

# Causal Inference in Voter Mobilization: Experimental vs. Selection-On-Observables Based Estimates

Naitik Poddar

May 9, 2025

## **Abstract**

This paper compares the effectiveness of two methods for estimating causal effects: randomized controlled trials (RCTs) and selection on observables. We evaluate these methods by using data from a 2002 voter mobilization field study, investigating whether receiving a phone call encouraging voting increases the likelihood of voting. Our analysis highlights the advantages of experimental designs in producing reliable estimates of causal effects. In contrast, the observational approach produces significantly inflated effect estimates. We conclude that the RCT design offers more reliable causal inference, while the selection-on-observables method remains vulnerable to omitted variable bias and requires cautious interpretation.

# 1 Introduction

Understanding the causal impact of different treatments is an extremely important idea in the real world. One of the most common statistical misunderstandings is that between causal and associational relationships. This difference guides the understanding and effectiveness of crucial treatments in public policy, medicine, and social sciences.

Paul Holland, in his 1986 journal article, emphasizes thinking about causal inference using ‘effects of causes’ rather than ‘causes of effects’. Modern day statistics uses a potential outcomes framework, first introduced as Rubin’s model, to define and estimate causal effects. The idea here is to consider, for each unit in a given population, what would the outcome be under a treatment vs. a control (no treatment) for any given unit. The model also emphasizes that causation cannot occur without manipulation of the variables. However this leads to the ‘fundamental problem of causal inference’ which essentially states that we can only observe either of the outcomes for the same unit[1].

To overcome this problem and estimate causal effects in practice, researchers rely on two dominant methods: randomized controlled trials (RCTs), and observational regression analysis based on selection on observables. Each has important implications for the credibility of real-world policy decisions. Statistically, it’s possible to estimate average causal effects across a certain population under careful assumptions: temporal stability, causal transience, unit homogeneity, and importantly, independence, which is often achieved through randomization in RCTs[1]. In contrast, the selection-on-observables method relies on the assumption that all relevant confounding variables affecting both treatment assignment and outcomes are observed and properly accounted for. By controlling for these observable characteristics through regression or matching techniques, researchers attempt to approximate the conditions of an experiment. However, this method remains susceptible to biases due to omitted or incorrectly measured variables, as it cannot guarantee the complete absence of

unobserved confounders. These assumptions and limitations help bridge the gap between the unobservable individual causal effects and estimable population-level effects. The average causal effect—the difference in outcomes between treatment and control groups across all units—is typically the main target of estimation. The strength of RCTs lies in their ability to eliminate bias from unobserved variables, while observational methods remain vulnerable to omitted variable bias even with extensive controls.

This paper evaluates these two approaches using data from a large-scale voter mobilization experiment conducted during the 2002 midterm elections. Our goal is to assess whether receiving a phone call encouraging voting has a causal effect on turnout, and how results from an observational approach compare to the experimental benchmark.

While the two approaches rely on different assumptions and identification strategies, their results diverge sharply. The experimental analysis finds only a minimal treatment effect, while the observational analysis suggests a substantially larger effect, likely inflated due to selection bias. To better compare the causal impact of receiving the call, we also implement an instrumental variables (IV) approach using assignment as an instrument for treatment receipt. This contrast across methods will show us the importance of research design in making credible causal claims.

## 2 Data

This paper systematically evaluates these two approaches using data derived from the large-scale voter mobilization experiment by Arceneaux, Gerber, and Green in 2006[2]. Increasing voter turnout remains an important concern for democratic societies around the world. Higher participation rates can make a big difference because they lead to more representative electoral outcomes and policy decisions that better reflect the preferences of the entire population, rather than just those of frequent or privileged voters. For example, evidence from

Australia’s adoption of compulsory voting shows that increasing turnout by 24 percentage points resulted in a 7–10 percentage point increase in the vote and seat share for the Labor Party, and also led to greater pension spending at the national level[3]. This demonstrates that higher turnout can shift both electoral results and public policy, ensuring that government decisions are more closely aligned with the broader electorate’s interests. The 2002 study used the strategy of voter mobilization calls (brief phone calls encouraging people to vote) to attempt to boost voter turnout.

Our analysis utilized extensive data collected from this field experiment implemented across Iowa and Michigan during the 2002 midterm elections. The original study aimed to test the performance of matching methods against randomized experiments in estimating causal effects. Their large-scale experiment involved over two million observations, where households were randomly assigned to treatment and control groups. The treatment group received phone calls which attempted to deliver the following get out the vote message.

“Hello, may I speak with (name of person) please? Hi. This is (caller’s name) calling from Vote 2002, a nonpartisan effort working to encourage citizens to vote. We just wanted to remind you that elections are being held this Tuesday. The success of our democracy depends on whether we exercise our right to vote or not, so we hope you’ll come out and vote this Tuesday. Can I count on you to vote next Tuesday?”

The control group did not receive this call. The experiment was structured to provide a benchmark for comparing experimental methods with non-experimental estimators.

The data generation process began with the collection of large voter registration lists from both states. These lists provided detailed demographic and historical voting information for each registered voter, including age, gender, household size, and past election participation - which are important characteristic variables for our analysis. Households

were divided into "competitive" and "uncompetitive" congressional districts, ensuring that random assignment occurred within these political groups to maintain balance across regions with varying political dynamics.

Once stratification was complete, households containing one or two registered voters were randomly assigned to either the treatment or control group. For households with two people, only one individual was randomly selected for treatment to avoid confounding effects. Two national phone banks were contracted to carry out the treatment. The phone calls were made close to the election date, and compliance with the script was monitored to ensure consistency across calls.

Following the election, researchers obtained public voting records to verify whether each individual voted in the 2002 election. This approach ensured high data quality and avoided common issues such as bias and inaccurate self-reports often found in survey data. This resulted was a rich dataset which is ideal for causal inference analysis.

For the purposes of the experimental analysis in this paper, we use a subset of this dataset, specifically focusing on Iowa residents. This sub-dataset comprised 161,490 observations, which is still a substantial sample size for analysis. Within this subset, approximately 85,988 individuals were in the control group, and 15,083 individuals were assigned to the treatment group. In contrast, for our observational analysis, the data set shifts focus from treatment assignment to actual treatment status - specifically, the individuals who answered and listened to the phone calls. This introduces an extremely important distinction: while the experimental data relies on random assignment, the observational data is a result of individuals *choosing to opt into treatment*.

We define several key variables used in our analysis. The outcome variable, *vote02*, is an indicator for whether an individual voted in the 2002 general election. Our primary treatment variables differ by method: for the experimental (RCT) analysis, we use *treat\_real*, which equals 1 if the individual was randomly assigned to receive a mobilization phone call,

regardless of whether they answered it. For the observational analysis, we use *contact*, an indicator for whether the individual actually received and listened to the call. The covariates used to improve estimation precision include age, gender, prior voting indicators (*vote*<sub>00</sub> for voting in the 2000 general election and *vote*<sub>98</sub> for the 1998 midterms), and *newreg*, which flags newly registered voters. These variables are used to construct balance tables, control for observable differences, and improve our regression estimates.

### 3 Methods

Causal inference provides us with a framework to understand the impact of interventions such as voter mobilization calls [1]. At its core is the counterfactual framework: what would have happened to the same individual in the absence of treatment? Since it is impossible to observe both potential outcomes for the same person, we estimate the average treatment effect (ATE) statistically.

$$ATE = E[Y_{1,i} - Y_{0,i}]$$

where  $Y_{1,i}$  is the potential outcome if individual  $i$  receives the treatment, and  $Y_{0,i}$  if they do not.

Our analysis uses three main empirical tools to estimate causal effects:

- **Balance tables** to assess covariate comparability between groups.
- **OLS regressions** to estimate treatment effects under experimental and observational assumptions.
- **Instrumental variables (2SLS)** regression to estimate the local average treatment effect (LATE).

Across all methods, the outcome variable is voting in the 2002 election (*vote02*), and the treatment variable is defined either as assignment to treatment (*treat\_real*) or as actual

contact with the call (`contact`), depending on context. All regressions include the same covariates: `age`, `gender`, `vote98`, `vote00`, and `newreg`.

### 3.1 Balance Tables

Before estimating treatment effects, we assess whether the treatment and control groups are comparable in terms of baseline characteristics. This is important because meaningful causal inference requires that the groups being compared differ only in their treatment status, not in underlying covariates. If the groups differ systematically on pre-treatment variables, any observed outcome differences could be confounded.

To evaluate this, we construct balance tables (seen in Table 1 and 2) that report the means of key baseline variables across two groups, along with the differences in means and their associated p-values from independent t-tests. In the experimental setting, we compare units assigned to treatment versus control. In the observational setting, we compare units who actually received the treatment (e.g., were contacted) versus those who did not. The same set of covariates is used in both settings.

This approach allows us to formally assess whether covariate differences are present, and to justify whether further adjustments (e.g., regression controls) may be necessary in our analysis. Balance tables are an important diagnostic tool in both experimental and quasi-experimental designs.

### 3.2 OLS Regression Estimation

We use ordinary least squares (OLS) regression to estimate the effect of treatment on voting behavior. Under the randomized design, the treatment is `treat_real`, indicating assignment to receive the call. The baseline model is:

$$\text{Vote}_{2002,i} = \beta_0 + \beta_1 \text{Treatment}_i + \varepsilon_i$$

Randomization ensures that  $\hat{\beta}_1$  is an unbiased estimator of the causal effect of assignment. To increase precision, we include a vector of covariates (Table 3):

$$\text{Vote}_{2002,i} = \beta_0 + \beta_1 \text{Treatment}_i + \boldsymbol{\beta} \mathbf{X}_i + \epsilon_i$$

where  $\mathbf{X}_i$  is the vector of covariates. This improves model fit and reduces the variance of the estimator:

$$\text{Var}(\hat{\beta}_1) = \frac{\sigma^2}{\sum (\text{Treatment}_i - \bar{\text{Treatment}})^2}$$

Under the observational approach, we change the treatment variable to **contact**, which indicates whether the person actually answered and listened to the call. This version of the regression estimates (Table 4):

$$\text{Vote}_{2002,i} = \beta_0 + \beta_1 \text{Treatment}_i + \boldsymbol{\beta} \mathbf{X}_i + v_i$$

where “Treatment” now refers to **contact**, i.e., actual exposure to the call. Because this treatment is not randomized,  $\hat{\beta}_1$  is likely biased due to unobserved confounders. We attempt to adjust for observable differences using covariates, but omitted variable bias likely remains.

### 3.3 Instrumental Variables Estimation (2SLS)

In our setting, actual contact with the phone call cannot be randomly assigned — individuals may choose to answer or ignore the call based on factors we do not observe. This makes it difficult to compare outcomes between contacted and non-contacted individuals, as they may differ systematically. To address this problem, we use an instrumental variables (IV) strategy to estimate the causal effect of contact on turnout.

We use treatment assignment (**treat\_real**) as an instrument for actual contact (**contact**). Since assignment is randomized, it provides a valid instrument under standard assumptions:



(1) relevance (assignment affects contact), (2) the exclusion restriction (assignment affects voting only through contact), and (3) monotonicity (no one would avoid contact because they were assigned).

We begin with the **first stage** regression, which estimates the relationship between assignment and actual treatment (Table 5):

$$\text{Treatment}_i = \pi_0 + \pi_1 \text{Instrument}_i + \boldsymbol{\pi} \mathbf{X}_i + u_i$$

We then estimate the **second stage** regression, where we use the predicted treatment values from the first stage to estimate the effect on voting (Table 6):

$$\text{Vote}_{2002,i} = \gamma_0 + \gamma_1 \widehat{\text{Treatment}}_i + \boldsymbol{\gamma} \mathbf{X}_i + \varepsilon_i$$

The coefficient  $\gamma_1$  in this second stage represents our IV estimate — the causal effect of being contacted on voting for the group of **compliers** (those who only received the call because they were assigned to). This is also known as the **Local Average Treatment Effect (LATE)**.

For interpretation, the second stage estimate  $\hat{\gamma}_1$  can also be viewed as the ratio of two quantities: the reduced form and the first stage:

$$\tau_{IV} = \frac{\text{Effect of assignment on outcome (reduced form)}}{\text{Effect of assignment on treatment (first stage)}} = \frac{\delta_1}{\pi_1}$$

This makes clear that the estimate from the 2SLS regression corresponds directly to the instrumental variables estimator.

## 4 Results

### 4.1 Experimental Analysis

We begin by examining the experimental analysis. Table 1 presents the results of the balance check between treatment and control groups. Across key baseline covariates—age, gender, and prior voting history in 1998 and 2000—we can see that there are no *statistically significant* differences. For example, the mean age of individuals in the treatment group was 55.80 years compared to 55.76 years in the control group, this difference can be interpreted to be negligible even without significance testing. Similarly, the proportion of women and the rates of prior voting were almost identical across groups. These findings confirm that randomization successfully balanced both observed and, implies that it also balanced unobserved covariates across the two groups, increasing the validity of the experiment.

From our initial model seen in Table 3, the treatment group showed a voter turnout rate that was 0.6 percentage points higher than the control group. However, this difference, with a p-value of approximately 0.165, is not statistically significant at any conventional level. The observed increase, equivalent to  $\approx 6$  additional voters per 1,000 contacted, is too small and uncertain to conclude practical significance.

Our extended model, with the additional covariates slightly increased the estimated treatment effect to  $\approx 1$  percentage point. This estimate became statistically significant at the 1% level. However, despite statistical significance, the effect size remained small in magnitude, suggesting that being assigned to receive a phone call had minimal practical impact on voting behavior. This stability across specifications for the treatment effect implies that covariates like past voting behavior (e.g.,  $Vote_{00} \beta = 50.985^{***}$ ) are strongly correlated with the outcome but not correlated with treatment status, which is consistent with successful randomization. Therefore the minimal change in the treatment effect suggests covariates do not confound the relationship between the phone call intervention and voting behavior.

To go a step further, we implemented an instrumental variables (IV) analysis to estimate the effect of actually being contacted, using random assignment as an instrument for contact. Table 5 shows the first-stage results, confirming that assignment to treatment is a strong predictor of actual contact, with a coefficient of 0.459 ( $p < 0.01$ ). The corresponding 2SLS results are shown in Table 6. The estimated local average treatment effect (LATE) was approximately 2.1 percentage points—larger than the intent-to-treat effect, but still much smaller than the observational estimates. This estimate captures the causal effect of contact for the subset of individuals whose treatment status was affected by the random assignment (compliers). The IV estimate lies between the ITT estimate and the observational estimate, suggesting that the observational model suffers from upward bias due to unobserved confounders, while the IV estimate provides a more credible causal effect.

## 4.2 Selection-on-Observables Analysis

Table 2 summarizes the baseline characteristics of individuals who were contacted versus those who were not. In contrast to the experimental sample, substantial differences are clearly evident. Contacted individuals were, on average, significantly older (mean age 58.6 years versus 53.1 years for non-contacted individuals), more likely to be female (58.1% vs. 55.2%), and more likely to have voted in prior elections (higher rates of turnout in 1998 and 2000). The balance table shows statistically (at the  $p < 0.01$  level), and by inference-practically, significant differences across all baselines characteristics. These imbalances suggest clear selection into treatment, violating the comparability that randomization provides and indicating a high likelihood of selection bias.

Our findings from adjusted regression model in Table 4 further confirm this bias. The coefficient  $\hat{\beta}_1$  on *Contact* was approximately 10.6 percentage points, and this estimate was highly statistically significant ( $p\text{-value} < 0.01$ ). This suggests that individuals who answered the call were much more likely to vote than those who did not. However, given the dif-

ferences in baseline characteristics, this estimate is likely to have substantial selection bias. The treatment group has pre-existing characteristics strongly associated with higher voting probability regardless of the treatment. Without accounting for these systematic differences, and other variables we don't see here, we would likely overestimate the impact of receiving the phone call.

Incrementally introducing covariates into the regression model (controlling for the observable confounders) reduced the estimated effect of *Contact* from 10.6 percentage points to approximately 5.0 percentage points. This represents a reduction of about 53%, which matches with the expectation that controlling for observable differences between contacted and non-contacted individuals removes part of the bias. However the observational estimate remains considerably larger than both the intent-to-treat (ITT) estimate from the experiment (1.0%) and the instrumental variables (IV) estimate of the treatment effect (2.1%). This persistent gap strongly suggests that important unobserved confounders—such as political interest, civic duty, or social affinity—continue to bias the observational estimates upward. The fact that the observational model estimates an effect more than twice as large as the IV estimate highlights the need for caution when interpreting non-experimental results.

## 5 Conclusion

This study shows the critical role of the methods used in estimating valid causal effects. Using the data from the voter mobilization study, we compared the effectiveness of two approaches for estimating the causal effect of phone calls encouraging voting: randomized controlled trials, and a selection-on-observables based regression approach.

Our results show that the RCT, which effectively uses random assignment to treatment provides the most accurate benchmark. It shows that the *assignment to receive a call* had a very small and practically insignificant effect on turnout (after being controlled for the covari-

ates). This reinforces the conclusion that voter mobilization calls, at least as implemented in this setting, have minimal influence on voter behavior.

In contrast, the observational analysis produced considerably larger estimates of treatment effects. While the outcome variable—voting in the 2002 election—remained consistent, the treatment definitions differed. In the experimental analysis, treatment referred to assignment (intent-to-treat), while in the observational analysis, it referred to actual contact. This mismatch, combined with selection into treatment, introduces bias when treatment receipt is not random. Before controlling for covariates, the estimated effect of receiving the call was over 10 percentage points—far larger than the experimental estimate. Even after adjusting for observed confounders such as age, gender, and prior voting, the estimate remained approximately five times higher.

To address this discrepancy, we used an instrumental variables (IV) approach that leverages random assignment as an instrument for actual contact. The IV estimate, which isolates the effect for compliers—those whose treatment status was influenced by assignment—provided a more credible causal estimate of approximately 2.1 percentage points. This value falls between the small intent-to-treat estimate and the inflated observational estimate, suggesting that IV captures the causal effect of treatment while avoiding the selection bias seen in the non-experimental approach.

Theoretically this difference can be understood through selection bias. As seen in the potential outcomes framework, when individuals self-select into treatment, the average untreated potential outcome among the treated group ( $\frac{1}{n_T} \sum_{i \in \text{treat}} Y_{0i}$ ) probably differs from that of the control group ( $\frac{1}{n_C} \sum_{i \in \text{notreat}} Y_{0i}$ ). This difference arises because, as we saw, individuals who choose to answer calls have similar characteristics and are likely more politically engaged and have matching characteristics that lead to them selecting the treatment. Such unobserved differences violate the assumptions required for unbiased estimation under selection-on-observables and result in an upward bias in the estimated treatment effect, even

after adjusting for observable differences. Formally, our observational estimator can be seen as:

$$\hat{\tau}_{ATE} = \tau_{ATE} + \left( \frac{1}{n_T} \sum_{i \in treat} Y_{0i} - \frac{1}{n_C} \sum_{i \in notreat} Y_{0i} \right)$$

where the second term reflects bias due to differences in untreated outcomes across groups. This aligns with what we found in our regression results, even after controlling for the observed variables, the non-experimental estimate remained much larger than the causal effect estimated through randomization. Thus, omitted variable bias (OVB) resulting from unobserved characteristics like political interest remains a crucial threat to the validity of causal inference in non-experimental settings.

In conclusion, our results underscore that randomized controlled trials offer the most reliable method for estimating causal effects when feasible. Observational methods, while useful, are vulnerable to omitted variable bias, particularly when self-selection into treatment is present, and rely heavily on controlling for observable variables. The addition of the IV approach helps bridge the gap between ITT and observational estimates, providing a more realistic picture of the causal effect for a meaningful subset of the population. Researchers and policymakers should interpret observational estimates with caution and consider experimental or quasi-experimental methods where possible.

## References

- [1] Paul W. Holland, “Statistics and Causal Inference,” *Journal of the American Statistical Association*, vol. 81, no. 396, pp. 945–960, 1986.
- [2] Kevin Arceneaux, Alan S. Gerber, and Donald P. Green, “Comparing Experimental and Matching Methods Using a Large-Scale Voter Mobilization Experiment,” *Political*

*Analysis*, vol. 14, no. 1, pp. 37–62, 2006.

- [3] Anthony George Fowler, “Five Studies on the Causes and Consequences of Voter Turnout,” Doctoral dissertation, Harvard University, 2013. <http://nrs.harvard.edu/urn-3:HUL.InstRepos:11156810>
- [4] Overleaf, “Learn LaTeX in 30 minutes,” [https://www.overleaf.com/learn/latex/Learn\\_LaTeX\\_in\\_30\\_minutes](https://www.overleaf.com/learn/latex/Learn_LaTeX_in_30_minutes)

## Appendix: Tables

Table 1: Balance of Baseline Characteristics: Assigned to Treatment vs. Control (Experimental)

|                                | Control Group | Treatment Group | Difference | p-value |
|--------------------------------|---------------|-----------------|------------|---------|
|                                | Mean          | Mean            |            |         |
| Age (years)                    | 55.827        | 55.628          | 0.199      | 0.2304  |
| Gender (1=Female)              | 0.562         | 0.566           | -0.004     | 0.3582  |
| Voted in 2000 General Election | 0.733         | 0.728           | 0.005      | 0.2107  |
| Voted in 1998 Midterm Election | 0.572         | 0.567           | 0.005      | 0.2235  |
| Newly Registered Voter         | 0.048         | 0.048           | -0.001     | 0.6683  |
| Observations                   | 85,988        | 15,083          |            |         |

Note: Standard errors in parentheses; \*\*\* $p \leq 0.001$ , \*\* $p \leq 0.01$ , \* $p \leq 0.05$

Table 2: Balance of Baseline Characteristics: No Contact vs. Received Call (Observational)

|                                | No Contact | Received Call | Difference | p-value |
|--------------------------------|------------|---------------|------------|---------|
|                                | Mean       | Mean          |            |         |
| Age (years)                    | 53.082     | 58.626        | -5.544***  | 0.0000  |
| Gender (1=Female)              | 0.552      | 0.581         | -0.029***  | 0.0004  |
| Voted in 2000 General Election | 0.693      | 0.769         | -0.077***  | 0.0000  |
| Voted in 1998 Midterm Election | 0.536      | 0.604         | -0.068***  | 0.0000  |
| Newly Registered Voter         | 0.054      | 0.042         | 0.012***   | 0.0005  |
| Observations                   | 8,156      | 6,927         |            |         |

Note: Standard errors in parentheses; \*\*\* $p \leq 0.001$ , \*\* $p \leq 0.01$ , \* $p \leq 0.05$



Table 3: Regression Analysis of Voting Behavior (Experimental: Assigned to Treatment)

| VARIABLES                      | Model 1              | Model 2              | Model 3              | Model 4              | Model 5              | Model 6              |
|--------------------------------|----------------------|----------------------|----------------------|----------------------|----------------------|----------------------|
| Assignment to treatment group  | 0.602<br>(0.433)     | 0.714*<br>(0.423)    | 0.806*<br>(0.424)    | 0.978***<br>(0.379)  | 0.996***<br>(0.365)  | 1.009***<br>(0.364)  |
| Age (years)                    |                      | 0.566***<br>(0.008)  | 0.617***<br>(0.008)  | 0.323***<br>(0.007)  | 0.136***<br>(0.008)  | 0.155***<br>(0.008)  |
| Gender (1=Female)              |                      |                      | -3.338***<br>(0.306) | -2.799***<br>(0.273) | -2.439***<br>(0.264) | -2.415***<br>(0.263) |
| Voted in 2000 General Election |                      |                      |                      | 50.985***<br>(0.323) | 37.553***<br>(0.349) | 40.037***<br>(0.363) |
| Voted in 1998 Midterm Election |                      |                      |                      |                      | 27.132***<br>(0.317) | 27.470***<br>(0.317) |
| Newly Registered Voter         |                      |                      |                      |                      |                      | 16.035***<br>(0.663) |
| Constant                       | 59.340***<br>(0.167) | 27.761***<br>(0.474) | 28.469***<br>(0.494) | 6.138***<br>(0.463)  | 10.862***<br>(0.450) | 6.995***<br>(0.476)  |
| Observations                   | 101,071              | 101,071              | 98,367               | 98,367               | 98,367               | 98,367               |
| R-squared                      | 0.000                | 0.048                | 0.056                | 0.247                | 0.299                | 0.303                |

Standard errors in parentheses, \*\*\* $p < 0.01$ , \*\* $p < 0.05$ , \* $p < 0.1$ .

Note: (vote02) was multiplied by 100; coefficients represent percentage point changes.

Table 4: Regression Analysis of Voting Behavior (Observational: Received Call)

| VARIABLES                      | Model 1              | Model 2              | Model 3              | Model 4              | Model 5              | Model 6              |
|--------------------------------|----------------------|----------------------|----------------------|----------------------|----------------------|----------------------|
| Received and listened to call  | 10.595***<br>(0.796) | 7.913***<br>(0.791)  | 6.523***<br>(0.793)  | 5.041***<br>(0.713)  | 5.101***<br>(0.689)  | 4.998***<br>(0.688)  |
| Age (years)                    |                      | 0.484***<br>(0.021)  | 0.551***<br>(0.021)  | 0.262***<br>(0.020)  | 0.077***<br>(0.020)  | 0.093***<br>(0.020)  |
| Gender (1=Female)              |                      |                      | -3.119***<br>(0.793) | -2.408***<br>(0.713) | -2.008***<br>(0.689) | -1.929***<br>(0.688) |
| Voted in 2000 General Election |                      |                      |                      | 49.459***<br>(0.837) | 36.700***<br>(0.900) | 38.711***<br>(0.936) |
| Voted in 1998 Midterm Election |                      |                      |                      |                      | 26.485***<br>(0.821) | 26.800***<br>(0.820) |
| Newly Registered Voter         |                      |                      |                      |                      |                      | 13.213***<br>(1.726) |
| Constant                       | 55.076***<br>(0.539) | 29.405***<br>(1.235) | 29.783***<br>(1.286) | 9.103***<br>(1.207)  | 13.538***<br>(1.175) | 10.349***<br>(1.244) |
| Observations                   | 15,083               | 15,083               | 14,679               | 14,679               | 14,679               | 14,679               |
| R-squared                      | 0.012                | 0.045                | 0.053                | 0.235                | 0.285                | 0.288                |

Standard errors in parentheses, \*\*\* $p < 0.01$ , \*\* $p < 0.05$ , \* $p < 0.1$ .

Note: (vote02) was multiplied by 100; coefficients represent percentage point changes.

Table 5: First Stage Regression: Effect of Assignment on Treatment Receipt

| VARIABLES  | First Stage: Instrument = Assignment |
|--|--------------------------------------|
| Assignment to treatment group  | 0.459***<br>(0.002)                  |
| Constant   | -0.000<br>(0.001)                    |
| Observations   | 101,071                              |
| R-squared  | 0.419                                |
| Standard errors in parentheses. *** $p < 0.01$ , ** $p < 0.05$ , * $p < 0.1$ |                                      |
| Dependent variable: contact (received and listened to call)                  |                                      |

Table 6: Instrumental Variable (2SLS) Regression Analysis of Voting Behavior

| VARIABLES                      | Model 1              | Model 2              | Model 3              | Model 4              | Model 5              |
|--------------------------------|----------------------|----------------------|----------------------|----------------------|----------------------|
| Received and listened to call  | 1.311<br>(0.944)     | 1.555*<br>(0.921)    | 1.735*<br>(0.913)    | 2.106***<br>(0.815)  | 2.144***<br>(0.786)  |
| Age (years)                    |                      | 0.565***<br>(0.008)  | 0.616***<br>(0.008)  | 0.322***<br>(0.007)  | 0.135***<br>(0.008)  |
| Gender (1=Female)              |                      |                      | -3.342***<br>(0.306) | -2.803***<br>(0.273) | -2.444***<br>(0.264) |
| Voted in 2000 General Election |                      |                      |                      | 50.971***<br>(0.323) | 37.539***<br>(0.349) |
| Voted in 1998 Midterm Election |                      |                      |                      |                      | 27.133***<br>(0.317) |
| Constant                       | 59.340***<br>(0.167) | 27.811***<br>(0.470) | 28.529***<br>(0.490) | 6.217***<br>(0.460)  | 10.943***<br>(0.447) |
| Observations                   | 101,071              | 101,071              | 98,367               | 98,367               | 98,367               |
| R-squared                      | 0.000                | 0.048                | 0.056                | 0.247                | 0.299                |

Standard errors in parentheses, \*\*\* $p < 0.01$ , \*\* $p < 0.05$ , \* $p < 0.1$ .

Note: (vote02) was multiplied by 100; coefficients represent percentage point changes.

Instrument: treat\_real (assignment to treatment group)