |  | (0.00500) |  | (0.00443) |
| T6 |  | 0.0145\*\*\* |  | -0.00294 |
|  |  | (0.00504) |  | (0.00445) |
| All Treatments | 0.00946\*\* |  | 0.00277 |  |
|  | (0.00410) |  | (0.00364) |  |
|  |  |  |  |  |
| Subject controls | Yes | Yes | Yes | Yes |
| Observations | 85,275 | 85,275 | 95,557 | 95,557 |
| R-squared | 0.248 | 0.248 | 0.337 | 0.337 |
| F-test | 1339 | 1244 | 2541 | 2362 |
| Prob > F | < 0.001 | < 0.001 | < 0.001 | < 0.001 |

Note: The three panels of this table show OLS estimates for the difference in voting rates between subjects in the treatment group and subjects in the control group for ecah state. The subset of subjects used to estimate each specification is shown with the “Population” label. Subject controls include whether or not a subject voted in the 2010 or 2012 general elections or both and subject age, gender, education, income, number of subjects in the household, and state of residence. Standard errors are clustered by household and displayed in parentheses. \*\*\* p<0.01, \*\* p<0.05, \* p<0.1

# Appendix B: Message Scripts

Call scripts were used based on the following schedule:

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| Group | Oct 30 | Oct 31 | Nov 1 | Nov 1 | Nov 3 | Nov 4 |
| Control | - | - | - | - | - | - |
| T1 | - | - | - | - | - | Script 6 |
| T3 | - | - | - | Script 4 | Script 5 | Script 6 |
| T6 | Script 1 | Script 2 | Script 3 | Script 4 | Script 5 | Script 6 |

**Script #1:**

Hello, this is \_\_\_\_\_\_\_\_\_ calling with \_\_\_\_\_\_ PAC reminding you to vote on November 4th for our Republican Candidates. Help send a message to Washington that you want your leaders to focus on conservative solutions for Jobs and the Economy. Vote on Tuesday, November 4th for the Republican Ticket.

Thank you. This call was paid for by \_\_\_\_\_\_ PAC. Not authorized by any candidate or candidate's committee. ###-###-####

**Script #2:**

Hello, this is \_\_\_\_\_\_\_\_\_ calling with \_\_\_\_\_\_ PAC.

Have you, or someone you know, lost your healthcare coverage or seen your premiums skyrocket under Obama's Healthcare Law?

Make your voice heard by voting Republican and letting Congress know you want government out of our healthcare. Remember to vote for our Republican team on Tuesday, November 4th.

Thank you. This call was paid for by \_\_\_\_\_\_ PAC. Not authorized by any candidate or candidate's committee. ###-###-####

**Script #3:**

Hello, this is \_\_\_\_\_\_\_\_\_ calling with \_\_\_\_\_\_ PAC.

We need your help to stop the gridlock in Washington D.C. and get our government leaders to focus on the issues important to all of us, not just special interest groups.

Send Congress a message that you want conservative leadership who will look out for you. Vote for the Republican Ticket on Tuesday, November 4th.

Thank you. This call was paid for by \_\_\_\_\_\_ PAC. Not authorized by any candidate or candidate's committee. ###-###-####

**Script #4:**

Hello, this is \_\_\_\_\_\_\_\_\_ calling with \_\_\_\_\_\_ PAC reminding you to vote on November 4th for your Republican Candidates.

With our National Debt spiraling out of control and a weak economy, we need strong candidates more than ever, and this election we have a chance to do something about it.

You can send a message to Washington by voting Republican on Tuesday November 4th.

Thank you. This call was paid for by \_\_\_\_\_\_ PAC. Not authorized by any candidate or candidate's committee. ###-###-####

**Script #5:**

Hello, this is \_\_\_\_\_\_\_\_\_ calling with \_\_\_\_\_\_ PAC.

This past year, Americans have seen their healthcare premiums skyrocket, while thousands have lost their coverage. We were promised by the President, and nearly every Democrat in Congress, that costs would go down and everyone would be able to keep the coverage they liked. The fact is: they lied.

This Election Day, we can change Washington.

Make your voice heard and tell Congress you want them to get government out of our healthcare by voting for your Republican Ticket on Tuesday, November 4th.

Thank you. This call was paid for by \_\_\_\_\_\_ PAC. Not authorized by any candidate or candidate's committee. ###-###-####

**Script #6:**

Hello, this is \_\_\_\_\_\_\_\_\_ calling with \_\_\_\_\_\_ PAC.

Today is Election Day and we have a chance to elect Republican candidates who understand that balanced budgets and lower taxes are far better for all of us than increasing the size of government.

Together, we can ensure Congress and our elected leaders focus on what matters.

Be sure to vote for the Republican Ticket *today*, Tuesday, November 4th.

Thank you. This call was paid for by \_\_\_\_\_\_ PAC. Not authorized by any candidate or candidate's committee. ###-###-####

# Appendix C: Detailed Treatment Call Outcomes

The following three tables show the number of calls that resulted in a live answer, answering machine, or either one, for subjects (not households) in each treatment group.

|  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- |
| Number of calls with live answer | T1 | | T3 | | T6 | | Total | |
| 0 | 82,565 | 61% | 54,529 | 40% | 43,858 | 33% | 180,952 | 45% |
| 1 | 53,355 | 39% | 23,907 | 18% | 12,861 | 10% | 90,123 | 22% |
| 2 | 0 | 0% | 26,847 | 20% | 12,079 | 9% | 38,926 | 10% |
| 3 | 0 | 0% | 29,543 | 22% | 13,167 | 10% | 42,710 | 11% |
| 4 | 0 | 0% | 0 | 0% | 15,382 | 12% | 15,382 | 4% |
| 5 | 0 | 0% | 0 | 0% | 16,935 | 13% | 16,935 | 4% |
| 6 | 0 | 0% | 0 | 0% | 19,208 | 14% | 19,208 | 5% |
| Total | 135,920 | 100% | 134,826 | 100% | 133,490 | 100% | 404,236 |  |

|  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- |
| Number of calls with AM | T1 | | T3 | | T6 | | Total | |
| 0 | 89,925 | 66% | 69,137 | 51% | 59,143 | 44% | 218,205 | 54% |
| 1 | 45,995 | 34% | 25,626 | 19% | 17,459 | 13% | 89,080 | 22% |
| 2 | 0 | 0% | 20,814 | 15% | 14,170 | 11% | 34,984 | 9% |
| 3 | 0 | 0% | 19,249 | 14% | 11,528 | 9% | 30,777 | 8% |
| 4 | 0 | 0% | 0 | 0% | 10,452 | 8% | 10,452 | 3% |
| 5 | 0 | 0% | 0 | 0% | 10,051 | 8% | 10,051 | 2% |
| 6 | 0 | 0% | 0 | 0% | 10,687 | 8% | 10,687 | 3% |
| Total | 135,920 | 100% | 134,826 | 100% | 133,490 | 100% | 404,236 |  |

|  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- |
| Number of calls with a live answer or AM | T1 | | T3 | | T6 | | Total | |
| 0 | 36,570 | 27% | 33,451 | 25% | 31,302 | 23% | 101,323 | 25% |
| 1 | 99,350 | 73% | 3,997 | 3% | 2,852 | 2% | 106,199 | 26% |
| 2 | 0 | 0% | 4,900 | 4% | 1,886 | 1% | 6,786 | 2% |
| 3 | 0 | 0% | 92,478 | 69% | 2,012 | 2% | 94,490 | 23% |
| 4 | 0 | 0% | 0 | 0% | 2,890 | 2% | 2,890 | 1% |
| 5 | 0 | 0% | 0 | 0% | 4,969 | 4% | 4,969 | 1% |
| 6 | 0 | 0% | 0 | 0% | 87,579 | 66% | 87,579 | 22% |
| Total | 135,920 | 100% | 134,826 | 100% | 133,490 | 100% | 404,236 |  |

The following ten tables show call outcomes for each phone number (household – not subjects) by treatment, call number, and state.

|  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- |
|  | T1 - Call 1 - Tuesday, Nov. 4th 09:00 Local | | | | | | | |
|  | VA | PA | GA | OH | NE | NM | Total | Averages |
| Live | 5,097 | 3,441 | 3,726 | 4,144 | 4,037 | 3,527 | 23,972 | 38.41% |
| Ans. mach. | 3,445 | 4,556 | 3,230 | 2,856 | 3,556 | 2,996 | 20,639 | 33.07% |
| Fax | 97 | 138 | 161 | 116 | 104 | 123 | 739 | 1.18% |
| Busy | 97 | 57 | 85 | 66 | 49 | 82 | 436 | 0.70% |
| No answer | 840 | 1086 | 877 | 836 | 998 | 683 | 5,320 | 8.52% |
| Other | 834 | 1,176 | 2,256 | 2,311 | 1,696 | 3,039 | 11,312 | 18.12% |
| Total | 10,410 | 10,454 | 10,335 | 10,329 | 10,440 | 10,450 | 62,418 | 100.00% |
|  |  |  |  |  |  |  |  |  |
|  | T3 - Call 1 - Saturday, Nov. 1st 17:00 Local | | | | | | | |
|  | VA | PA | GA | OH | NE | NM | Total | Averages |
| Live | 5,891 | 4,029 | 4,377 | 4,411 | 4,695 | 3,834 | 27,237 | 43.64% |
| Ans. mach. | 2,572 | 3,988 | 2,453 | 2,819 | 2,868 | 2,202 | 16,902 | 27.08% |
| Fax | 98 | 145 | 182 | 110 | 108 | 132 | 775 | 1.24% |
| Busy | 66 | 45 | 66 | 50 | 68 | 109 | 404 | 0.65% |
| No answer | 630 | 1080 | 958 | 712 | 1049 | 825 | 5,254 | 8.42% |
| Other | 1,142 | 1,170 | 2,300 | 2,223 | 1,653 | 3,349 | 11,837 | 18.97% |
| Total | 10,399 | 10,457 | 10,336 | 10,325 | 10,441 | 10,451 | 62,409 | 100.00% |
|  |  |  |  |  |  |  |  |  |
|  | T3 - Call 2 - Monday, Nov. 3rd 18:00 Local | | | | | | | |
|  | VA | PA | GA | OH | NE | NM | Total | Averages |
| Live | 5,002 | 3,875 | 3,829 | 3,889 | 4,185 | 3,426 | 24,206 | 38.80% |
| Ans. mach. | 3,507 | 4,160 | 3,048 | 3,378 | 3,333 | 2,623 | 20,049 | 32.14% |
| Fax | 104 | 159 | 203 | 110 | 97 | 133 | 806 | 1.29% |
| Busy | 66 | 62 | 97 | 67 | 80 | 97 | 469 | 0.75% |
| No answer | 561 | 1013 | 920 | 707 | 960 | 821 | 4,982 | 7.99% |
| Other | 1,156 | 1,188 | 2,236 | 2,171 | 1,780 | 3,341 | 11,872 | 19.03% |
| Total | 10,396 | 10,457 | 10,333 | 10,322 | 10,435 | 10,441 | 62,384 | 100.00% |

|  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- |
|  | T3 - Call 3 - Tuesday, Nov. 4th 09:00 Local | | | | | | | |
|  | VA | PA | GA | OH | NE | NM | Total | Averages |
| Live | 5,343 | 3,490 | 3,830 | 4,041 | 3,844 | 3,411 | 23,959 | 38.42% |
| Ans. mach. | 3,059 | 4,473 | 2,882 | 3,110 | 3,491 | 2,573 | 19,588 | 31.41% |
| Fax | 100 | 151 | 188 | 103 | 112 | 116 | 770 | 1.23% |
| Busy | 74 | 49 | 80 | 58 | 67 | 125 | 453 | 0.73% |
| No answer | 693 | 1133 | 1,098 | 824 | 1232 | 866 | 5,846 | 9.37% |
| Other | 1,123 | 1,157 | 2,252 | 2,183 | 1,685 | 3,350 | 11,750 | 18.84% |
| Total | 10,392 | 10,453 | 10,330 | 10,319 | 10,431 | 10,441 | 62,366 | 100.00% |
|  | T6 - Call 1 - Thursday, Oct. 30th 18:00 Local | | | | | | | |
|  | VA | PA | GA | OH | NE | NM | Total | Averages |
| Live | 5,627 | 3,930 | 4,079 | 4,169 | 4,811 | 3,963 | 26,579 | 42.60% |
| Ans. mach. | 2,688 | 3,989 | 2,498 | 3,073 | 2,939 | 2,719 | 17,906 | 28.70% |
| Fax | 108 | 125 | 147 | 98 | 102 | 129 | 709 | 1.14% |
| Busy | 92 | 64 | 95 | 68 | 52 | 68 | 439 | 0.70% |
| No answer | 679 | 844 | 1,085 | 724 | 972 | 567 | 4,871 | 7.81% |
| Other | 1,209 | 1,499 | 2,429 | 2,196 | 1,559 | 2,995 | 11,887 | 19.05% |
| Total | 10,403 | 10,451 | 10,333 | 10,328 | 10,435 | 10,441 | 62,391 | 100.00% |
|  |  |  |  |  |  |  |  |  |
|  | T6 - Call 2 - Friday, Oct. 31st 17:30 Local | | | | | | | |
|  | VA | PA | GA | OH | NE | NM | Total | Averages |
| Live | 6,028 | 3,641 | 4,064 | 4,223 | 4,546 | 3,780 | 26,282 | 42.13% |
| Ans. mach. | 2,255 | 4,263 | 2,476 | 3,019 | 3,145 | 2,887 | 18,045 | 28.93% |
| Fax | 112 | 116 | 149 | 99 | 103 | 123 | 702 | 1.13% |
| Busy | 55 | 59 | 83 | 45 | 59 | 63 | 364 | 0.58% |
| No answer | 732 | 1143 | 1,127 | 689 | 1010 | 607 | 5,308 | 8.51% |
| Other | 1,220 | 1,229 | 2,431 | 2,252 | 1,568 | 2,979 | 11,679 | 18.72% |
| Total | 10,402 | 10,451 | 10,330 | 10,327 | 10,431 | 10,439 | 62,380 | 100.00% |

|  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- |
|  | T6 - Call 3 - Saturday, Nov. 1st 11:00 Local | | | | | | | |
|  | VA | PA | GA | OH | NE | NM | Total | Averages |
| Live | 5,943 | 3,740 | 4,158 | 4,191 | 4,265 | 3,518 | 25,815 | 41.40% |
| Ans. mach. | 2,313 | 4,117 | 2,364 | 3,015 | 3,144 | 3,102 | 18,055 | 28.95% |
| Fax | 102 | 121 | 149 | 94 | 100 | 133 | 699 | 1.12% |
| Busy | 70 | 72 | 99 | 59 | 68 | 62 | 430 | 0.69% |
| No answer | 757 | 714 | 1,115 | 716 | 1034 | 628 | 4,964 | 7.96% |
| Other | 1,210 | 1,686 | 2,443 | 2,246 | 1,818 | 2,994 | 12,397 | 19.88% |
| Total | 10,395 | 10,450 | 10,328 | 10,321 | 10,429 | 10,437 | 62,360 | 100.00% |
|  |  |  |  |  |  |  |  |  |
|  | T6 - Call 4 - Saturday, Nov. 1st 17:00 Local | | | | | | | |
|  | VA | PA | GA | OH | NE | NM | Total | Averages |
| Live | 6,204 | 3,754 | 4,331 | 4,422 | 4,768 | 4,015 | 27,494 | 44.11% |
| Ans. mach. | 2,060 | 4,055 | 2,172 | 2,771 | 2,841 | 2,611 | 16,510 | 26.49% |
| Fax | 95 | 132 | 153 | 100 | 107 | 124 | 711 | 1.14% |
| Busy | 64 | 43 | 118 | 54 | 66 | 96 | 441 | 0.71% |
| No answer | 796 | 1211 | 1,115 | 711 | 1006 | 622 | 5,461 | 8.76% |
| Other | 1,172 | 1,251 | 2,437 | 2,259 | 1,635 | 2,965 | 11,719 | 18.80% |
| Total | 10,391 | 10,446 | 10,326 | 10,317 | 10,423 | 10,433 | 62,336 | 100.00% |
|  |  |  |  |  |  |  |  |  |
|  | T6 - Call 5 - Monday, Nov. 3rd 18:00 Local | | | | | | | |
|  | VA | PA | GA | OH | NE | NM | Total | Averages |
| Live | 4,916 | 3,763 | 3,939 | 3,999 | 4,273 | 3,638 | 24,528 | 39.38% |
| Ans. mach. | 3,355 | 4,126 | 2,593 | 3,219 | 3,303 | 3,023 | 19,619 | 31.50% |
| Fax | 111 | 129 | 157 | 109 | 95 | 129 | 730 | 1.17% |
| Busy | 117 | 67 | 155 | 78 | 70 | 77 | 564 | 0.91% |
| No answer | 753 | 1037 | 1,094 | 704 | 956 | 575 | 5,119 | 8.22% |
| Other | 1,125 | 1,314 | 2,382 | 2,200 | 1,723 | 2,984 | 11,728 | 18.83% |
| Total | 10,377 | 10,436 | 10,320 | 10,309 | 10,420 | 10,426 | 62,288 | 100.00% |

|  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- |
|  | T6 - Call 6 - Tuesday, Nov. 4th 09:00 Local | | | | | | | |
|  | VA | PA | GA | OH | NE | NM | Total | Averages |
| Live | 5,511 | 3,400 | 3,848 | 3,952 | 3,973 | 3,631 | 24,315 | 39.05% |
| Ans. mach. | 2,599 | 4,374 | 2,539 | 3,125 | 3,462 | 2,953 | 19,052 | 30.60% |
| Fax | 110 | 126 | 152 | 96 | 100 | 132 | 716 | 1.15% |
| Busy | 119 | 75 | 130 | 53 | 72 | 80 | 529 | 0.85% |
| No answer | 916 | 1207 | 1,258 | 846 | 1172 | 632 | 6,031 | 9.69% |
| Other | 1,120 | 1,252 | 2,392 | 2,228 | 1,638 | 2,996 | 11,626 | 18.67% |
| Total | 10,375 | 10,434 | 10,319 | 10,300 | 10,417 | 10,424 | 62,269 | 100.00% |

# Appendix D: Follow-up call and selection effects

One concern about the estimates in GOTV experiments is the difficulty of measuring treatment effectiveness when we do not observe perfect compliance – not all subjects that we want to treat are actually treated. In this experiment that means that not all subjects even answered any calls – let alone listened to the message. This does not undermine the validity of an intent-to-treat comparison (as reported in the main body of the paper), but it can be vexing for practitioners who must decide when and how-much to try to contact the uncontacted voters that they’re trying to mobilize. In this appendix we report some analysis that uses an unrelated call to the subject group from after the election to attempt to obtain an estimate that is closer to the treatment on the treated effect. We do not claim that the method reported here is as effective as a standard design with placebo calls to the control group a la Gerber, Green, Kaplan, and Kern (2010). Similarly, it involves more work than simply calculating the local average treatment effect (LATE) or complier average causal effect (CACE) as discussed in Appendix E. However, we believe that the method described here has is an interesting way to use ancillary data to divide the subject population into subgroups that are more directly comparable and may have applications in other contexts where a placebo design is not feasible.

Since the decision not to comply is correlated with subject characteristics and since we do not know which subjects in the control group would have answered the phone if they had been called, a simple comparison of subjects who complied to subjects in the control group would result in biased estimates.[[2]](#footnote-2) The correlation between answering treatment calls and voting is clearly demonstrated by subjects’ historical voting records. We find that 51.7 percent of 2014 treatment live answerers voted in the 2010 election but only 46.1 percent of 2014 treatment non-answerers voted in 2010.[[3]](#footnote-3)

As shown in Table 2 of the manuscript, 55 percent of the subjects in the treatment groups were in households that selected themselves into “live answerer” treatment because someone in that household answered at least one call live. The calls to the remaining subjects in the treatment groups went to answering machines, had busy signals, or were otherwise non-connections.[[4]](#footnote-4)

The control group includes subjects who would have been answerers and subjects who would have been non-answerers if they had been called. However, we do not know which subjects in the control group belong to the answerers or non-answerers groups had they been called. Subjects in the control group have a higher voter participation rate than treatment non-answerers. Live answerers in the treatment (referred to as “treatment live answerers” or simply “treatment answerers” in this appendix to distinguish from the follow-up call) have a higher propensity to vote than the general population, as represented by the control group. This suggests that answering the phone is correlated with the same traits that make a subject more likely to vote. For example, a subject may both answer a phone call and vote in the election because that subject is extroverted.[[5]](#footnote-5)

To address this selection effect issue and estimate the causal effect of robo calls on voter participation, we evaluated data from a follow-up call that our partner organization placed to households in all treatment and control groups.[[6]](#footnote-6) This call contained a partisan political message not related to voting in the 2014 general election. We define as follow-up answerers subjects residing in households where the follow-up call resulted in a live answer. We classify all other subjects as follow-up non-answerers. Table 11 reports answering rates from this follow-up call, for the treatment and control groups. In both the treatment and control groups, 28 percent of subjects were follow-up answerers.[[7]](#footnote-7)

Table : Follow-up answering rates in control and treatment groups

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| Follow-up result | Control | T1 | T3 | T6 | Total |
| Live answer | 28%  (37,622) | 28%  (38,282) | 28%  (37,112) | 28%  (37,948) | 28%  (150,964) |
| Answering machine | 41%  (55,900) | 42%  (57,314) | 41%  (55,212) | 40%  (53,320) | 41%  (221,746) |
| No answer | 31%  (41,809) | 30%  (40,324) | 32%  (42,502) | 32%  (42,222) | 31%  (166,857) |
| Total | 100%  (135,331) | 100%  (135,920) | 100%  (134,826) | 100%  (133,490) | 100%  (539,567) |

Note: The table shows the percent of subjects in each group for whom the follow-up call to their household resulted in a live answer, an answering machine responding, or in no answer. These follow-up call outcome categories are listed in the rows. Percents are relative to the total number of subjects in each treatment group. The raw number of subjects in each cell is given in parentheses.

This suggests that answering the follow-up call is a reasonable proxy for being inclined to answer the treatment call. There are not meaningful differences among treatment groups in terms of likeliness to answer the follow-up call. Table 12 shows a cross tabulation for whether subjects in the treatment groups answered the follow-up call and/or at least one treatment call. Among follow-up answerers across all groups, 82 percent of subjects were also treatment live answerers. Among follow-up non-answerers, only 45 percent were treatment live answerers. Figure 2 shows this relationship visually.

Table : Cross tabulation between treatment live answerers and follow-up answerers by treatment group

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
|  | T1 | | T3 | | T6 | |
|  | Treatment live answerer | Treatment non-answerer | Treatment live answerer | Treatment non-answerer | Treatment live answerer | Treatment non-answerer |
| Follow-up answerer | 63%  (24,033) | 37%  (14,249) | 89%  (32,904) | 11%  (4,208) | 94%  (35,673) | 6%  (2,275) |
| Follow-up non-answerer | 30%  (29,322) | 70%  (68,316) | 49%  (47,393) | 51%  (50,321) | 56%  (53,959) | 44%  (41,583) |
| Total | 39%  (53,355) | 61%  (82,565) | 60%  (80,297) | 40%  (54,529) | 67%  (89,632) | 33%  (43,858) |

Note: The table shows the percent of subjects in each treatment group who were treatment live answerers or treatment non-answerers for follow-up answerers and follow-up non-answerers. Percents are relative to the total number of subjects in each treatment group conditional on the follow-up call outcome. The raw number of subjects in each cell is given in parentheses.

Figure 2: Treatment success rates by follow up call response



Note: Column height indicates the percentage of subjects in a category. Subjects in the treatment group are categorized based on whether or not they answered a treatment call, and whether or not they answered the follow up-call. Subjects in the control group are categorized based on whether or not they answered the follow-up call.

We can use this data to obtain a treatment effect estimate that is closer to the treatment-on-the-treated effect by restricting the sample to follow-up answerers and comparing the treatment and control group within this subsample.[[8]](#footnote-8) This approach improves the comparability of control and treatment groups, as shown in Table 13, Table 14, and Figure 3. The figure shows the voting rates for follow-up answerers and non-answerers in the treatment and control. Table 13 is the balance table for subjects based on the outcome of the follow-up call. Table 14 is the balance table for the treatment and control groups within each follow-up call outcome (answering the call, answering machine, no answer). Within each of these groups, differences in the individual characteristics of the treatment group and control group are smaller than the differences between follow-up answerers and non-answerers.

Table : Balance table: Subject characteristics based on the outcome of the follow-up call

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
|  | Follow-up  answerers | Follow-up AM | Follow-up non-answerers | Prob > T |
| Mean Age | 52.7 | 50.1 | 47.1 | < 0.001 |
|  | (0.05) | (0.04) | (0.04) |  |
| Mean Income | 74,262 | 79,960 | 70,305 | < 0.001 |
|  | (196.9) | (167.4) | (166.2) |  |
| Percent Male | 49.3% | 49.6% | 49.3% | 0.008 |
|  | (0.1) | (0.1) | (0.1) |  |
| Percent in SVH | 9.0% | 8.6% | 13.3% | < 0.001 |
|  | (0.1) | (0.1) | (0.1) |  |
| Percent voted in 2010 | 51.8% | 51.4% | 44.2% | < 0.001 |
|  | (0.1) | (0.1) | (0.1) |  |
| Percent voted in 2012 | 78.3% | 79.1% | 73.1% | < 0.001 |
|  | (0.1) | (0.1) | (0.1) |  |
| Observations | 150,964 | 221,746 | 166,857 |  |

Note: The table shows average subject characteristics for follow-up answerers, subjects for whom the call went to an answering machine, and follow-up non-answerers in the entire subject pool. Percent in SVH is the percent of subjects in single-voter households. Column 4 shows the p-value from the t-test for the difference in means. Standard errors are clustered by household and displayed in parentheses.

Table : Balance table: Subject characteristics comparing all treatment groups to control based on the outcome of the follow-up call

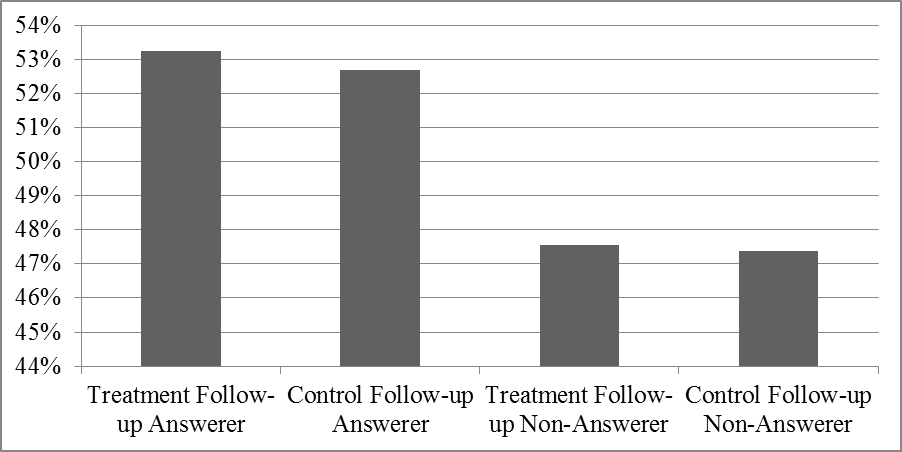
|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| Follow-up  answerers | Control | All treatments | Difference | Prob > T |
| Mean Age | 52.7 | 52.7 | 0.0 | 0.943 |
|  | (0.11) | (0.06) | (0.12) |  |
| Mean Income | 72309 | 74915 | -2605 | < 0.001 |
|  | (368.2) | (232.0) | (435.2) |  |
| Percent Male | 49.3% | 49.3% | 0.1% | 0.681 |
|  | (0.2) | (0.1) | (0.2) |  |
| Percent in SVH | 8.9% | 9.0% | 0.0% | 0.807 |
|  | (0.2) | (0.1) | (0.2) |  |
| Percent voted in 2010 | 52.1% | 51.7% | 0.3% | 0.308 |
|  | (0.3) | (0.2) | (0.3) |  |
| Percent voted in 2012 | 78.0% | 78.4% | -0.4% | 0.097 |
|  | (0.2) | (0.1) | (0.3) |  |
| Observations | 37,622 | 113,342 | 150,964 |  |

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| Follow-up answering machine | Control | All treatments | Difference | Prob > T |
| Mean Age | 50.0 | 50.2 | -0.1 | 0.148 |
|  | (0.08) | (0.05) | (0.09) |  |
| Mean Income | 79029 | 80276 | -1246 | 0.001 |
|  | (321.2) | (195.9) | (376.3) |  |
| Percent Male | 49.4% | 49.6% | -0.2% | 0.113 |
|  | (0.1) | (0.1) | (0.1) |  |
| Percent in SVH | 8.5% | 8.6% | -0.2% | 0.304 |
|  | (0.1) | (0.1) | (0.1) |  |
| Percent voted in 2010 | 51.9% | 51.2% | 0.7% | 0.016 |
|  | (0.2) | (0.1) | (0.3) |  |
| Percent voted in 2012 | 78.8% | 79.2% | -0.4% | 0.072 |
|  | (0.2) | (0.1) | (0.2) |  |
| Observations | 55,900 | 165,846 | 221,746 |  |

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| Follow-up non-answerers | Control | All treatments | Difference | Prob > T |
| Mean Age | 47.1 | 47.1 | -0.1 | 0.489 |
|  | (0.09) | (0.05) | (0.10) |  |
| Mean Income | 69234 | 70667 | -1433 | < 0.001 |
|  | (314.4) | (195.3) | (370.2) |  |
| Percent Male | 49.1% | 49.4% | -0.3% | 0.168 |
|  | (0.2) | (0.1) | (0.2) |  |
| Percent in SVH | 13.5% | 13.3% | 0.2% | 0.358 |
|  | (0.2) | (0.1) | (0.2) |  |
| Percent voted in 2010 | 43.9% | 44.2% | -0.3% | 0.279 |
|  | (0.3) | (0.2) | (0.3) |  |
| Percent voted in 2012 | 72.6% | 73.3% | -0.7% | 0.011 |
|  | (0.2) | (0.1) | (0.3) |  |
| Observations | 41,809 | 125,048 | 166,857 |  |

Note: The panels in this table show average subject characteristics for the control and combined treatment groups based on the outcome of the follow-up call. The top panel is the data for follow-up answerers, the middle panel is the data for subjects for whom the call went to an answering machine, and the bottom panel is the data for follow-up non-answerers. Percent in SVH is the percent of subjects in single-voter households. Column 3 shows the difference between the follow-up non-answerers and follow-up answerers averages for a given characteristic. Column 4 shows the p-value from the t-test for the difference in means. Standard errors are clustered by household and displayed in parentheses.

Figure 3: Voting rates by follow-up call outcome and subject group



Note: Column height indicates the percentage of subjects within each category who voted in the 2014 general election. Treatment live answerers and non-answerers are combined into groups of follow-up answerers and follow-up non-answerers.

Table 15 reports intent-to-treat effects from the subset of follow-up answerers in the treatment and control groups.[[9]](#footnote-9) Columns 1 and 2 report results for subjects who answered the follow-up with a live answer. Columns 3 and 4 report results for subjects whose follow-up call went to an answering machine. Columns 5 and 6 report results for subjects whose follow-up call was neither answered live nor went to an answering machine.

Table : Intent-to-treat effect among subjects by follow up call outcome

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
|  | (1) | (2) | (3) | (4) | (5) | (6) |
| Variables | voted | voted | voted | voted | voted | voted |
|  |  |  |  |  |  |  |
| Population | Follow-up answerer | Follow-up answerer | Follow-up AM | Follow-up AM | Follow-up non-answerer | Follow-up non-answerer |
|  |  |  |  |  |  |  |
| T1 |  | 0.00534 |  | -0.00318 |  | 0.00237 |
|  |  | (0.00419) |  | (0.00345) |  | (0.00408) |
| T3 |  | 0.0103\*\* |  | 0.00488 |  | 0.00738\* |
|  |  | (0.00424) |  | (0.00348) |  | (0.00405) |
| T6 |  | 0.00103 |  | -0.000844 |  | 0.00498 |
|  |  | (0.00420) |  | (0.00351) |  | (0.00405) |
| All treatments | 0.00553 |  | 0.000253 |  | 0.00495 |  |
|  | (0.00344) |  | (0.00284) |  | (0.00331) |  |
| Constant | 0.527\*\*\* | 0.527\*\*\* | 0.519\*\*\* | 0.519\*\*\* | 0.413\*\*\* | 0.413\*\*\* |
|  | (0.00298) | (0.00298) | (0.00245) | (0.00245) | (0.00287) | (0.00287) |
|  |  |  |  |  |  |  |
| Subject controls | No | No | No | No | No | No |
| Observations | 150,964 | 150,964 | 221,746 | 221,746 | 166,857 | 166,857 |
| R-squared | 2.293e-05 | 6.594e-05 | 4.826e-08 | 3.477e-05 | 1.896e-05 | 3.177e-05 |
| F-test | 2.580 | 2.458 | 0.00793 | 1.904 | 2.238 | 1.244 |
| Prob > F | 0.108 | 0.0609 | 0.929 | 0.127 | 0.135 | 0.292 |

Note: The table shows OLS estimates for the difference in voting rates between subjects in the treatment group and subjects in the control group. The intercept represents the baseline voting rate in the control group. The subset of subjects used to estimate each specification is described in the population row. Standard errors are clustered by household and displayed in parentheses.\*\*\* p<0.01, \*\* p<0.05, \* p<0.1

Compared to results from the unrestricted sample in Table 17, both the base voting rate and the average treatment effect are larger among the subset of follow-up answerers. Among all follow-up answers, the overall treatment coefficient is 0.00553, implying a more than half a percentage point increase in the probability of voter participation of a household member, when a subject in that household is a member of the treatment group. This estimate is not quite statistically significant at the ten percent level.

Column 2 reports results by treatment. The treatment effects for T1 and T6 are both positive and larger than the corresponding estimates in Table 17, while statistically insignificant. The coefficient for T3 is 0.0103 and statistically significant at the five percent level.[[10]](#footnote-10) The coefficients are jointly statistically significant at the ten percent level. The estimate for T3 implies that voter participation for live follow-up answerers in the T3 group is 1.03 percentage points higher than for live follow-up answerers in the control group. The magnitude of this estimate implies an additional ten individuals participating in an election for every 1,000 subjects called.[[11]](#footnote-11)

In Table 15, Columns 3 and 4 report the treatment effects for those subjects whose follow-up call went to an answering machine. These coefficients are all smaller and statistically insignificantly different from the control group. Similarly, Columns 5 and 6 show that for those who did not receive the follow-up call either live or via an answering machine, the point estimates on the treatment variables are small and are statistically insignificant, except for T3, where the size of the coefficient is 0.0074 and is statistically significant at the ten percent level.[[12]](#footnote-12) The estimates in these columns are not jointly significant.

Restricting the sample by the result of the follow-up call controls for some subject characteristics that were previously unobserved, but not necessarily for all important observable subject characteristics. Table 16 reports intent-to-treat effects from the subset of follow-up answerers in the treatment and control groups, after including subject characteristics in the regression model. As in previous specifications, we use whether or not a subject voted in the 2010 or 2012 general election or both elections, subject age, gender, educational attainment, estimated income, state, and number of subjects in household as control variables. Several estimates which were not significant in Table 15 are statistically significant with the introduction of control variables, including estimates for all treatment subjects and subjects T1 among follow-up answerers, and subjects in T3 among subjects with an answering machine for the follow-up call. Overall, although the inclusion of controls changes some coefficients, due to slight imbalances in subject characteristics across treatment groups, the primary results are robust to the inclusion of these controls. As before, the treatment effect is largest among follow-up answerers, specifically follow-up answerers in the T3 treatment.

Table : Intent-to-treat effect among subjects who answered the follow up call with subject controls

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
|  | (1) | (2) | (3) | (4) | (5) | (6) |
| Variables | voted | voted | voted | voted | voted | voted |
|  |  |  |  |  |  |  |
| Population | Follow-up answerer | Follow-up answerer | Follow-up AM | Follow-up AM | Follow-up non-answerer | Follow-up non-answerer |
|  |  |  |  |  |  |  |
| T1 |  | 0.00649\* |  | -0.00459 |  | 0.00311 |
|  |  | (0.00359) |  | (0.00297) |  | (0.00350) |
| T3 |  | 0.00893\*\* |  | 0.00657\*\* |  | 0.00704\*\* |
|  |  | (0.00362) |  | (0.00300) |  | (0.00346) |
| T6 |  | 0.00477 |  | 0.000500 |  | 0.000432 |
|  |  | (0.00361) |  | (0.00303) |  | (0.00348) |
| All treatments | 0.00671\*\* |  | 0.000756 |  | 0.00354 |  |
|  | (0.00295) |  | (0.00245) |  | (0.00284) |  |
|  |  |  |  |  |  |  |
| Subject controls | Yes | Yes | Yes | Yes | Yes | Yes |
| Observations | 150,964 | 150,964 | 221,746 | 221,746 | 166,857 | 166,857 |
| R-squared | 0.300 | 0.300 | 0.292 | 0.292 | 0.266 | 0.267 |
| F-test | 2656 | 2495 | 3782 | 3554 | 2312 | 2172 |
| Prob > F | < 0.001 | < 0.001 | < 0.001 | < 0.001 | < 0.001 | < 0.001 |

Note: The table shows OLS estimates for the difference in voting rates between subjects in the treatment group and subjects in the control group. The subset of subjects used to estimate each specification is described in the population row. Subject controls include whether or not a subject voted in the 2010 or 2012 general elections or both and subject age, gender, education, income, number of subjects in the household, and state. Standard errors are clustered by household and displayed in parentheses. \*\*\* p<0.01, \*\* p<0.05, \* p<0.1

By using the seemingly unrelated data generated by the follow-up call, this analysis helps mitigate the selection biases inherent to an experiment in which subjects opt in to treatment. The estimates that are generated are closer to the ostensible causal treatment effect on the treated subjects. This is a potential supplement to traditional methods of estimating treatment effects, such as the local average treatment effect discussed in Appendix E.

# Appendix E: Local Average Treatment Effects

In this appendix we compute the Local average treatment effect (LATE) (Angrist and Pischke 2014) to obtain an estimate of the treatment effect of our robo calls on subjects who were successfully treated (as opposed to the intent-to-treat effect reported in the main body of the paper). The LATE is a two-stage least squares (2SLS) IV estimate of the average causal effect for subjects who were treatment live answerers. That is, calculating the LATE allows us to estimate the magnitude of the treatment effect on the subjects who received the treatment.[[13]](#footnote-13) We use the random assignment to the treatment group as an instrument for answering a treatment call live or by answering machine. A valid instrument satisfies three criteria. One is that the instrument must have a causal effect on the causal variable of interest. Second the instrument must be randomly assigned and be unrelated to important omitted variables. Third, the instrument must only affect the dependent variable through the causal variable of interest. Our application meets these three criteria. Therefore, the LATE calculation is an estimate of the causal treatment effect on subjects who were treatment live answerers or received a message on their answering machine.

In Table 17 we report the results of the experiment without subject controls that form the basis of the LATE calculations in Table 18. We find that estimated LATE for all subjects is 0.42 percentage points, and that this estimate is statistically significant at the five percent level (Table 18 Column 1). This magnitude implies an expected four additional voters for every 1,000 subjects reached. Table 18, Columns 2 through 4 report the LATE estimates for each treatment group, and Columns 5 through 8 report the corresponding LATE estimates for follow-up answerers.[[14]](#footnote-14) Column 3 shows that the LATE for T3 is 0.86 percentage points and Column 7 shows that the LATE for T3 among only follow-up answerers is 1.1 percentage points. Both of these estimates are statistically significant at the one percent level. The latter estimate implies an expected eleven additional voters for every 1,000 subjects reached in T3. The estimates for T1 and T6 are not statistically significant. These estimates support our previous findings that the T3 treatment generates the largest treatment effect and automated calls are an effective tool to increase voter participation.

Table : Intent-to-treat effect for all subjects and single-voter households (SVH), no subject controls

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
|  | (1) | (2) | (3) | (4) |
| Variables | voted | voted | voted | voted |
|  |  |  |  |  |
| Population | All subjects | All subjects | SVH | SVH |
|  |  |  |  |  |
| T1 |  | 0.00219 |  | 0.00755 |
|  |  | (0.00225) |  | (0.00590) |
| T3 |  | 0.00647\*\*\* |  | 0.0136\*\* |
|  |  | (0.00226) |  | (0.00591) |
| T6 |  | 0.000798 |  | 0.00442 |
|  |  | (0.00226) |  | (0.00582) |
| All Treatments | 0.00316\* |  | 0.00845\* |  |
|  | (0.00184) |  | (0.00479) |  |
| Constant | 0.489\*\*\* | 0.489\*\*\* | 0.380\*\*\* | 0.380\*\*\* |
|  | (0.00159) | (0.00159) | (0.00414) | (0.00414) |
|  |  |  |  |  |
| Subject controls | No | No | No | No |
| Observations | 539,567 | 539,567 | 54,849 | 54,849 |
| R-squared | 7.510e-06 | 2.495e-05 | 5.648e-05 | 1.028e-04 |
| F-test | 2.951 | 3.253 | 3.111 | 1.876 |
| Prob > F | 0.0858 | 0.0207 | 0.0778 | 0.131 |

Note: The table shows OLS estimates for the difference in voting rates between subjects in the treatment group and subjects in the control group. The subset of subjects used to estimate each specification is shown with the “Population” label. Standard errors are clustered by household and displayed in parentheses. \*\*\* p<0.01, \*\* p<0.05, \* p<0.1

Table : Local Average Treatment Effect (LATE) estimates

|  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- |
|  | (1) | (2) | (3) | (4) | (5) | (6) | (7) | (8) |
| Variables | voted | voted | voted | voted | voted | voted | voted | voted |
|  |  |  |  |  |  |  |  |  |
| Population | All subjects | T1, control | T3, control | T6, control | Follow-up answerers | T1, control, follow-up answerers | T3, control, follow-up answerers | T6, control, follow-up answerers |
|  |  |  |  |  |  |  |  |  |
| Live answer or | 0.00422\*\* | 0.00300 | 0.00861\*\*\* | 0.00104 | 0.00578\* | 0.00574 | 0.0107\*\*\* | 0.00106 |
| AM | (0.00209) | (0.00263) | (0.00256) | (0.00252) | (0.00311) | (0.00390) | (0.00378) | (0.00373) |
| Constant | 0.489\*\*\* | 0.489\*\*\* | 0.489\*\*\* | 0.489\*\*\* | 0.527\*\*\* | 0.527\*\*\* | 0.527\*\*\* | 0.527\*\*\* |
|  | (0.00136) | (0.00136) | (0.00136) | (0.00136) | (0.00257) | (0.00257) | (0.00257) | (0.00257) |
|  |  |  |  |  |  |  |  |  |
| Subject controls | No | No | No | No | No | No | No | No |
| Observations | 539,567 | 271,251 | 270,157 | 268,821 | 150,964 | 75,904 | 74,734 | 75,570 |
| R-squared | 5.167e-04 | 2.517e-04 | 7.543e-04 | 8.010e-05 | 1.539e-04 | 1.445e-04 | 2.492e-04 | 1.365e-05 |

Note: The table shows LATE estimates for the effect of answering at least one treatment call. Being assigned to the treatment group is used as an intstrument for answering a treatment call. The subset of subjects used to estimate each specification is described in the population row. Robust standard errors are displayed in parentheses. \*\*\* p<0.01, \*\* p<0.05, \* p<0.1

# Appendix F: Multiple Hypothesis Testing

This appendix evaluates whether our results are robust given the multiple hypothesis testing to attempt to avoid spurious conclusions. Multiple hypothesis testing (MHT) “refers to any instance in which a family of hypotheses is carried out simultaneously and the analyst must decide which hypotheses to reject” (List et al. 2015, 2). In our analysis, we consider intent-to-treat effects for three distinct treatment groups, T1, T3, and T6. In Appendix D we estimated intent-to-treat effects for these three treatment groups among three distinct subgroups – follow-up answerers, subjects who received the follow-up call on an answering machine, and follow-up non-answerers who did not receive the follow-up call on an answering machine. In each of these cases, the fact that we test for an effect for T1, T3, and T6 and do so in the same regression, increases the probability that we would erroneously reject a true null hypothesis. That is, we would erroneously conclude that a treatment was effective even if the apparent effectiveness were only due to random variation in the data.

Following the methodology described in Romano and Wolf (2010), we performed MHT adjustments to check the robustness of our findings. Table 19 and Table 20 show MHT-adjusted p-values corresponding to the treatment effects reported in Table 5 of the manuscript and Table 15 of the appendices, respectively.

We find that the statistical significance of the point estimate for T3 that we report in the main text of this paper is robust to MHT adjustments. For example, in our empirical follow-up call analysis, we find that the point estimate for T3 among follow-up answerers remains significant at the five percent level, after MHT adjustments. However, the point estimate for T3 among follow-up non-answerers is no longer statistically significant at the ten percent level after MHT adjustments.

Table : Multiple hypothesis testing for estimates from Table 5 in the Manuscript

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| Population/subgroup | Treatment | Difference | Unadjusted p-value | MHT adjusted p-value |
| All subjects | T1 vs. Control | 0.00219 | 0.2443 | 0.3937 |
| All subjects | T3 vs. Control | 0.00647 | 0.0003 | 0.0003 |
| All subjects | T6 vs. Control | 0.00080 | 0.6687 | 0.6687 |

Note: The table shows adjusted p-values after correcting for multiple hypothesis testing. Column 1 shows the subject population in the sample, Column 2 shows the intent-to-treat comparison being made, Column 3 shows the difference in voter participation rates between the treatment group and the control group, Column 4 shows the unadjusted estimate of the statistical significance of the estimated treatment effect, and Column 5 shows the adjusted statistical significance of the estimated treatment effect after correcting for MHT.

Table : Multiple hypothesis testing for estimates from Table 15

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| Population/subgroup | Treatment | Difference | Unadjusted p-value | MHT adjusted p-value |
| Follow-up answerers | T1 vs. Control | 0.00534 | 0.1453 | 0.5277 |
| Follow-up answerers | T3 vs. Control | 0.01031 | 0.0033 | 0.0247 |
| Follow-up answerers | T6 vs. Control | 0.00103 | 0.7867 | 0.7867 |
| Follow-up AM | T1 vs. Control | 0.00318 | 0.2927 | 0.7397 |
| Follow-up AM | T3 vs. Control | 0.00488 | 0.1090 | 0.5133 |
| Follow-up AM | T6 vs. Control | 0.00084 | 0.7823 | 0.9507 |
| Follow-up Non-answer, no-AM | T1 vs. Control | 0.00237 | 0.4903 | 0.8737 |
| Follow-up Non-answer, no-AM | T3 vs. Control | 0.00738 | 0.0287 | 0.1863 |
| Follow-up Non-answer, no-AM | T6 vs. Control | 0.00498 | 0.1357 | 0.5507 |

Note: The table shows adjusted p-values after correcting for multiple hypothesis testing. Column 1 shows the subject population in the sample, Column 2 shows the intent-to-treat comparison being made, Column 3 shows the difference in voter participation rates between the treatment group and the control group, Column 4 shows the unadjusted estimate of the statistical significance of the estimated treatment effect, and Column 5 shows the adjusted statistical significance of the estimated treatment effect after correcting for MHT.

# Appendix G: APSA Reporting Standards Response

In this appendix we provide responses to the Recommended Reporting Standards for Experiments (Laboratory, Field, Survey) provided at <https://www.cambridge.org/core/services/aop-file-manager/file/57922fab308173a70b28ddc9/xps-reportingstandards.pdf>. The standards and recommendations are written in plain text and **our responses are bolded**.

1. Hypotheses
   * Specific objectives or hypotheses.
     + What question(s) was (were) the experiment designed to address?
     + **Are automated GOTV calls (robo calls) effective at increasing voter turnout? How does the number of calls affect effectiveness?**
     + What are the specific hypotheses to be tested?
     + **Do subjects in the treated groups vote at higher rates than subjects in the control group?**
2. Subjects and Context
   * Eligibility and exclusion criteria for participants.
     + Why was this subject pool selected? Who was eligible to participate in the study? What would result in the exclusion of a participant? Were any aspects of recruitment changed (such as the exclusion criteria) after recruitment began?
     + **The particular subject pool was selected because to a certain degree Republican subjects had been understudied in GOTV experiments (this appears to have balanced out some in the intervening years). Also, we were working with a generally Republican political consulting firm.**
     + **The pool of potential subjects was Republican registered voters who had voted in exactly two of the 2010 and 2012 primary and general elections in Georgia, Nebraska, New Mexico, Ohio, Pennsylvania, or Virginia and were registered prior to January 1, 2010. Household members of the selected subjects were also included in our analysis because we could not tell who (if anyone) in the household heard the treatment message.**
   * Procedures used to recruit and select participants.
     + If there is a survey: Identify the survey firm used and describe how they recruit respondents.
     + **Subjects were identified and selected based on state voting files. They were not explicitly recruited – they just received the GOTV calls at specified times before the election(or didn’t, if they were in the control group)**
   * Recruitment dates defining the periods of recruitment and when the experiments were conducted.
     + Also list dates of any repeated measurements as part of a follow-up.
     + **Subjects were called between October 30 and November 4, 2014. The follow-up call detailed in Appendix D took place on August 5, 2015.**
   * Settings and locations where the data were collected.
     + In the field, lab, classroom, or some other specialized setting?
     + **Field**
     + Other relevant specifics of the population: e.g., large public university vs. small private university; geographic location; etc.
     + **Republican registered voters in Georgia, Nebraska, New Mexico, Ohio, Pennsylvania, and Virginia**
   * If there is a survey: Provide response rate and how it was calculated.
     + **N/A**
3. Allocation Method
   * Details of the procedure used to generate the assignment sequence (e.g., randomization procedures).
     + **Subjects were allocated to the treatment group at the household level by the consulting firm that ran the experiment. The authors did not receive the data until after the randomization had occurred.**
   * If random assignment used, then details of procedure (e.g., any restrictions, blocking).
     + Note the unit of randomization (individuals, groups, households, etc). Pay careful attention to report clustered random assignment if subjects were assigned at some level other than the individual subject.
     + **The unit of randomization is households. Where applicable, standard errors were clustered at the household level.**
   * If random assignment used, provide evidence of random assignment.
     + If demographic or other pretreatment variables were collected, a table (in text or appendix) showing baseline means and standard deviations for demographic characteristics and other pre-treatment measures by experimental group.
     + **Subject characteristics are provided in balance tables in Appendix A**
     + If blocking was used, and group assignment proportions were not equal across blocks, provide table for each of the blocks. If there are too many blocks for this to be practical, combine blocks to present weighted averages of covariates using inverse probability weighting.
     + **N/A**
   * Blinding: Were participants, those administering the interventions, and those assessing the outcomes unaware of condition assignments?
   * **The firm implementing the experiment was aware of condition assignment. However, the implementation was prerecorded phone messages sent out to subjects via an automated system. Participants were not aware of the experiment, per se.**
     + If blinding took place, include a statement regarding how it was accomplished and how the success of blinding was evaluated.
     + **N/A**
4. Treatments
   * Description of the interventions in each treatment condition, as well as a description of the control group.
   * **Subjects in the treatments received automated recorded messages (robocalls) according to the schedule and script detailed in Appendix B. Subjects in the control group did not receive any calls from this experiment.** 
     + Descriptions should be sufficient to allow replication: Summary or paraphrasing of experimental instructions in the article text; verbatim instructions and/or other treatment materials provided in an appendix.
     + **The exact call-schedule and script are in appendix B**
   * How and when manipulations or interventions were administered.
     + Method of delivery: Pen-and-paper vs. computer or internet vs. face-to-face communication vs. over the telephone.
     + **Telephone**
     + If computerized, the software should be described and cited. (If possible, programs should be included in appendix so as to be available for purposes of replication.)
     + **N/A**
     + For lab experiments (and other experiments, when relevant):
       - Report the number of repetitions of the experimental task and the group rotation protocol. Report the ordering of treatments for within-subject designs. Any piggybacking of other protocols should be reported. Report any use of experienced subjects or subjects used in more than one session or treatment.
       - **N/A – not a lab experiment**
       - Time span: How long did each experiment last? How many sessions were subjects expected to attend? If there were multiple sessions, how much time passed between them?
       - **N/A – not a lab experiment**
       - Total number of sessions conducted and number of subjects used in each session.
       - **N/A – not a lab experiment**
       - Was deception used?
       - **N/A – not a lab experiment**
       - Treatment fidelity: Evidence on whether the treatment was delivered as intended.
       - **N/A – not a lab experiment**
         * Report any instructional anomalies or inaccuracies.
         * Were subjects given quizzes on the experimental instructions?
         * Were there practice rounds? If so, how many and what were the results?
         * Did subjects complete a post-experiment debriefing, interview, or questionnaire? If so, is there evidence that subjects understood the instructions and treatments?
         * Did the experimental team observe aspects of the intervention?
         * Provide description of manipulation checks, if any.
       - Were incentives given? If so, what were they and how were they administered.
       - **N/A – not a lab experiment**
5. Results
6. Outcome Measures and Covariates
   * Provide precise definition of all primary and secondary measures and covariates.
   * **The primary measure was whether or not a subject voted, and how that varied between treatment groups and the control.**
     + For indices, provide exact description of how they are formed. For survey items provide exact question wording in an appendix. Please provide a copy of the complete survey questionnaire (in an on-line appendix if it is long).
     + **N/A**
   * Clearly state which of the outcomes and subgroup analyses were specified prior to the experiment and which were the result of exploratory analysis.
   * **The intent-to-treat analysis by treatment group was specified prior (though not preregistered). Results related to single-voter households were the result of exploratory analysis**
7. Complete CONSORT Participant Flow Diagram
   * An example of a CONSORT flow diagram is attached. The flow diagram records the initial number of subjects deemed eligible for the experiment and all losses of subjects during the course of the experiment. The flow chart follows the subjects from initial recruitment to the sample used in the main analyses, providing readers clear information on the amount of attrition and exclusions. The chart also reports the portion of each treatment group that received the allocated intervention and if not, why this was not accomplished. Naturally, in the event that there is zero or very trivial non-compliance with group assignment or zero or very trivial attrition, researchers may decide it is more convenient to report the information that would otherwise be shown in the CONSORT diagram in the text and omit the diagram.
   * **We don’t believe that a CONSORT diagram will provide additional useful information here. The selection criteria and experiment uptake (i.e., whether or not any calls were answered) are described in the manuscript and in considerable detail in Appendix C. To the extent that subjects dropped out of the subject pool (as described in section 3 below), we do not have data on the number or identity of these subjects. We were told by the partner who performed the experiment that it was an insignificant number of subjects, as voter rolls do not change as much in the months following an election as they do in the months leading up to one.**
   * Number of subjects initially assessed for eligibility for the study.
   * Exclusions prior to random assignment and reasons for the exclusions.
   * Number of subjects initially assigned to each experimental group.
   * The proportion of each group that received its allocated intervention and the reasons why subjects did not receive the intended intervention.
   * The number of subjects in each group that dropped out or for other reasons do not have outcome data.
   * The number of subjects in each group that are included in the statistical analysis, and the reasons for any exclusions.
8. Statistical Analysis
   * Researchers will conduct statistical analysis and report their results in the manner they deem appropriate. We recommend that this reporting include the following:
     + Report sample means and standard deviations for the outcome variables using intent-to treat (ITT) analysis (means for the entire collection of subjects assigned to a group, whether the treatment is successfully delivered or not).
     + **This is done in the paper and appendices – frequently through the reporting of multivariate regression results**
       - If the experiment uses block randomization with unequal assignment rates, present ITT analysis by block or present overall means using inverse probability weighting.
       - **N/A**
     + Note if level of analysis differs from level of randomization and estimate appropriate standard errors.
     + **Subjects were effectively randomized at the household level but are analyzed at the individual level. Wherever possible, standard errors are clustered at the household level**
     + If there is attrition, discuss reasons for attrition and examine if attrition is related to pretreatment variables.
     + **There was some attrition between the primary experiment and the follow-up call detailed in Appendix D. It is also possible that there was attrition between the experiment conducted prior to the 2014 election and when the voting records were made available by the various state governments. In both cases, we believe that the primary reasons for attrition are subjects being removed from the registered voting rolls (voluntarily or otherwise), subjects moving (possibly out of state) or forming different households, name changes, or other such events in the intervening months. We suppose these variables could be related to age or income but do not have the data necessary to check that. Similarly, we have no reason to suspect any statistically meaningful impact from the attrition.**
     + Report for other missing data (not outcome variables):
       - Frequency or percentages of missing data by group.
       - **23,914 subjects (4.43% of the total) were missing data for sex, income, age, and/or education**
       - Methods for addressing missing data (e.g., listwise deletion, imputation methods).
       - **Missing data was imputed with a method appropriate to how it is used in the analysis. The missing data for sex and age was replaced with the mean values from the subject pool. The missing data for education and income were changed to an indicator (99 in this case) that allowed them to be analyzed as their own categorial group in the regression analysis**
       - For each primary and secondary outcome and for each subgroup, provide a summary of the number of cases deleted from each analysis and rationale for dropping the cases.
       - **After we did the data imputation (thank you for the suggestion, anonymous reviewer!) we did not need to drop/delete any subjects for whom we had data**
     + For survey experiments: Describe in detail any weighting procedures that are used.
     + **N/A**
9. Other Information
   * Was the experiment reviewed and approved by an IRB?
   * **Yes, as described in the manuscript**
   * If the experimental protocol was registered, where and how can the filing be accessed?
   * **It was not.**
   * What was the source of funding? What was the role of the funders in the analysis of the experiment?
   * **There were no external sources of funding**
     + Were there any restrictions or arrangements regarding what findings could be published? Any funding sources where conflict of interest might reasonably be an issue?
     + **There were no explicit conditions or restrictions on the use of the data. Our partner requested that we refrain from publishing the results for a year after the experiment, but that did not prove to be a remotely binding constraint.**
   * If a replication data set is available, provide the URL.
   * **The data, code, and any additional materials required to replicate all analyses in this article are available at the Journal of Experimental Political Science Dataverse within the Harvard Dataverse Network, at: https://doi.org/10.7910/DVN/DMJ7EA**

1. Once a landline is assigned to a treatment or control group, every voter in that household is assigned to the same group. The number of subjects per treatment group varies because the number of subjects per household differs. [↑](#footnote-ref-1)
2. For example, Gerber et al. (2011) find that individuals measuring high on an extraversion and emotional stability scale are more likely to vote. It is plausible that those individuals are also more likely to answer a phone call. See also Sagarin et al. (2014). [↑](#footnote-ref-2)
3. In unreported results, we find that the difference in voting rates between answerers and non-answerers is persistent going back to at least the 2000 general election. [↑](#footnote-ref-3)
4. Receiving a message on an answering machine could be considered as a form of treatment compliance. Since a treatment message left on an answering machine could induce a subject to vote, we include answering machines for the computation of IV/LATE estimates in Appendix E. However, for our discussion of selection effects, we consider answering machines and live answer to be distinct outcomes. [↑](#footnote-ref-4)
5. Similarly, Gerber et al., (2010) find that the likelihood of not answering the phone is correlated with the likelihood of not voting. [↑](#footnote-ref-5)
6. As with the initial treatment calls, this follow-up call included up to three call attempts if the first and second call outcomes were operator, no answer, busy, fax machine, or otherwise uncompleted. [↑](#footnote-ref-6)
7. This live answer rate is lower than the live answer rate for T1 subjects when they received our automated call prior to the 2014 election, even though in both cases calls were only placed on one day. One possible explanation for the lower answering rate is the increased likelihood of relocation as well as travel since the follow-up call occurred in the summer. [↑](#footnote-ref-7)
8. One alternative to our approach is the placebo method of placing contemporaneous non-political calls to a subset of control group subjects. The goal of that method is to compare voter participation of subjects with similar propensities to answer phone calls. For a detailed description of the placebo method see Gerber et al. (2010). [↑](#footnote-ref-8)
9. The results in Table 15 measure intent-to-treat effects because 18 percent of follow-up answerers were treatment non-answerers. [↑](#footnote-ref-9)
10. After correcting for multiple hypothesis testing with three subgroups and three treatment groups, we find that this estimate remains statistically significant at the five percent level. See Appendix F. This estimate is also similar to the estimate found through an instrumental variable (IV) estimate of the local average treatment effect (LATE) as discussed in Appendix E. [↑](#footnote-ref-10)
11. As in Table 5 of the manuscript, we estimated treatment effects for subjects in single-voter households who were follow-up answerers. However, this subsample contained fewer than 14,000 subjects. In this restricted sample, the coefficient for all treatments is 0.0100, but is not statistically significant. The coefficient for T3 among follow-up answerers in single-voter households is 0.025 and is statistically significant at the five percent level. [↑](#footnote-ref-11)
12. After correcting for multiple hypothesis testing with three subgroups and three treatment groups, we find that this estimate is not statistically significant at the ten percent level. [↑](#footnote-ref-12)
13. This estimate may also be of more interest to GOTV practitioners, as it is common for campaigns to pay only for calls that are answered by a person or by an answering machine. [↑](#footnote-ref-13)
14. We compute the LATE for the sub-sample of follow-up answerers since this group does not have full compliance, though compliance rates are higher than the overall sample. As noted above, the rate of treatment answering for follow-up answerers compliance is 82 percent. [↑](#footnote-ref-14)