Increasing Inequality: The Effect of GOTV Mobilization on the Composition of the Electorate

Ryan D. Enos Harvard University
Anthony Fowler University of Chicago
Lynn Vavreck University of California, Los Angeles

Numerous get-out-the-vote (GOTV) interventions are successful in raising voter turnout. However, these increases may not be evenly distributed across the electorate and could potentially increase the differences between voters and nonvoters. By analyzing individual level-data, we reassess previous GOTV experiments to determine which interventions mobilize under-represented citizens versus those who regularly turn out. We develop a generalized and exportable test which indicates whether a particular intervention reduces or exacerbates disparities in political participation and apply it to 24 previous experimental interventions. On average, current mobilization strategies significantly widen disparities in participation by mobilizing high-propensity individuals more than the under-represented, low-propensity citizens. The results hold troubling implications for the study and improvement of political inequality, but the methodological procedures laid out in this study may assist the development and testing of future strategies which reverse this pattern.

Scholars are increasingly interested in inequalities in political participation and their consequences for political outcomes (APSA Task Force 2004; Bartels 2008, 2009; Dahl 2006; Gilens 2005, 2012; Schlozman, Verba, and Brady 2012).[1] At the same time, political scientists increasingly use large-scale field experiments to test methods for increasing political participation. The findings of these experiments have been adopted by political campaigns and are an increasingly important feature of electoral politics (Gerber et al. 2011; Green and Gerber 2008; Issenberg 2010, 2012). Often (but not always), a goal of these get-out-the-vote (GOTV) strategies is to reduce the participation gap—the extent to which the electorate differs from the voting-eligible population. While many experiments successfully increase average levels of voter participation, they may not affect all citizens equally. Instead, an experiment may affect the types of citizens already represented in the political process more than underrepresented citizens. In this article, we explicitly test whether GOTV treatments tend to reduce or exacerbate the gap in political participation. We find that GOTV interventions, on average, tend to magnify the participation gap. This finding is widely important for the study and equalization of political representation, political campaigns, and voting behavior.

In a typical GOTV experiment, a researcher will randomly assign individuals, households, or geographic regions to receive a particular treatment. Then, the researcher will typically collect voting data from public records and estimate the average effect of the treatment on voter turnout by calculating a simple difference-in-means. To our knowledge, there are four primary motivations for conducting GOTV experiments. First, scholars and civic groups may want to increase political participation as an end in and of itself. Second, field experiments can improve the efficiency of political campaigns. Third, these experiments might inform the study of voting behavior and human behavior in general. Fourth, the participation gap may hold critical political and policy outcomes, and GOTV experiments can identify methods for reducing this gap. For all but the first motivation, the traditional GOTV analysis of calculating average

[1]An online appendix for this article is available at http://dx.doi.org/10.1017/S0022381613001308 containing more details on the data and analyses and additional results. Data and supporting materials necessary to reproduce the numerical results and figures in the article are currently available at www.dropbox.com/sh/b6257dcqkr27ald/ZuQfx-KUzc.

The Journal of Politics, Vol. 76, No. 1, January 2014, Pp. 273–288 doi:10.1017/S0022381613001308

ISSN 0022-3816

treatment effects is usually insufficient. In addition to knowing *how many* voters were mobilized by a particular treatment, we would also like to know *which types* of voters were mobilized. We offer a generalized method for assessing one particular type of heterogeneity in experimental treatment effects which can lend insights for all three of the preceding motivations, although we focus mainly on the last.

Because voters are systematically unrepresentative of the rest of the population, some individuals and groups may be underrepresented relative to others (Lijphart 1997; Verba, Schlozman, and Brady 1995). By studying the determinants of voter turnout, scholars may be able to propose policies to reduce differences in political participation. We develop a single statistical procedure that can be applied to any GOTV experiment to explicitly test whether the intervention reduced or exacerbated disparities in voter turnout. The test can be easily implemented for any previous experiment, often without the collection of any additional data. We apply this method to 24 experimental interventions from 11 published papers to assess the overall implications of the get-out-the-vote paradigm for equality and representation. Interpreting the point estimates directly, 16 of the interventions are found to have exacerbated the participation gap while only eight are found to have reduced it. Moreover, eight of the exacerbating effects are statistically significant while only two of the reducing effects cross this threshold. A pooled analysis of all interventions reveals a large and statistically significant exacerbating effect of GOTV interventions, on average, on disparities in political participation.

By analyzing many experiments at once, we build upon the work of Arceneaux and Nickerson (2009) who reanalyze 11 previous door-to-door canvassing experiments to determine which subset of citizens should be targeted to achieve the most efficient allocation of campaign resources. They argue that campaigns should target higher-propensity voters in low-salience elections and lower-propensity voters in high-salience elections. In this article, we expand the number of interventions to 24, examine many different types of mobilization strategies, and move beyond the focus of campaigns and electioneering to assess the effects of voter mobilization on inequality and by extension political representation. Some of our subsequent analyses are consistent with the results of Arceneaux and Nickerson. For example, the ease with which low-propensity individuals can be mobilized increases with the salience of the election (shown in the online appendix). However, in the typical election, high-propensity citizens are much easier to mobilize, and this remains true even in many high-salience elections. As a result, modern mobilization strategies, which are being actively used by interest groups and campaigns, can dramatically and systematically change the composition of the electorate.[2]

Get-Out-the-Vote and the Political Consequences of the Participation Gap

"The existence of political equality is the fundamental premise of democracy" (Dahl 2006). Dahl's assertion presents a difficult challenge for democratic societies because political equality is undermined by disparities in voter participation. As Verba, Schlozman, and Brady put it: "since democracy implies not only government responsiveness to citizen interests but also equal consideration of the interests of each citizen, democratic participation must also be equal. . . . No democratic nation—certainly not the United States—lives up to the ideal of participatory equality" (1995, 1). As Citrin, Schickler, and Sides explain, "there are meaningful differences between voters and non-voters," and "tiny changes in the distribution of the popular vote (and thus small differences in turnout) can make an enormous difference to the nation's politics" (2003, 88). Scholars have shown that increased turnout is associated with more equitable municipal spending (Hajnal and Trounstine 2005), altered election results (Hansford and Gomez 2010), less interest-group capture (Anzia 2011), increased working-class representation (Fowler 2013), and general legislative responsiveness (Griffin and Newman 2005).

Recognizing the importance of participation to a healthy democracy and electoral outcomes, political scientists have long studied the determinants of political participation (e.g., Putnam 2000; Verba, Schlozman, and Brady 1995; Wolfinger and Rosenstone 1980). The introduction of field experiments gives political scientists another tool for understanding how campaigns and policy makers can decrease inequalities in turnout, and scholars note the potential for field experiments to

[2]Our empirical strategy only allows us to assess the effect of randomized GOTV interventions, which have been primarily conducted by academic researchers often in conjunction with interest groups or campaigns. We cannot explicitly test for the effects of other GOTV interventions which were not randomized. Nonetheless, to the extent that these GOTV experiments are similar to other mobilization tactics, our results inform us about the effects of the broader voter mobilization enterprise outside of academic research.

increase participation among underrepresented groups. For example, Avey (1989) cites door-to-door canvassing as a method to increase participation among regular nonvoters; Michelson, Garcia Bedolla, and Green point to a "multi-year effort to increase voting rates among infrequent voters" (2008, 198) that relied entirely on GOTV experiments; and Green and Michelson (2009) argue that GOTV campaigns can alter the "age, socio-economic, and racial/ethnic disparities between voters and nonvoters" (2009, 245; see also Garcia Bedolla and Michelson 2012).

What has not been acknowledged by practitioners of voter mobilization is that GOTV methods may actually exacerbate the differences between voters and the voting-eligible population. Furthermore, no previous study has explicitly assessed the effects of voter mobilization on the participation gap and representational inequality. Given the prevalence of GOTV field experiments and the fact that political campaigns are now adopting the methods that these experiments have shown to be effective (Gerber et al. 2011; Green and Gerber 2008; Issenberg 2010, 2012), an assessment is warranted.

How Can GOTV Treatments Exacerbate the Participation Gap?

For both theoretical and empirical reasons, our claim that get-out-the-vote treatments may exacerbate the participation gap may be counterintuitive. First, many scholars assume that increased turnout is unambiguously good for democracy and equality (Key 1949; Lijphart 1997; Schattschneider 1960). For example, Schattschneider writes "Unquestionably, the addition of forty million voters (or any major fraction of them) would make a tremendous difference [in the quality of representation]" (1960, 101). Schattschneider's observation is obviously true in the extreme because with universal turnout there would be no difference between the population of voters and the voting-eligible population. However, when not everyone votes regularly, especially in systems with widespread nonparticipation like the United States, a marginal increase in voter turnout may actually exacerbate differences if the increase is concentrated among high-propensity voters, the types of citizens who are voting at high rates anyway.

To illustrate this possibility, Figure 1 shows three hypothetical experimental treatments. For all three treatments and a control group, the lines plot voter turnout across varying underlying propensities to vote. The underlying propensity to vote can be thought of as the probability that the voter would have voted absent a treatment. All three treatments have the same average effect. For an individual with an average propensity, the treatments all increase the probability of voting by 10 percentage points. Despite this large, positive average effect, the treatments have starkly different implications for equality of participation. One treatment effect (blue/dotted; color online) is concentrated among low-propensity voters and therefore reduces the participation gap. Another treatment (red/dashed) is concentrated among high-propensity voters, thereby exacerbating differences. The other treatment effect (green/solid) is homogeneous and is therefore neutral in regards to these disparities.

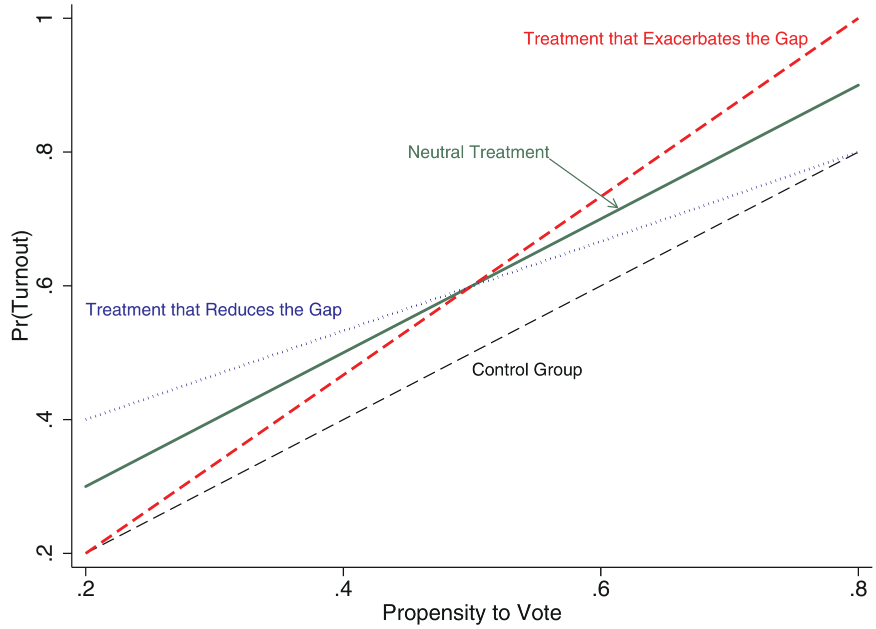
Our findings may also be counterintuitive because at high-propensity levels we expect to see a ceiling effect. If an individual is sure to vote in the absence of a treatment, then no treatment can exhibit a positive effect. For this reason, we might expect most treatments to naturally favor equality. After all, low-propensity voters have more room to increase their turnout probabilities because they are voting at lower rates. These factors make our subsequent results all the more surprising. Despite the possibility of a ceiling effect and despite the large numbers of low-propensity citizens that can be mobilized, we still find that most GOTV interventions *exacerbate* the already stark disparities in voter turnout.

Previous Studies of "Who Is Mobilized?"

The varying treatment effects across different underlying propensities to vote that are illustrated in Figure 1 can be thought of as interactions between the propensity to vote and the treatment. Some previous GOTV studies have examined interactive effects in attempts to test mechanisms or to look for stronger or weaker effects across subgroups. For example, Nickerson and Rogers (2010) show that discussing a "voting plan" tends to mobilize individuals who live alone but has little effect for others; Alvarez, Hopkins, and Sinclair (2010) find that partisan campaign contacts are most effective in mobilizing new registrants; Gerber and Green (2000b) find that nonpartisan leaflets are most effective among nonpartisans who have recently voted; and Green and Gerber (2008) find that canvassing typically boosts turnout among those who voted in the previous election.

Despite these examples, this type of analysis is rare. Experimental researchers do not typically explore all possible interactive effects of their treatment—and rightfully so, because atheoretical testing of multiple

FIGURE 1 **Three Hypothetical Experiments with the Same Average Effect**



Note: The figure presents three hypothetical experimental treatments. Each one significantly boosts the average level of turnout, but they have starkly different implications for the participation gap. The blue/dotted line reduces the participation gap by primarily mobilizing low-propensity citizens. The green/solid line mobilizes citizens equally at all propensity levels and is therefore neutral in regard to the participation gap. The red/dashed line actually exacerbates the participation gap by primarily mobilizing high-propensity citizens, thereby making the electorate less representative of the voting-eligible population.

interactive effects will generate false-positive findings (Gabler et al. 2009; Pocock et al. 2002). While researchers could reduce this problem by testing for interactions with split samples (Green and Kern 2012), the atheoretical testing of many hypotheses is not recommended. This methodological issue poses a problem for those interested in the participation gap and the related effects on representation and policy outcomes if this interest is not factored into the design phase of an experiment. As previously discussed, the typical GOTV study simply tells us how many people were mobilized by a particular treatment but not which types of individuals were mobilized. To learn who is mobilized, an interactive test is called for, but this increases the danger of producing false positives.

An Empirical Test for the Effect of a Treatment on the Participation Gap

We develop a single test which explicitly assesses whether an experimental treatment exacerbates or reduces disparities in participation. We propose a procedure which reduces an individual's pretreatment characteristics onto a single dimension, her propensity to vote. Having estimated this propensity, we test whether the experimental treatment effect varies across this single variable.

Many demographic factors predict an individual's propensity to vote. In our analysis, we reduce all of the factors onto one scale—a propensity score—that indicates the a priori probability that a person with those characteristics will vote in the absence of any GOTV intervention. Of course, not all relevant demographic variables are available through public records used in GOTV experiments. Voter files typically indicate an individual's age, gender, geographic location, and previous turnout history. Variables like race and party registration are only available in certain states, and personal information such as income, education, or church attendance is never available. Even the limited variables that are available serve as a proxy for the extent to which a particular individual may be represented in the political process. For the purposes of this article, we do not care *why* these demographic variables predict voter turnout. The fact that any variables can systematically predict the propensity to vote suggests that there are meaningful disparities in political participation, and our test assesses whether a particular treatment reduces or exacerbates those disparities. Later in the article, we employ survey data to show that this propensity variable is strongly correlated with demographic factors that portend specific policy positions such as income, education, church attendance, and marital status as well as political attitudes on taxes, minimum wage, federal spending, affirmative action, and other important issues.

If a specific experimental treatment mobilizes low-propensity citizens more than high-propensity citizens, then we will say that the treatment has reduced the participation gap. In other words, the demographic gap between voters and the greater population has been reduced. However, if a treatment mobilizes high-propensity citizens more than low-propensity citizens, we will say that the intervention has exacerbated the participation gap. In this scenario, the voting population has become even more unrepresentative of the general population.

Our estimation procedure involves three specific steps. First, we estimate a propensity score for every individual in the sample by regressing voter turnout on every available demographic variable for each individual in the control group.[3] The specification for estimation should be as flexible as possible given the amount of available data. We restrict this part of the analysis to subjects assigned to the control group because we want to estimate the propensity of each individual in the absence of any treatment. Because individuals have been randomly assigned, we know that the propensity scores of the control group are representative of the treatment group as well. Then, for every individual in the sample, we calculate their predicted probability of voting. For each individual, this score indicates their a priori probability of voting in the absence of a treatment. This score represents our single propensity variable that we employ to assess the effect of each treatment on differences in participation. For the second step, we rescale the propensity variable such that the mean equals 0 and the standard deviation equals 1. This is done by subtracting the mean from each individual score and then dividing by the standard deviation. This step allows us to reasonably compare treatments across different types of elections and populations and also improves the interpretation of our subsequent estimates. Lastly, we estimate the following interactive model by ordinary least squares (OLS) to test whether the treatment effect increases or decreases as the propensity score increases:

$$\begin{aligned}\text{Turnout}_i = \alpha + \beta^{*}\text{Treatment}_i + \gamma^{*}\text{Propensity}_i \\ + \delta^{*}\text{Treatment}_i{}^{*}\text{Propensity}_i + \varepsilon_i.\end{aligned} \tag{1}$$

The two coefficients of interest are β and δ. The *treatment* coefficient, β, represents the treatment effect for an individual with an average propensity score (this is not necessarily the same as the average treatment effect). The interactive or multiplicative coefficient, δ, represents the extent to which the treatment effect varies with the propensity score. For example, a δ of .01 indicates that the treatment effect increases by 1 percentage point, on average, for every standard deviation increase in propensity. If δ is greater than 0, then the treatment has exacerbated the participation gap, while a δ less than 0 indicates that the treatment has reduced the gap. If the interactive effect is nonlinear, then we should not interpret the model literally by imputing the treatment effect at various levels of propensity. Nonetheless, even relaxing any assumption of linearity, the sign of δ still indicates whether the treatment is on average greater or smaller for high-propensity citizens.

Two tricky methodological issues could theoretically pose a problem for our analysis but do not. First, the propensity variable is estimated from one regression and incorporated into a second regression. As a result, the second regression may yield incorrect standard errors by failing to incorporate the uncertainty from the first regression. We can assess this additional uncertainty with a nonparametric bootstrap, and in each case in our data, the bootstrapped standard errors are virtually identical to those where we ignore the initial uncertainty of the propensity variable. Second, we might worry that our initial regression "overfits" the data by fitting random noise rather than an underlying relationship. If this were the case, our subsequent estimates could be biased. Specifically, if we overfit the data in the first regression, that would lead to a downward bias in the interaction term in the second regression because our propensity variable would not predict turnout out-of-sample as well as it does within-sample. We can address this issue by randomly partitioning the control group into two groups, estimating the propensity variable with one group, and then running the second regression using only the second

[3]We use OLS, a linear probability model, for this step but the results are virtually unchanged if we use Logit or Probit instead. Advantages of the linear probability model in the setting as compared to other models include computational speed, fewer assumptions and the avoidance of the "incidental parameters problem" (Angrist and Pischke 2009).

group. This procedure also produces results nearly identical to the original setup.[4]

Having estimated this parametric model, we also explore the nonlinearity of the interactive effect through kernel regressions. By plotting the probability of turnout as a nonparametric function of the propensity score for both the treated and control observations, we obtain a more detailed picture of how the treatment effect varies across different propensity scores.

This analysis provides a single, theoretically based method for evaluating the effect of a treatment on participation differences. In the following sections, we apply this procedure to 24 different GOTV interventions and then pool all of the data to assess the overall effect of these programs on the participation gap.

A Detailed Example: The Neighbors Treatment

To clarify our procedure, we we will discuss in greater detail one experiment by Gerber, Green, and Larimer (2008) that we will call the "neighbors" treatment. Following this discussion, we will provide a broader analysis of the others. In this experiment, the authors randomly assigned registered voters to receive postcards in the run up to a low-salience 2006 primary election in Michigan. The sample included registered voters in Michigan who voted in the 2004 general election. Subjects in the control condition received no contact at all, while subjects in the treatment condition received a memorable postcard, indicating the previous turnout behavior of individuals in their neighborhood. All at once, this powerful treatment informed citizens about the upcoming election, indicated that a researcher was watching them, and threatened to share their turnout behavior with others in the neighborhood.

The neighbors treatment had a massive average-treatment effect, raising turnout by 8 percentage points. However, despite a strong average effect, the treatment may not be equally effective for all types of citizens. For the purposes of our study, we want to know whether this treatment tends to reduce or exacerbate the participation gap by mobilizing low- or high-propensity citizens.

To estimate our propensity variable for each subject in the experiment, we utilize all the pretreatment variables available in the data provided by the authors. We know whether each subject voted in the primary and general elections of 2000, 2002, and 2004, the gender of each subject, their age, and their household size (calculated as the number of voters registered at the same address). Restricting our analysis to the control group, we divided subjects' ages into 13 categories and then regressed the dependent variable (turnout in a 2006 primary) on turnout in each of the six previous elections, gender, age-category fixed effects, and household-size fixed effects. We could have chosen a more flexible model with interactive effects, but we must be careful to avoid overfitting. Having estimated the model by OLS, we predict for each individual in both the control and treatment groups their a priori probability of voting. These predicted probabilities have a mean of .30 and standard deviation of .13, which we then rescale so that the mean is 0 and the standard deviation is 1. The r-squared of .21 indicates that our model explains a significant proportion but not all of the variation in turnout.[5]

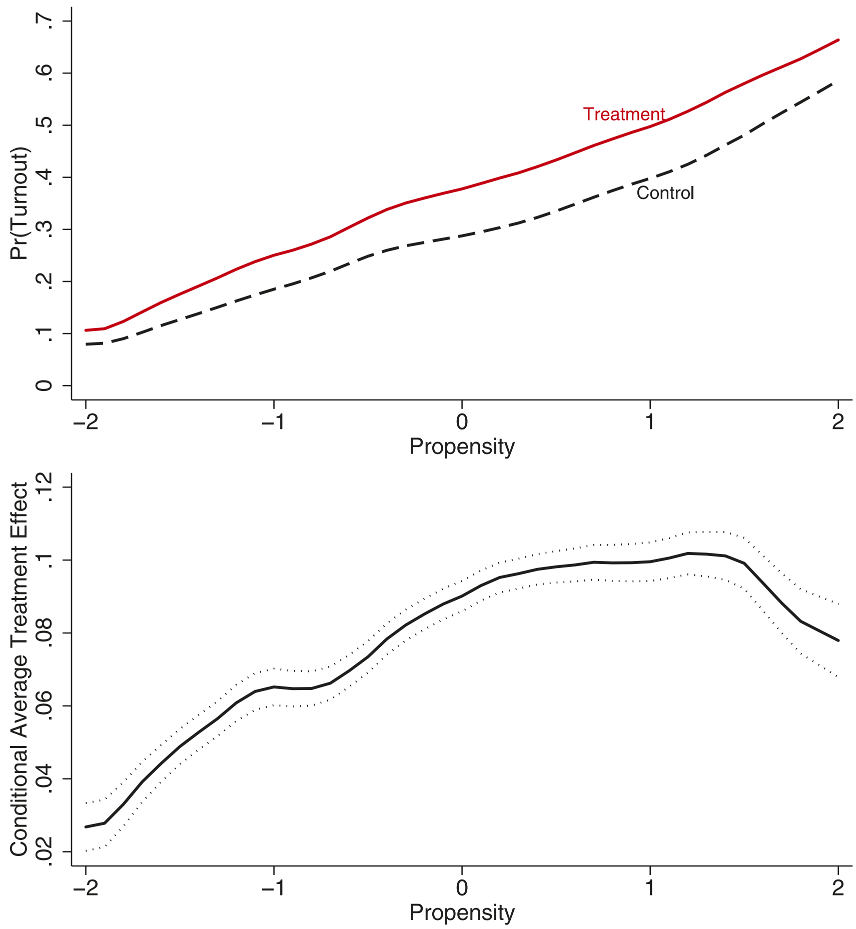
Having estimated the propensity variable, we test whether the treatment effect varies across different propensities. As described earlier, we regress turnout on the treatment, propensity, and the interaction of the two. The coefficient on the treatment variable is .080 with a standard error of .003 (clustered by household), indicating that the treatment effect was 8.0 percentage points for the hypothetical individual with an average propensity score.[6] This is remarkably close to the average treatment effect reported by Gerber, Green, and Larimer (2008), but this need not be the case. The coefficient on the interaction term is of greatest interest because it indicates how the treatment effect changes as propensity changes. Here, we estimate a coefficient of .016 with a standard error of .003 ($p < .01$), indicating that on average the treatment effect increases by 1.6 percentage points for every standard deviation increase in

[4]Future scholars who implement this test should be sensitive to these issues, particularly if the sample size is small.

[5]Given the natural variation in voter turnout within demographic subgroups, this r-squared statistic fits well within the bounds of r-squared statistics found in other published models of voter turnout. To obtain a better sense of the extent to which our propensity model captures meaningful differences within the electorate, we can examine the range and distribution of the propensity variable before rescaling. The top panel of Figure 2 shows that our propensity variable identifies some subgroups with probabilities of voting as low as .1 or as high as .6. Results for the other experiments, shown in the online appendix, are similar.

[6]Standard errors are clustered by household or block where applicable and are heteroskedasticity robust in all cases. See the online appendix for more details.

Figure 2 Variation in Neighbors Treatment Effect



Note: The figure shows variation in the "neighbors" treatment across different propensity levels. The top panel presents kernel regressions of voter turnout on propensity separately for the treatment and control group. The bottom panel presents the difference between these two kernel regressions, presenting the conditional average treatment effect at each propensity level. The dotted lines indicate standard errors. We see that the "neighbors" treatment is much larger for high-propensity citizens relative to low-propensity citizens, thereby exacerbating the participation gap.

propensity. According to our statistical test, the neighbors treatment exacerbated the participation gap despite significantly raising the average level of turnout.

One way to assess the substantive size of this interactive effect is to compare the treatment effects for two hypothetical groups whose propensity scores are, respectively, two standard deviations above and below the mean. According to our model, the conditional average treatment effect of the neighbors intervention was 11.2 percentage points (.080 + 2*.016 = .112) for the high-propensity group, while the effect was only 4.8 percentage points for the low-propensity group. As such, the neighbors treatment effect is more than two times greater for the highest propensity individuals in the sample compared to the lowest propensity individuals. This difference is substantively large and statistically significant, so we conclude that this experimental intervention significantly exacerbates the participation gap.

Having conducted our formal statistical test, we take a closer look at this interactive effect with a graphical, nonparametric approach. In Figure 2, we employ kernel regressions to plot the probability of turnout across individuals' a priori probabilities of voting. We plot these relationships separately for both the control condition and the treatment condition. We display these curves in the top panel of Figure 2. For low-propensity voters, there is little difference between the curves, indicating a small conditional average treatment effect. However, as propensities increase, the gap between the groups increases as well, indicating that the treatment effect is becoming stronger for high-propensity citizens. In the bottom panel of Figure 2, we plot the difference between these two kernel regressions, showing how the treatment effect varies across different propensities. In this panel, the higher the curve, the greater the treatment effect, so a positively sloping curve indicates that the treatment effect increases as propensity increases. The curve in the bottom panel turns downward at the highest propensities, demonstrating the expected ceiling effect. However, prior to this downward turn, the curve is positive and increasing, indicating that the treatment has an increasingly strong effect as the underlying propensity to vote increases—thereby exacerbating existing differences in voter turnout.

Later, we apply these same procedures to 24 different experimental treatments. This allows us to statistically assess the effects of GOTV treatments on the participation gap and also to visualize the way in which these effects vary across different experiments. The Neighbors treatment is not an anomaly: numerous experimental interventions exacerbate the participation gap while few interventions significantly reduce it. First, however, we establish that the propensity-score measure we have described is associated with politically meaningful variation in the electorate.

Does the Propensity Score Capture Meaningful Differences Between Voters?

We may not care whether GOTV interventions mobilize high- or low-propensity citizens if our propensity variable does not capture meaningful characteristics of the electorate that are associated with policy preferences. Political scientists have extensively studied the correlates of turnout (e.g., Putnam 2000; Verba, Schlozman, and Brady 1995; Wolfinger and Rosenstone 1980), and this literature suggests that high-propensity citizens will be systematically different from low-propensity citizens across a number of politically relevant variables. Even though we only have data on vote history and a few demographics, we argue that our propensity variable is a proxy for many characteristics that we care about such as socioeconomic status (SES) and issue positions.

We test the political relevance of our propensity variable using survey data. We employ data from the 2008 versions of the Cooperative Congressional Election Study (CCES) and the Cooperative Campaign Analysis Project (CCAP). For each survey respondent, we generate our propensity variable, using turnout in the 2008 general election as the dependent variable and only using information that would be available in public records as the independent variables: age, race, gender, household size, state, party registration, and vote history in previous elections. To keep in line with typical GOTV samples, we only include registered voters in the samples. Also, because survey respondents were matched to statewide voter files, we use validated turnout data instead of reported behavior in the analysis.

Having generated our propensity variable using only those variables available in public records, we test whether this variable captures other meaningful features of the citizenry. Table 1 reports the results of 33 separate regressions. In each case, we regress a demographic characteristic or political attitude on the propensity variable. The coefficient indicates the extent to which the characteristic changes, on average, for every standard deviation increase in propensity. With the exceptions of family income, ideology, and party identification, all dependent variables are coded as dummies, so the coefficients can be interpreted as changes in probability. For example, a single standard deviation increase in propensity corresponds with a $6,000 increase in family income, a 6 percentage point increase in the probability of a college degree, and a 3 percentage point increase in the probability of approving of George W. Bush. To summarize, high-propensity citizens, as identified by our method, are wealthier, more educated, more likely to attend church, more likely to be employed, more likely to approve of Bush, more conservative, and more Republican. They are more supportive of abortion rights and less supportive of withdrawing troops from Iraq, domestic spending, affirmative action, minimum wage, gay marriage, federal housing assistance, and taxes on wealthy families.

Even though our propensity variable relies on a small number of sparse measures that are readily available from public records, it corresponds strongly with numerous demographic characteristics and issue positions which are highly relevant in American politics. As a result, if GOTV interventions tend to mobilize high-propensity citizens over those with low propensities, they will make the electorate wealthier, more educated, more religious, and more conservative on a number of important issues. However, high-propensity citizens are not more conservative on all issues. In fact, they are more liberal on abortion, and they are no different on health care and immigration. The abortion result is consistent with previous findings that high socioeconomic-status citizens are more supportive of abortion rights (Bartels 2008). The propensity variable is also highly correlated with intensity of preferences. For example, high-propensity citizens are much more likely to have an extreme ideology or a strong party identification. As a result, GOTV interventions

Table 1 Does the Propensity Variable Capture Meaningful Political Characteristics?

	CCES	CCAP
Family income	6117 (305)**	6022 (399)**
College	.065 (.003)**	.059 (.003)**
Church attendance	.052 (.003)**	.040 (.005)**
Married	.097 (.004)**	.041 (.006)**
Employed or retired	.086 (.004)**	.068 (.006)**
Bush approval	.027 (.003)**	.026 (.004)**
Allow abortion	.013 (.004)**	.029 (.004)**
Withdraw troops from Iraq	-.033 (.004)**	-.031 (.006)**
5-point ideology (0 = Very Liberal, 1 = Very Conservative)	.013 (.002)**	.010 (.003)**
Extreme ideology	.037 (.003)**	.015 (.005)**
7-point party ID (0 = Strong Dem., 1 = Strong Rep.)	.009 (.003)**	.007 (.004)*
Strong party ID	.091 (.004)**	.062 (.005)**
Cut domestic spending	.014 (.004)**	
Support affirmative action	-.018 (.004)**	
Raise minimum wage	-.038 (.003)**	
Ban gay marriage	.034 (.004)**	
Support federal housing assistance	-.020 (.004)**	
Support civil unions		.006 (.006)
Support universal healthcare		-.007 (.006)
Raise taxes on wealthy families		-.037 (.005)**
Citizenship for illegal immigrants		.000 (.006)
Sample Size	25,481	16,518

Note: This analysis employs survey data from the 2008 CCES and 2008 CCAP studies to assess whether our propensity variable is meaningfully capturing politically relevant variables. The samples include verified registered voters who were successfully matched to statewide voter records. For both survey samples, we generate our propensity variable using validated turnout in the 2008 presidential election as our dependent variable and using only variables available on voter files (age, race, gender, household size, state, and validated turnout in previous elections) as our independent variables. Then, we test whether this propensity variable is correlated with political characteristics and attitudes by regressing those variables on our propensity variable. Each cell in the table represents a separate regression of a dependent variable on the propensity variable. The first cell entry of 6,117 indicates that a single standard deviation increase in propensity corresponds with an increase in family income of $6,117. Aside from income, ideology, and party ID, all other variables are coded as dummy variables. The propensity variable is positively related to income, education, church attendance, marriage, employment, Bush approval, conservative ideology, Republican party identification, having an extreme ideology, having a strong party identification, and a desire to ban gay marriage. Propensity is negatively correlated with a desire to withdraw troops from Iraq, support for affirmative action, a desire to raise minimum wage, support for federal housing assistance, and support for raising taxes on wealthy families. This analysis shows that when GOTV treatments increase the prevalence of high-propensity citizens in the electorate, they are meaningfully changing the demographic and attitudinal composition of the electorate. Robust standard errors in parentheses; *significant at 0.05, **significant at 0.01.

which mobilize high-propensity citizens are likely to increase the polarization of the electorate in addition to changing its demographic composition. This analysis demonstrates that our propensity variable captures meaningful political differences. Having demonstrated the wealth of information captured by our propensity variable, we apply our test to experimental data.

Applying the Test to 24 Different GOTV Treatments

We obtained our sample by identifying all GOTV field experiments published since 2000 in 10 leading journals where the data are available online.[7] We augmented the sample by directly requesting data from authors. In principle, we can apply our test to any GOTV effort. However, when the average effect of a get-out-the-vote effort is zero, we should not expect to find any interactive effect unless the treatment demobilizes some subset of individuals. For this reason, we restrict our analysis to available experiments with

[7]Our list of journals was comprised of *American Journal of Political Science*, *American Political Science Review*, *American Politics Research*, *British Journal of Political Science*, *Electoral Studies*, *Journal of Politics*, *Political Analysis*, *Political Behavior*, *Political Research Quarterly*, and *Quarterly Journal of Political Science*.

positive and statistically significant average treatment effects.[8] Even within a particular study, we only analyze those experimental treatments which exhibit a statistically significant effect on the average level of voter turnout. Table 2 presents a summary of the published studies utilized in this study. For each study, the table reports the delivery method of the experimental treatment, the electoral context, and the set of covariates available for the calculation of our propensity variables. See the online appendix for more details on each study and the nuanced aspects of our analyses in each setting.

The regression results for each treatment are presented in Table 3. For each experiment, the table presents the coefficient on the *Treatment* variable, indicating the effect of the treatment for the average citizen in the sample. More importantly, the table presents the coefficient on *Treatment*Propensity*, indicating the extent to which the treatment effect changes as propensity increases. Because the *Propensity* variable is recoded so that the mean equals 0 and the standard deviation equals 1, we can interpret the interactive coefficient as the extent to which the treatment effect increases for every standard deviation increase in propensity. In Figure 3, we show nonparametric analyses, similar to that in Figure 2, for each of these 24 interventions. Overall, we see that GOTV interventions tend to exacerbate the participation gap. The interactive coefficient is positive in 16 out of 24 cases and positive and statistically significant for eight of these cases. Alternatively, we find significant evidence for a reduction in the participation gap for only two out of 24 interventions.

How should we interpret the 14 cases where we find no statistically significant interactive coefficient? In many cases, our test is underpowered and unlikely to detect a real, substantively significant effect. For example, in the Stonybrook experiment (N06), we are unable to statistically reject interactive effects as low as −7.5 percentage points or as high as 4.7 percentage points—in other words, the interactive effect could be hugely positive or negative. However, in other cases where we have more statistical power, our lack of a statistically significant coefficient is a sign that even if there is an interactive effect, its substantive magnitude is likely small. For example, for the professional phone-bank experiment (N07), we can statistically reject any interactive effect that is less than −0.6 percentage points or greater than 0.9 percentage points. For this reason, we urge readers to interpret the substantive size of the coefficients and standard errors in Table 3 without overly emphasizing statistical significance.

In the final row of Table 3, we conduct a pooled analysis by combining observations from all experiments in a single regression. In total, 319,251 individuals received one of these GOTV treatments, and 848,521 were assigned to a control group, receiving no such treatment. With these pooled data, we assess the overall effect of these GOTV treatments on the participation gap. For nearly 1.2 million individuals, we regress voter turnout on a treatment dummy variable, the interaction of the treatment with each propensity score, study fixed effects, and the interaction of each study with the propensity score. The inclusion of study dummy variables and study-propensity interactions is necessary because the treatment is random within each study but not between studies.[9] Overall, these treatments exhibit a large positive effect on voter turnout for the average individual, 3.3 percentage points. However, this treatment effect is much stronger for high-propensity individuals. For every standard deviation increase in propensity, the treatment effect increases by 0.5 of a percentage point, on average. This suggests that GOTV treatments, on average, increase turnout by 4.3 percentage points for those whose propensity score is 2 standard deviations above the mean. However, these treatment effects are much weaker, only 2.3 percentage points, for low-propensity citizens. On average, GOTV mobilization effects are more than 85% greater for the highest propensity individuals compared to the lowest. As a result, the typical treatment exacerbates the participation gap, despite the fact that GOTV interventions are often designed to reduce this gap. In the online appendix, we describe additional tests which ensure that our results are not driven by "deadweight" or other issues with the quality of data available on voter files.

[8]While the demobilization of a subset of GOTV subjects is theoretically possible (e.g., Mann 2010), we find no systematic evidence of this. We checked for interactive effects in the experiments reported in Nickerson (2007a), Vavreck (2007), Green and Vavreck (2008), Gerber, Karlan, and Bergan (2010), and Shaw et al. (2012), all of which had close to zero average effect, and we found no interactive effects.

[9]Suppose 50% of subjects are assigned to treatment in one experiment, and only 10% of subjects are assigned to treatment in another. The treated and control individuals would not be comparable between studies, but they are comparable within studies. To prevent the different treatment rates from influencing our estimates, we include study fixed effects along with study-propensity interactions. This ensures that only the within-study variation (that which is randomly assigned) contributes to our estimates.

TABLE 2 **Summary of Studies and Available Data**

			Available Covariates								
Study	**Method**	**Context**	**Age**	**Race**	**Gender**	**Household Size**	**Geography**	**Vote History**	**Party Registration**	**Registration Year**	**Survey Responses**
Gerber and Green 2000a (GG00)	Multiple	1998 General—New Haven, CT	X				X	X	X		
Gerber, Green, and Nickerson 2003 (GGN03)	Door	2001 Local—Multiple Cities	X	X	X			X	X		
Nickerson 2006 (N06)	Phone	2000 and 2001—Multiple Cities	X		X		X	X	X	X	
Nickerson 2007b (N07)	Phone	2002 General—Multiple Cities	X			X	X	X	X		
Gerber, Green, and Larimer 2008 (GGL08)	Mail	2006 Primary—Michigan	X		X	X		X			
Middleton and Green 2008 (MG08)	Door	2004 General—Multiple States	X		X		X	X			
Nickerson 2008 (N08)	Door	2002 Primary—Multiple Cities	X		X		X	X	X		
Dale and Strauss 2009 (DS09)	Text	2006 General—Multiple States	X	X	X		X	X		X	
Gerber, Green, and Larimer 2010 (GGL10)	Mail	2007 Local—Michigan						X			
Gerber, Huber, and Washington 2010 (GHW10)	Mail	2008 Primary—Connecticut	X		X	X		X		X	X
Nickerson and Rogers 2010 (NR10)	Phone	2008 Primary—Pennsylvania	X	X	X	X	X	X			

Note: The table summarizes the available data employed for our study. Our data is drawn from 11 previously published studies, numerous different methods of delivery, and numerous types of electoral settings. To construct our propensity variable, we employ any pretreatment variables that can help us to predict voter turnout: age, race, gender, household size, geography, vote history, party registration, registration year, and in one case, survey responses.

TABLE 3 **Summary of Results**

Intervention	Treatment	Treatment*Propensity	N-Treated	N-Control	Study
Mail	.016 (.008)*	.002 (.007)	7,679	11,665	GG00
Door	.040 (.011)**	-.006 (.009)	2,877	11,665	–
Mail+Door	.037 (.012)**	.004 (.010)	1,853	11,665	–
Phone+Mail+Door	.031 (.015)*	.026 (.013)*	1,207	11,665	–
Bridgeport	.049 (.020)*	.052 (.025)*	895	911	GGN03
Detroit	.027 (.009)**	-.020 (.006)**	2,472	2,482	–
Minneapolis	.027 (.013)*	.027 (.010)**	1,409	1,418	–
St. Paul	.035 (.016)*	-.015 (.011)	1,104	1,104	–
Stonybrook	.071 (.031)*	-.014 (.031)	680	279	N06
Volunteer	.008 (.004)*	-.004 (.004)	26,565	27,221	N07
Professional	.016 (.004)**	.001 (.004)	27,496	27,221	–
Prof.+Vol.	.015 (.004)**	-.003 (.004)	27,452	27,221	–
Civic Duty	.018 (.003)**	.002 (.003)	38,218	191,243	GGL08
Hawthorne	.025 (.003)**	.008 (.003)**	38,204	191,243	–
Self	.048 (.003)**	.008 (.003)**	38,218	191,243	–
Neighbors	.080 (.003)**	.016 (.003)**	38,201	191,243	–
MoveOn	.016 (.004)**	-.010 (.005)*	23,384	22,893	MG08
Minneapolis	.038 (.016)*	.051 (.017)**	876	1,748	N08
Text Message	.030 (.010)**	-.017 (.010)	4,007	4,046	DS09
Civic Duty	.017 (.008)*	.017 (.009)	3,238	353,341	GGL10
Shame	.064 (.006)**	.029 (.007)**	6,325	353,341	–
Pride	.041 (.006)**	.005 (.006)	6,307	353,341	–
Party Reg.	.034 (.008)**	.011 (.010)	1,173	1,175	GHW10
Planning	.010 (.004)**	.000 (.003)	19,411	228,995	NR10
Pooled	**.033 (.001)****	**.005 (.001)****	**319,251**	**848,521**	–

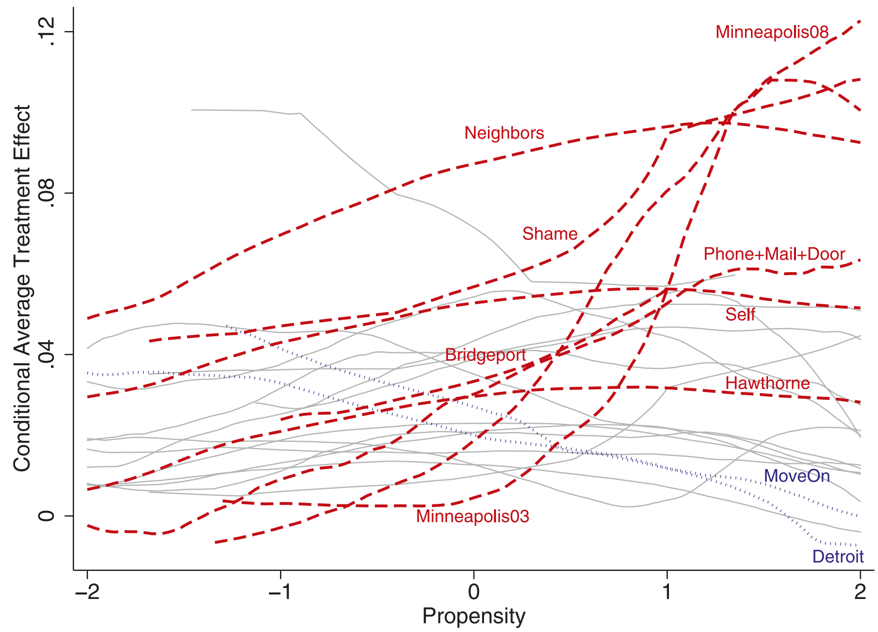
Note: The table presents our regression results for 24 different experimental interventions (descriptions of each intervention are listed in column one of the table). Standard errors are clustered by household or block where applicable and heteroskedasticity-robust in all other cases. We only include those treatments which demonstrate a statistically significant effect on the mean level of voter turnout. The *Treatment* coefficient indicates the treatment effect for the average citizen in the sample. The *Treatment*Propensity* coefficient indicates the extent to which the treatment effect changes as *Propensity* increases. Because the propensity variable is recoded so that the mean equals 0 and the standard deviation equals 1, we can interpret the coefficient as the extent to which the treatment effect increases as propensity increases by one standard deviation. In eight cases, we see a statistically significant, positive coefficient, indicating that the treatment exacerbated in the participation gap. In only two cases do we see statistically significant evidence that the participation gap was reduced. The final row presents a pooled analysis, showing the overall effect of GOTV experiments on the participation gap. Robust standard errors in parentheses; *significant at 0.05, **significant at 0.01.

What Is the Mechanism Behind This Effect?

Readers surprised by our results will naturally ask why these experiments tend to exacerbate the participation gap. Determining the mechanism is difficult to answer in any setting (Bullock, Green, and Ha 2010) but deserves attention nonetheless. Here, we provide several hypotheses and provide evidence that while high-propensity individuals are easier to contact, even within the contacted set, the relationship between propensity to vote and response to treatment persists. First, the correlation of voting propensity with education and political knowledge might result in high-propensity citizens better understanding the treatment. Also, low-propensity citizens may have higher costs involved with voting, so a treatment that reminds voters of the benefits of voting may have to be more powerful to stimulate low-propensity voters than high-propensity voters. Moreover, psychological, social, and economic differences between high- and low-propensity citizens may explain why some people are simply more likely to comply with any policy or experimental intervention. Previous research shows that higher SES individuals are more likely to respond to many interventions including electoral reform (Berinsky 2005), public-health campaigns (Pickett, Luo, and Lauderdale 2005), medical screenings (Wee et al. 2012), public housing (Blundell, Fry, and Walker 1988), and Medicaid (Aizer 2003) even when such interventions are specifically designed to benefit low-SES individuals.

A related potential mechanism lies in citizens' differential probabilities of being successfully contacted.

Figure 3 Variation in 24 GOTV Treatment Effects



Note: The figure presents the conditional average treatment effect for all 24 experimental interventions across different levels of propensity. These plots mimic the bottom panel of Figure 2. The red/dashed curves indicate treatments which significantly exacerbate the participation gap, the blue/dotted curves indicate treatments which significantly reduce the participation gap, and the gray/solid curves indicate treatments which had no statistically significant effect. The curves are cut off in order to avoid extrapolation. For example, if there are no observations in a particular setting with propensities below −1.5, then we show no curve below that point.

High-propensity individuals may be more likely to read their mail, answer their phone, or talk to a canvasser. With our data, we explicitly test whether high-propensity citizens are easier to contact via phone or door-to-door canvassing. For many studies, including direct-mail studies, the researcher is unable to know who actually received the treatment, but for phone and canvassing studies, the researcher can record which individuals or households were contacted. The results of these tests for all available studies are available in the online appendix. On average, high-propensity individuals are much easier to contact via door-to-door canvassing or phone calls. For example, in Gerber and Green's (2000a) New Haven study, a standard deviation increase in propensity corresponds to an extra 5 percentage point chance of canvassing contact and an extra 12 percentage point chance of phone contact. These results suggest that differential contact rates may explain much of the variation in intention-to-treat effects between high- and low-propensity voters. If subsequent interventions hope to improve the participation and representation of low-propensity voters, they must first overcome the challenge or reaching them in the first place.

While differential contact rates are one primary reason that low-propensity citizens are more difficult to mobilize, they are not the only reason. Nickerson (2008) conducted door-to-door canvassing experiments with placebo treatments, allowing us to run separate analyses where we confine our observations only to individuals or households that were contacted by the experimenters. Because of the placebo treatments, we know which individuals in this control group would have been contacted by the canvassers had they been in the treatment group, and we know that they are comparable in expectation to those who were contacted in the treatment group. When we apply our test to only the sample of individuals or households who answered the door, we still estimate large, positive interactive coefficients. Even among the sample of individuals or households willing to answer the door, high-propensity citizens were more responsive to the get-out-the-vote message. More details on these tests are provided in the online appendix. While contact is a significant barrier to mobilizing low-propensity citizens, further barriers remain. Even conditional on receiving the message, low-propensity citizens are less likely to respond to it.

Further information regarding the mechanisms behind our overall findings can be garnered by examining variation in our interactive effects across different interventions. In the online appendix, we explore several sources of this variation. Consistent with Arceneaux and Nickerson (2009), we find that the exacerbating effect of GOTV treatment is greatest in low-salience elections. However, even in high-salience elections, these treatments, on average, increase disparities in participation. We also find that the most effective treatments in terms of increasing overall participation also exhibit the greatest exacerbating effects. Finally, we discuss the peculiar phenomenon that among the studies we examine, the few experimental settings in which the participation gap was reduced involved largely African American samples.

In one sense, why GOTV interventions tend to exacerbate the participation gap is of secondary importance. Whether the mechanism involves knowledge, costs, psychology, contact, or something else, the positive and normative implications of our study are the same. However, knowing that low-propensity citizens are harder to contact provides one promising avenue for future researchers to design interventions that may mobilize them. Reaching the population of interest would significantly mitigate (but not remove entirely) the exacerbating phenomenon that we identify. We hope that the continued application of our test on future experiments will allow us to understand the reasons why and the conditions under which campaign interventions exacerbate or reduce the participation gap.

Conclusion

In analyzing 24 field experiments, we find that two-thirds of GOTV experiments mobilized high-propensity voters to a greater degree than low-propensity voters—thereby exacerbating the participation gap. Moreover, this exacerbating effect is statistically significant in eight cases. Our pooled analysis demonstrates an average exacerbating effect that is substantively large and statistically significant. On average, GOTV methods developed by political scientists, many of which have been subsequently adopted by political campaigns (Gerber et al. 2011; Green and Gerber 2008; Issenberg 2010, 2012), appear to exacerbate the disparities between voters and the voting-eligible population. Because turnout is politically consequential and because our *propensity* variable captures meaningful political differences between individuals and groups, this exacerbating effect should be of concern to both political scientists and practitioners.

These results pose and clarify a challenge to scholars and practitioners interested in political participation. Current mobilization methods employed by researchers and political campaigns appear to be mainly effective in bringing high-propensity citizens to the polls. The search for a reliable mobilization method for bringing underrepresented, low-propensity citizens to the polls continues, largely without resolution, and these results underscore the need for an increased focus on this goal. The method we present for assessing differential treatment effects provides a metric for future experimenters interested in the effects of GOTV interventions on the composition of the electorate. We hope that this tool will aid in the development and assessment of new GOTV methods that are effective in mobilizing low-propensity citizens and reducing the demographic gaps between the electorate and the greater, eligible population.

Even for those uninterested in the participation gap per se, our findings hold practical implications for campaigns and scholars of voter mobilization. Some individuals, particularly those with low propensities, are harder to mobilize than others. As a result, the combinatorial effects of multiple GOTV efforts are called into question. Practitioners may wrongly assume that the effects of one mobilization will combine additively with the effects of others. However, once a campaign has hit a ceiling among high-propensity voters, it may be unable to further increase participation using traditional methods.

The findings of this study also raise an ethical concern for experimenters and practitioners because experimental interventions and mobilization efforts are often conducted with the assumption that raising average participation levels can only be good for democracy. However, the evidence in this article—that voter mobilization tends to exacerbate existing inequalities in the electorate—necessitates a more nuanced perspective. Despite their unquestionable accomplishments in raising mean levels of participation, current GOTV efforts are not the solution to persistent inequalities in the political process. On the contrary, these efforts may contribute to the problem by making the electorate more polarized and less representative of the greater population.

Acknowledgments

Author order is alphabetical. We thank Kevin Arceneaux, Steve Ansolabehere, Peter Aronow, Catherine Choi, Lisa Garcia Bedolla, Don Green, Andy Hall, Ben Lauderdale, Melissa Michelson, Joel Middleton, David Nickerson, Todd Rogers, Jim Snyder, Aaron Strauss, and all authors who made their data available to us.

References

Aizer, Anna. 2003. "Low Take-up in Medicaid: Does Outreach Matter and for Whom?" *American Economic Review Papers and Proceedings* 3 (2): 238–41.

Alvarez, R. Michael, Asa Hopkins, and Betsy Sinclair. 2010. "Mobilizing Pasadena Democrats: Measuring the Effects of Partisan Campaign Contacts." *Journal of Politics* 72 (1): 31–44.

American Political Science Association Task Force (APSA). 2004. "American Democracy in an Age of Rising Inequality." *Perspectives on Politics* 2 (4): 651–66.

Angrist, Joshua D., and Jorn-Steffen Pischke. 2009. *Mostly Harmless Econometrics*. Princeton, NY: Princeton University Press.

Ansolabehere, Stephen D. 2011. "CCES Common Content, 2008." http://hdl.handle.net/1902.1/14003 V4.

Anzia, Sarah F. 2011. "Election Timing and the Electoral Influence of Interest Groups." *Journal of Politics* 73 (2): 412–27.

Arceneaux, Kevin, and David W. Nickerson. 2009. "Who Is Mobilized to Vote? A Re-Analysis of 11 Field Experiments." *American Journal of Political Science* 53 (1): 1–16.

Avey, Michael J. 1989. *The Demobilization of American Voters*. Westport, CT: Greenwood Press.

Bartels, Larry M. 2008. *Unequal Democracy*. Princeton, NJ: Princeton University Press.

Bartels, Larry M. 2009. "Economic Inequality and Political Representation." *The Unsustainable American State*, ed. Lawrence Jacobs and Desmond King. Oxford: Oxford University Press, 167–98.

Blundell, Richard, Vanessa Fry, and Ian Walker. 1988. "Modelling the Take-up of Means-tested Benefits: The Case of Housing Benefits in the United Kingdom." *Economic Journal* 98: 58–74.

Berinsky, Adam J. 2005. "The Perverse Consequences of Electoral Reform in the United States." *American Politics Research* 33 (4): 471–91.

Bullock, John G., Donald P. Green, and Shang E. Ha. 2010. "Yes, But What's the Mechanism? (Don't Expect an Easy Answer)." *Journal of Personality and Social Psychology* 98 (4): 550–58.

Citrin, Jack, Eric Schickler, and John Sides. 2003. "What if Everyone Voted? Simulating the Impact of Increased Turnout in Senate Elections." *American Journal of Political Science* 47 (1): 75–90.

Dahl, Robert A. 2006. *On Political Equality*. New Haven, CT: Yale University Press.

Dale, Allison, and Aaron Strauss. 2009. "Don't Forget to Vote: Text Message Reminders as a Mobilization Tool." *American Journal of Political Science* 53 (4): 787–804.

Fowler, Anthony. 2013. "Electoral and Policy Consequences of Voter Turnout: Evidence from Compulsory Voting in Australia." *Quarterly Journal of Political Science* 8 (2): 159–82.

Gabler, Nicole, Naihua Duan, Diana Liao, Joann G. Elmore, Theodore G. Ganiats, and Richard L. Kravits. 2009. "Dealing with Heterogeneity of Treatment Effects: Is the Literature Up to the Challenge?" *Trials* 10: 43.

Garcia Bedolla, Lisa, and Melissa R. Michelson. 2012. *Mobilizing Inclusion: Transforming the Electorate through Get–out-the-vote Campaigns*. New Haven, CT: Yale University Press.

Gerber, Alan S., and Donald P. Green. 2000a. "The Effects of Canvassing, Direct Mail, and Telephone Contact on Voter Turnout: A Field Experiment." *American Political Science Review* 94: 653–63.

Gerber, Alan S., and Donald P. Green. 2000b. "The Effect of a Nonpartisan Get-Out-The-Vote Drive: An Experimental Study of Leafletting." *Journal of Politics* 62 (3): 846–57.

Gerber, Alan S., Donald P. Green, and David Nickerson. 2003. "Getting Out the Vote in Local Elections: Results from Six Door-to-Door Canvassing Experiments." *Journal of Politics* 65 (4): 1083–96.

Gerber, Alan S., Donald P. Green, and Christopher Larimer. 2008. "Social Pressure and Voter Turnout: Evidence from a Large-Scale Field Experiment." *American Political Science Review* 102 (1): 33–48.

Gerber, Alan S., Dean Karlan, and Daniel Bergan. 2009. "Does the Media Matter? A Field Experiment Measuring the Effect of Newspapers on Voting Behavior and Political Opinions." *American Economic Journal: Applied Economics* 1 (2): 35–52.

Gerber, Alan S., Donald P. Green, and Christopher W. Larimer. 2010. "An Experiment Testing the Relative Effectiveness of Encouraging Voter Participation by Inducing Feelings of Pride or Shame." *Political Behavior* 32 (3): 409–22.

Gerber, Alan S., Gregory A. Huber, and Ebonya Washington. 2010. "Party Affiliation, Partisanship, and Political Beliefs: A Field Experiment." *American Political Science Review* 104 (4): 720–44.

Gerber, Alan S., James G. Gimpel, Donald P. Green, and Daron R. Shaw. 2011. "How Large and Long-Lasting Are the Persuasive Effects of Televised Campaign Ads? Results from a Randomized Field Experiment." *American Political Science Review* 105 (1): 135–50.

Gilens, Martin. 2005. "Inequality and Democratic Responsiveness." *Public Opinion Quarterly* 69 (5)778–796.

Gilens, Martin. 2012. *Affluence and Influence: Economic Inequality and Political Power in America*. Princeton, NJ: Princeton University Press.

Green, Donald P., and Holger L. Kern. 2012. "Modeling Heterogeneous Treatment Effects in Survey Experiments with Bayesian Additive Regression Trees." *Public Opinion Quarterly*. Forthcoming.

Green, Donald P., and Melissa R. Michelson. 2009. "ACORN Experiments in Minority Voter Mobilization." In *The People Shall Rule*, ed. Robert Fisher. Nashville, TN: Vanderbilt University Press, 235–50.

Green, Donald P., and Alan S. Gerber. 2008. *Get Out the Vote: How to Increase Voter Turnout*. Washington, DC: Brookings Press.

Green, Donald P., and Lynn Vavreck. 2008. "Analysis of Cluster-Randomized Field Experiments: A Comparison of Alternative Estimation Approaches." *Political Analysis* 16 (2): 138–52.

Griffin, John D., and Brian Newman. 2005. "Are Voters Better Represented?" *Journal of Politics* 67 (4): 1206–27.

Hajnal, Zoltan, and Jessica Trounstine. 2005. "Where Turnout Matters: The Consequences of Uneven Turnout in City Politics." *Journal of Politics* 67 (2): 515–36.

Hansford, Thomas G., and Brad T. Gomez. 2010. "Estimating the Electoral Effects of Voter Turnout." *American Political Science Review* 104 (2): 268–88.

Issenberg, Sasha. 2010. "Nudge the Vote." *New York Times Magazine*, October 29, MM28.

Issenberg, Sasha. 2012. *The Victory Lab: The Secret of Winning Campaigns*. New York: Crown.

Jackman, Simon, and Lynn Vavreck. 2010. Cooperative Campaign Analysis Project, Release 2.1. YouGov, Inc., Palo Alto, CA.

Key, V. O. 1949. *Southern Politics in State and Nation*. New York: Alfred A. Knopf.

Lijphart, Arend. 1997. "Unequal Participation: Democracy's Unresolved Dilemma." *American Political Science Review* 91 (1): 1–14.

Mann, Christopher B. 2010. "Is There Backlash to Social Pressure? A Large-Scale Field Experiment on Voter Mobilization." *Political Behavior* 32 (3): 387–407.

Michelson, Melissa R., Lisa García Bedolla, and Donald P. Green. 2008. "Mobilizing the Registered: Tactics for Improving Voter Turnout." In *Voting in America*, ed. Morgan E. Felchner. New York: Praeger, 198–211.

Middleton, Joel A., and Donald P. Green. 2008. "Do Community-Based Mobilization Campaigns Work Even in Battleground States? Evaluating the Effectiveness of MoveOn's 2004 Outreach Campaign." *Quarterly Journal of Political Science* 3 (1): 63–82.

Nickerson, David W. 2006. "Volunteer Phone Calls Can Increase Turnout: Evidence from Eight Field Experiments." *American Politics Research* 34 (3): 271–92.

Nickerson, David W. 2007a. "Does Email Boost Turnout?" *Quarterly Journal of Political Science* 2: 369–79.

Nickerson, David W. 2007b. "Quality is Job One: Professional and Volunteer Voter Mobilization Calls." *American Journal of Political Science* 51 (2): 269–82.

Nickerson, David W. 2008. "Is Voting Contagious? Evidence from Two Field Experiments." *American Political Science Review* 102 (1): 49–57.

Nickerson, David W., and Todd Rogers. 2010. "Do You Have a Voting Plan? Implementation Intentions, Voter Turnout, and Organic Plan Making." *Psychological Science* 21 (2): 194–99.

Pickett, Kate E., Ye Luo, and Diane S. Lauderdale. 2005. "Widening Social Inequalities in Risk for Sudden Infant Death Syndrome." *American Journal of Public Health* 95 (11): 1976–81.

Pocock, Stuart J., Susan E. Assmann, Laura E. Enos, and Linda E. Kasten. 2002. "Subgroup Analysis, Covariate Adjustment, and Baseline Comparisons in Clinical Trial Reporting: Current Practice and Problems." *Statistics in Medicine* 23: 2509–25.

Putnam, Robert D. 2000. *Bowling Alone*. New York: Simon and Schuster.

Schattschneider, E. E. 1960. *The Semisovereign People: A Realist's View of Democracy in America*. Austin, TX: Holt, Rinehart, and Winston.

Schlozman, Kay Lehman, Sidney Verba, and Henry E. Brady. 2012. *The Unheavenly Chorus: Unequal Political Voice and the Broken Promise of American Democracy*. Princeton, NJ: Princeton University Press.

Shaw, Daron R., Donald P. Green, James G. Gimpel, and Alan S. Gerber. 2012. "Do Robotic Calls from Credible Sources Influence Voter Turnout or Vote Choice? Evidence from a Randomized Field Experiment." *Journal of Political Marketing* 11 (4): 231–45.

Vavreck, Lynn. 2007. "The Exaggerated Effects of Advertising on Turnout: The Dangers of Self-Reports." *Quarterly Journal of Political Science* 2 (4): 325–43.

Verba, Sidney, Kay Lehman Schlozman, and Henry E. Brady. 1995. *Voice and Equality*. Cambridge, MA: Harvard University Press.

Wee, Liang En, Gerald Choon Huat Koh, Wei Xin Yeo, Run Ting Chin, Jolene Wong, and Braden Seow. 2012. "Screening for Cardiovascular Disease Risk Factors in an Urban Low-Income Setting at Baseline and Post Intervention: A Prospective Intervention Study." *European Journal of Cardiovascular Prevention and Rehabilitation*. Forthcoming.

Wolfinger, Raymond E., and Steven J. Rosenstone. 1980. *Who Votes?* New Haven, CT: Yale University Press.

Ryan D. Enos is Assistant Professor in the Department of Government at Harvard University, Cambridge, MA 02138.

Anthony Fowler is Assistant Professor in the Harris School of Public Policy Studies at the University of Chicago, Chicago, IL 60637.

Lynn Vavreck is Associate Professor of Political Science and Communication at the University of California at Los Angeles, Los Angeles, CA 90065.