Supplementary Materials

Table of Contents

[***1.*** ***Theoretical Framework: The Calculus of Voting*** 1](#_Toc129961277)

[***2.***  ***Theoretical Framework: “Voting-as-a-Habit” Literature and the PMBC Model*** 5](#_Toc129961278)

[***3.*** ***Data*** 13](#_Toc129961279)

[***4.*** ***Justifying the Lack of Pre-Registration*** 18](#_Toc129961280)

[***5.*** ***Analysis*** 20](#_Toc129961281)

[*5.1: Upstream Effects* 20](#_Toc129961282)

[*5.2: Checking for Weak Instruments* 21](#_Toc129961283)

[*5.3: Additional Predictions* 24](#_Toc129961284)

[*5.4: Robustness Checks* 25](#_Toc129961285)

[***6.*** ***Additional Discussion*** 27](#_Toc129961286)

[***7.*** ***Additional Tables and Figures*** 28](#_Toc129961287)

[***8.*** ***Anonymized Canvasser Script and Training Documents*** 32](#_Toc129961288)

[***9.*** ***Additional References*** 41](#_Toc129961289)

1. ***Theoretical Framework: The Calculus of Voting***

To construct a theoretical model of why casting a vote makes an individual more likely to vote in subsequent election, we must first establish the mechanism for why any citizen votes in the first place. Our understanding of why voters turn out to the polls has developed considerably since the canonical formalizations of Riker and Ordershook (1968) and Downs (1957). Their pioneering work noted that there is some cost of voting and, in large elections, a single vote is highly unlikely to matter, so what is the benefit of turning out to vote? Why do so many people still incur the personal cost (e.g., time costs, transportation costs, etc.) of voting even though it is highly improbable that their vote will decide the outcome of an election? Downs (1957) wrote of the perceived likelihood of being the pivotal voter and Riker and Ordershook (1968) added the “warm glow” variable to the model. In other words, a voter considers how close the election will be in their mental calculus, but also anticipates a warm glow after fulfilling one’s civic duty at the polls. And so, the voter may incur the cost of voting even in elections that aren’t close. Their reliance on the concept of the “rational voter” and insistence on including a “pivotality” variable (i.e., the expected probability that one’s vote will decide an election) set off a heated debate between the Downsian political economists and psychology-influenced political scientists (for summaries see Blais, 2000; Green and Shapiro, 1996). As such, there was a proliferation of theoretical models and empirical research that attempted to better understand what the Downsian school called the “paradox of voting” (for a summary of the relevant approaches and a meta-analysis of empirical research, see Smets & Van Hem, 2013). But though some elements of the original model continue to be controversial (i.e., do people really think about the probability that they can swing a presidential election?), even some of the leading critics of the rational voter model came to use cost-benefit equations to model an individual’s decision to vote or abstain in their own work (e.g., see Gerber, Green and Larimer, 2008). Therefore, as a first step, I assume that a cost-benefit equation is one key determinant of a would-be voter’s decision-making process.

There is also some consensus as to several of the variables included in that equation: there likely exists some cost of voting and that there may be some *intrinsic* benefit of voting (e.g., Gerber, Green & Larimer, 2008; Meredith, 2009; DellaVigna et al., 2016; Fujiwara, Meng and Vogl, 2016) which we conceptualize as the original “warm glow” variable from Riker and Ordershook (1968).[[1]](#footnote-1) Critics of the “pivotality” variable (i.e., the variable that denotes the expected probability of casting the vote that swings an election) maintain that it is likely so small that it could be treated as zero (e.g., Gerber, Green, & Larimer, 2008), but because this variable is still included in many economic models of voting (e.g., DellaVigna et al., 2016; Fujiwara, Meng and Vogl, 2016), I include it in my cost-benefit equation with the acknowledgement that it may be zero (for a complete formalization, see DellaVigna et al., 2016).

Aside from considering these well-established cost and benefits of voting, it has become increasingly clear that there is a powerful social element to voting. Gerber, Green and Larimer (2008) experimentally induced prospective voters to turn out via social pressure by publicizing their neighbors’ voting records.[[2]](#footnote-2) The social pressure intervention has been replicated in many other field experiments (for overviews and meta-analyses, see Green & Gerber, 2019; Green, McGrath and Aronow, 2013) and the concept of “social image” was finally incorporated into formal models of voting (e.g., DellaVigna et al., 2016). In other words, there is a widespread belief that voting is an important part of what it means to be a good citizen,[[3]](#footnote-3) so there may be negative social repercussions for abstaining and potentially positive social rewards for turning out to the polls. And indeed, scholars have found that voters relinquish monetary rewards to tell others that they vote (DellaVigna et al., 2016) and pay to find out if others voted (Gerber et al., 2016). This is all to say that there are strong norms that citizens should vote, so people may turn out to vote because they perceive or anticipate some form of real or imagined social sanctions for being a non-voter (Fujiwara, Meng, and Vogl, 2016; DellaVigna, 2016; Gerber et al., 2016).

For the purposes of this paper, I use the DellaVigna et al. (2016) model as the core cost-benefit model[[4]](#footnote-4) for three reasons: the costs and benefits in this model are commonly used in other formalizations, it has extensive empirical validation backing its assertions, and it is parsimonious. However, this paper does not hinge on the choice of cost-benefit models and alternative models could easily supplant the DellaVigna et al. (2016) model in my analyses.[[5]](#footnote-5)

In the DellaVigna et al. (2016) model, there are three benefits and two costs. The benefits are

1. the expected utility of being the pivotal voter (i.e., your vote decides the election),
2. the “warm glow” of voting,
3. and the social utility of voting (e.g., being able to truthfully report having had voted to peers and family)

whereas the costs are:

1. the transactional cost of voting (time, transportation costs, opportunity cost, etc.),
2. and the social cost of not voting (e.g., either telling the truth and being socially sanctioned for not voting or the cognitive cost of lying about one’s abstention).[[6]](#footnote-6)

***2.***  ***Theoretical Framework: “Voting-as-a-Habit” Literature and the PMBC Model***

Habitual actions are “automatic,” which implies an “absence of awareness, conscious control, mental effort and deliberation” (Lally and Gardner, 2013; 137). It is unlikely that even the most consistent of voters are “unaware” going to the polls as one would drive down a familiar route.[[7]](#footnote-7) The second issue is that to translate a deliberative behavior to an automatic behavior, substantial repetition is required to achieve any level of automaticity—for instance, a study looking at the adoption of healthy daily behaviors found that the behavior became automatic only after 66 days on average (Lally et al., 2010). If we assume any habitual action needs to be repeated 66 times, even the individual who votes in three elections every two years (e.g., a municipal election, a party primary, and a general election) would only be expected to develop a voting “habit” after 44 years of being a voter. This stands in stark contrast to the fact that much of the voting-as-a-habit literature examined much shorter time horizons but still found turnout persistence. The economic model is compatible with vote persistence (i.e., if it incorporates past election turnout in the cost-benefit calculus described in Section 1), but it so narrow that it provides no guidance as to the mechanism behind why one vote is more likely to lead to another. [[8]](#footnote-8)

The PMBC is able to bridge the two literatures by incorporating both deliberate and heuristic[[9]](#footnote-9) behaviors in a more generalized way. The advantages of applying the PMBC to turnout persistence are as follows. First, unlike narrow definitions of turnout persistence, the PMBC provides clear, testable predictions as to which circumstances are more favorable for turnout persistence. Second, the PMBC avoids the paradox of how such an infrequent and clearly non-automatic behavior can be construed as habitual in the strict psychological sense. Third, the PMBC is a parsimonious, rich, and coherently self-contained theoretical framework. Applying the PMBC to turnout persistence allows us to reconcile inconsistencies in the empirical voting-as-a-habit literature. For instance, some studies found that previous voting in one type of election does not translate to downstream voting in a different type of election (Michelson, 2003; Hill & Kousser, 2016), but other studies found turnout persistence across different types of elections (Garcia Bedolla and Michelson, 2012; Gerber, Green, & Shachar 2003).

There have been two major approaches to isolating turnout persistence from confounding variables: 1) a regression discontinuity at the point when a voter becomes eligible to vote, 2) a 2SLS regression on downstream turnout among compliers in randomized (e.g., GOTV nudges) or natural experiments (e.g., rainfall). I discuss each of these approaches in turn and apply the PMBC model to make sense of their findings.

The first approach is a temporal regression discontinuity design that uses the date at which a voter becomes eligible to vote (Meredith, 2009; Franklin & Hobolt, 2011; Dinas, 2012; Coppock and Green, 2016). In other words, this approach compares the long-term vote record of the voter who turned 18 just before an election to the long-term vote record of the voter who turned 18 just after the same election. This design exploits the fact that people cannot self-select to either group and assumes that the only meaningful difference between the two groups is eligibility to vote. With this approach, researchers have found that voting habits persist for at least twenty years (Coppock and Green, 2016), which is consistent with the broader psychological literature on habit-formation—repeated habitual actions should be self-reinforcing (e.g., Lally and Gardner, 2013).

To better see how this research would be interpreted by the proposed PMBC model, it is useful to explicate each step for a young voter who becomes eligible to vote shortly before an election. Under the PMBC model the Situation Stage is the upcoming election that a young voter will be eligible for. Next, that young voter may be made aware of the upcoming elections via campaign and non-profit outreach, media coverage, and conversations with friends and family (i.e., the Attention Stage). If they are not made aware of the upcoming election, they will therefore abstain without ever assessing the costs and benefits of voting. If they are made aware of the upcoming election that they are eligible to vote in, the young voter will proceed to the Appraisal Stage, since they have not voted before and, thus, have no cognitive shortcuts past the Appraisal Stage. As mentioned earlier, for the Appraisal Stage, I use a variant of DellaVigna et al.’s (2016) model, which has three benefits and two costs. The benefits are the expected utility of being the pivotal voter (i.e., your vote decides the election), the “warm glow” of voting, and the social utility of voting (e.g., being able to truthfully report having had voted to peers and family). The costs are the transactional cost of voting (time, transportation costs, opportunity cost, etc.) and the social cost of not voting (e.g., either telling the truth and being socially sanctioned for not voting or the cognitive cost of lying about one’s abstention).

What is important to note is that the new voters must estimate these variables with limited information and those estimates may be wildly inaccurate. Additionally, these variables can be manipulated exogenously. For instance, social pressure GOTV outreach explicitly (e.g., Gerber, Green and Larimer, 2008) or implicitly (e.g., Rogers, Ternovski and Yoeli, 2016) emphasizes that whether or not someone voted is public record and this tactic has consistently increased turnout even in competitive elections (e.g., Green, McGrath and Aronow, 2013). This is to say that, under this theoretical model, receiving a piece of mail that emphasizes that whether or not you voted is public record and people may ask you about it is likely to increase the expected cost of not voting. The key takeaway is that social and campaign outreach that calls attention to an upcoming election can also affect the perceived variables in the cost-benefit model used in the Appraisal stage of the PMBC decision-making model. Hence, the young voter in our example would either choose to vote or abstain and proceed to the Situation Stage of the PMBC model. In this example, the Situation Stage can be construed as encompassing both the experience at the polls and the subsequent social interactions about voting.[[10]](#footnote-10)

At the Situation Stage, the voter will therefore get important information that updates the values of each variable used in the cost-benefit analysis, but the voter and the non-voter have systematically different sets of information. For one, the voter can update the “warm glow” variable depending on one’s experiences at the polls. Someone who had a positive experience may assign a higher value to their “warm glow” benefit, while someone who waited in a long line and was affected by voter ID laws (e.g., they cast a “provisional” ballot because their identification did not meet state requirements), may decrease the value assigned to the “warm glow” variable.[[11]](#footnote-11) The non-voter would thus only be able to update this variable through second-hand information (e.g., talking to friends and family who voted). And as Meredith (2009) points out, the voter may also be able to reduce the transaction cost of voting by getting important first-hand logistical information about voting (e.g., knowing where to go, what the ballot looks like, how voting machines work, knowing more about the candidates). The non-voter would not get the benefit of this cost reduction. The voter and non-voter will get markedly different pieces of information that affect the social cost and social benefit of voting. The non-voter will get more information about the social cost of not voting and the voter will get more information about the social benefit.

On the next election day (which may be years later), our example youth voter encounters a similar situation and must again get some reminder to bring their attention to the election. If they are made aware of the election they may again proceed to the Appraisal stage with updated cost-benefit values, which may be updated again in a similar way. Ultimately, this updating of cost-benefit terms can also be considered in a reduced form: was the decision to vote or not vote “correct”? In other words, in casting a vote did the voter feel that the benefits justified the costs and vice versa. This aspect is critical in habit-formation because a desirable behavior (i.e., voting) only starts to become a habit if there is a positive benefit to that behavior and the contextual cues drawing attention to the upcoming election are similar (Duckworth and Gross, 2020). At some point, an individual can therefore sidestep the Appraisal Stage because the cost-benefit maximization equation has led to positive results in the past.

The results from this RDD literature could imply that over the course of multiple elections, newly eligible voters develop a persistent heuristic-driven shortcut past the Appraisal stage. But there does remain one alternative explanation that I alluded to earlier: divergences in the voting record of the voter who became eligible to vote just before an election versus that of the voter who didn’t may be driven by divergences in the information boluses the two types of individuals receive (i.e., this is an exclusion violation). The key part is that newly eligible voters may have highly inaccurate estimates of the costs and benefits of turning out to the polls and voting for the first time could potentially provide a major correction to a misestimated cost or benefit. The amount of information a new voter receives from voting in each successive election is likely diminishing on average. As such, to better disentangle whether it is the act of voting or the information received from the voting experience that is responsible for future votes, it would be useful to examine the development of heuristic shortcuts among more seasoned voters.

Some of the empirical research in the voting-as-a-habit literature does indeed make use of randomized controlled experiments to include the development of turnout persistence among older voters (e.g., Coppock & Green, 2016). This population is composed of people who have already made it a habit to abstain in elections or those individuals who have gotten a good sense of the costs and benefits of voting but did not become frequent voters. It is worth emphasizing that the PMBC model was constructed to gain a better understanding of how to modify existing, undesirable behaviors and supplant them with healthy, desirable behaviors (e.g., healthier eating choices) (Duckworth and Gross, 2020), which makes the PMBC model particularly suitable for studying the long-term implications of inducements to vote among a broader population of individuals who have “unhealthy voting habits.”

For this population of infrequent voters, the voting-as-a-habit literature makes use of 2SLS regression to look at the long-term vote histories of individuals who were induced to vote due to random or as-if random interventions. Most of this literature using this approach made use of randomized field experiments that were originally conducted not to build a voting habit but to simply get voters to vote in a particular upstream election (e.g., Gerber, Green, Shachar, 2003; Michelson, 2003; Garcia Bedolla and Michelson, 2012; Coppock and Green, 2016). Among the most successful of these interventions is the social pressure nudge (Green, McGrath, and Aronow, 2013) and this outreach explicitly targets the social cost and social benefit terms in the cost-benefit equation at the Appraisal Stage.

While these experiments analyze a more experienced voting base and therefore likely avoid the problem of large information boluses from turning up at the polls for the first time, they may have a different exclusion restriction violation. As Fujiwara, Meng, and Vogl (2016) note, it is possible that the observed downstream effects of these nudges are due to lasting changes in the cost-benefit equation during the Appraisal Stage as opposed to a cognitive shortcut past the Appraisal Stage. To be clear, social pressure nudges aim to increase both the social cost and social benefit of voting by emphasizing that whether or not someone voted is public record. The treated individual may thus anticipate more people asking about whether they voted, but after a few elections they may realize that being asked about voting is not as frequent a situation as they expect, which may lead to cessation of voting after just a few elections. And indeed, this is what Coppock and Green (2016) see in their analysis of downstream voting among compliers of GOTV field experiments, which stands in stark contrast to their findings using eligibility criteria of young voters (where frequent voting persists two decades later).[[12]](#footnote-12) It is also worth noting that the initial social pressure outreach studied in the experiments analyzed by Coppock and Green (2016) had some level of novelty and may have triggered awareness at the Attention Stage in a way that temporarily broke a habit of non-voting, by both providing a different contextual cue and changing the expected social costs of continuing to abstain.

One additional strategy used in the past to overcome this exclusion restriction is the use of unexpected shocks like rainfall (Fujiwara, Meng and Vogl, 2016), but as the authors themselves note, it is still very possible that people will still commit a misattribution error and allow an obviously unrelated event color their perceptions of future events.[[13]](#footnote-13) This rain study is also limited by the fact that the unit of observation was county-level data and as the authors note their habit effects may be amplified by social interactions among voters. Indeed, it is worth noting that anticipated social costs of abstention correlate strongly with local turnout rates (Gerber et al., 2016), so a lower-than-anticipated turnout in a county after heavy rain may be communicated by peers and the media, thereby causing an adjustment of the social value of voting in the Appraisal Stage of the PMBC model.

The data used in this paper does not necessarily circumvent all potential violations of the exclusion restriction, but it does provide a different perspective in that the exogenous shock is not solely focused on increasing upstream turnout (as with social pressure nudges),[[14]](#footnote-14) but also cultivating a habit in the same way public health interventions attempt to make health behaviors automatic for individuals (e.g., Lally et al 2010). How the cost-benefit equation of voting would change from this intervention is less clear aside from potentially reducing the transaction cost of voting.[[15]](#footnote-15)

1. ***Data***

I partnered with a large-scale US labor organization that continually runs door-to-door outreach programs with millions of working class Americans. Their outreach strategy focuses on building a relationship with voters through recurring contact via in-person conversations; importantly, both individuals in the control and treatment groups get subsequent in-person contact.[[16]](#footnote-16) Aside from recurring non-election contact ranging from attempts to increase health insurance uptake to issue-based advocacy, the organization also does voter mobilization during election years. Election year contact consists of freeform conversations at the door with several guiding questions, a Get-Out-the-Vote (GOTV) component, and a persuasion component regarding some issue and/or candidate. The GOTV component is rooted in behavioral psychology research on encouraging habit-formation. Particularly, the script has the canvasser ask about the voter’s plan to vote (i.e., “When do you normally go- on your way to work? When you get off? Just before dinner?”), which has proven to be an effective GOTV intervention even in high salience election contexts (Nickerson and Rogers, 2010).[[17]](#footnote-17) The script also includes a component encouraging voters to talk to their friends and family about the election. (A de-identified sample script can be found in Section 7.)

This treatment is distinct from prior analyses of downstream turnout effects for three reasons. First, the treatments are specifically tailored around building a voting-habit. Implementation intentions (commonly known as “plan-making nudges”) have proven to have medium to large effects on translating intentions to actions (Gollwitzer and Sheeran, 2008). This component is therefore designed to affect the formation of a cognitive shortcut unlike many of the prior field experiments in the voting-as-a-habit literature.[[18]](#footnote-18) That said, it should be noted that the component of the script that encourages participants to talk to their friends and family about the election could conceivably alter one’s social cost/benefit of voting in the Appraisal Stage. Still, because it does not explicitly reference that whether or not someone votes is public record, it is less likely to affect the cost-benefit equation as compared to social pressure GOTV nudges. Social pressure nudges, in contrast, may illustrate that the cost of not voting is higher than expected. An encouragement to simply talk about the election is only likely to affect the social cost of not voting in very indirect ways and is therefore much less likely to affect the cost-benefit calculation than the social pressure intervention. Additionally, recurring contact by the same organization[[19]](#footnote-19) and encouragement to talk about the election with friends and family attempts to keep the same contextual cues for the Attention Stage in future elections, so that the Appraisal stage can more easily be sidestepped. In summary, one key goal of this program was to affect long-term voting behavior of individuals; this stands in contrast to most field experiments in the voting-as-a-habit literature, which were designed to maximize short-term (i.e., upstream) voting behavior.

The sample is balanced across treatment and control conditions in terms of household income, age, marriage status, sex, race, political party, and whether they voted in the 2012, 2010, and 2008 general elections. When treatment condition is regressed against these covariates for each experiment, the joint p-value does not exceed 0.05, with the exception of Washington, which is already excluded from all primary analyses because the upstream effect is null (see SM, Section 5.2). Some precincts were not canvassed due to programmatic and/or logistical reasons after randomization but before treatment implementation. Since treatment and control units were balanced within precinct, we can exclude these units from the analysis without biasing our results.

As seen in Tables 3.1-3.2, the demographics of the sample in each experiment differed substantially. These differences are partially an artifact of geography. (e.g., individuals randomized in the Philadelphia experiment are predominantly black, because Philadelphia is a black majority city). But a major part of these differences is due to strategic considerations: namely, the organization is tasked with engaging with working class voters and so their targets will not be representative of the wider electorate. Furthermore, the organization’s sampling criteria reflected their expectations as to people who are likely to be receptive to their GOTV outreach (which explains why such a high proportion of individuals across all experiments voted in the 2012 general election). To account for the influence of demographics on turnout persistence in my analyses, I control for all available observable covariates.[[20]](#footnote-20) I also note that there may be intrinsic differences between populations in the subsequent subgroup analyses. For example, if all targets in low salience elections happen to be people who never take cognitive shortcuts, while all targets in the high salience elections happen to always take cognitive shortcuts, we would observe turnout persistence in high salience elections not because of the salience of the election but because of certain unobservable, intrinsic characteristics of the populations being studied. However, there is no concrete evidence to indicate that this may be the case.

**Table 3.1**: Summary statistics (demographics)

|  |  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
|  |  | **Which experiment?** | | | | | | | | | |
|  |  | **Whole Sample** | **IA 2014** | **IL 2014** | **MI 2014** | **NY 2014** | **WA 2014** | **PA 2015** | **Philly 2015** | **NC 2016** | **VA 2017** |
| ***Age*** |  |  |  |  |  |  |  |  |  |  |  |
|  | ***18-29*** | 6% | 10% | 3% | 3% | 8% | 6% | 7% | 1% | 27% | 8% |
|  | ***30-49*** | 28% | 35% | 21% | 32% | 26% | 27% | 34% | 29% | 40% | 25% |
|  | ***50-64*** | 36% | 30% | 37% | 43% | 34% | 33% | 34% | 35% | 21% | 29% |
|  | ***65+*** | 30% | 25% | 38% | 22% | 32% | 33% | 25% | 35% | 12% | 38% |
| ***Gender*** |  |  |  |  |  |  |  |  |  |  |  |
|  | ***% Female*** | 54% | 60% | 53% | 48% | 54% | 53% | 60% | 60% | 56% | 59% |
| ***Race*** |  |  |  |  |  |  |  |  |  |  |  |
|  | ***White*** | 72% | 0% | 83% | 83% | 91% | 86% | 41% | 36% | 34% | 81% |
|  | ***Black*** | 17% | 0% | 8% | 11% | 2% | 3% | 52% | 58% | 51% | 18% |
|  | ***Hispanic*** | 4% | 0% | 6% | 1% | 5% | 3% | 5% | 5% | 6% | 0% |
| ***Partisanship*** |  |  |  |  |  |  |  |  |  |  |  |
|  | ***Democrat*** | 24% | 56% | 0% | 0% | 30% | 0% | 98% | 98% | 67% | 0% |
|  | ***Republican*** | 6% | 5% | 0% | 0% | 39% | 0% | 2% | 2% | 5% | 0% |
|  | ***Party Unknown*** | 66% | 0% | 100% | 100% | 23% | 100% | 0% | 0% | 0% | 100% |
| ***Median Household Income*** |  | $50,000 - $75,000 | $50,000 - $75,000 | $50,000 - $75,000 | $50,000 - $75,000 | $50,000 - $75,000 | $50,000 - $75,000 | $30,000 - $50,000 | $30,000 - $50,000 | $30,000 - $50,000 | $50,000 - $75,000 |
| ***% Married*** |  | 47% | 51% | 66% | 66% | 66% | 0% | 0% | 0% | 28% | 0% |
| ***Total N*** |  | 1,875,739 | 57,991 | 535,726 | 449,733 | 252,258 | 119,769 | 49,278 | 205,981 | 137,032 | 67,971 |

**Table 3.2**: Summary statistics (vote history)

|  |  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
|  |  | **Which experiment?** | | | | | | | | | |
|  |  | **Whole Sample** | **IA 2014** | **IL 2014** | **MI 2014** | **NY 2014** | **WA 2014** | **PA 2015** | **Philly 2015** | **NC 2016** | **VA 2017** |
| ***% Voted in 2012 General*** |  | 90% | 99% | 93% | 98% | 83% | 92% | 92% | 90% | 60% | 91% |
| ***% Voted in 2010 General*** |  | 73% | 76% | 87% | 79% | 56% | 73% | 67% | 74% | 28% | 62% |
| ***% Voted in 2008 General*** |  | 85% | 85% | 89% | 92% | 81% | 80% | 85% | 91% | 52% | 84% |
| ***Total N*** |  | 1,875,739 | 57,991 | 535,726 | 449,733 | 252,258 | 119,769 | 49,278 | 205,981 | 137,032 | 67,971 |

1. ***Justifying the Lack of Pre-Registration***

The empirical data was analyzed by me over the course of several years as part of a consulting project. While I did not analyze the specific subgroups discussed in the paper, there is no way for me to verify that this was the case. Before undertaking my test of PMBC on this empirical data, I carefully considered submitting a pre-analysis plan (PAP). The benefits of PAPs are widely documented (Olken 2015; Nosek et. al 2018; Ofosu & Posner 2021), so my final decision to not pre-register deserves justification.

First, the analysis was conducted on proprietary historical data, so the situation is similar to the argument against pre-registering observational analyses of historical data (Ofosu & Posner 2021). There is no concrete evidence that would definitively show that subgroup analysis was performed only after the hypotheses were pre-registered. It would simply be a restatement of my assurance that I did not analyze the subgroups that test the PMBC, but with the additionally risk of providing a casual reader with the erroneous impression that the findings are more rigorous. Indeed, some scholars bring up that PAPs may inadvertently act as a signal of rigor without any actual accountability (e.g., Laitin 2013). This criticism has turned out to be partially true in practice; Ofosu & Posner’s (2021) empirical research on PAPs has found that only 34% of reviewers consult PAPs. It is likely that the average reader is even less likely to pay attention to the specifics of a PAP (including the fact that the data has been analyzed previously), but may still consciously or subconsciously ascribe a higher level rigor to pre-registered studies.

Second, my goal is to emphasize that though this study uses a large-N dataset, my findings should be first replicated before the PMBC interpretation of turnout persistence is accepted (cf., Coffman & Niederle, 2015). As noted in Green & Gerber (2019), campaigns and non-profits do adopt findings from political science research as “best practices.” Emphasizing the exploratory nature of these findings may dampen overenthusiasm by practitioners.

Finally, to compensate for a lack of pre-registration, I attempt to maximize analytical and data transparency. I attempt to illustrate all the possible ways the subgroup analysis can be conducted with the available data and provide robustness checks where possible. I also relegate an exploratory, less intuitive prediction to the Supplementary Materials as a precaution, despite it also comporting with PMBC. Finally, all data will be made available in a de-identified format for public use upon publication.

1. ***Analysis***

*5.1: Upstream Effects*

As a first step, I analyze all the field experiments together in an OLS regression (with experiment fixed effects and robust standard errors) to measure the impact of the treatment on upstream elections. Overall, the organization’s GOTV outreach increased turnout by 0.3 percentage points (Intent-to-Treat, se = 0.0006, p<0.001). It is standard practice (e.g., Gerber and Green 2012) to take into account that not all households were successfully contacted and evaluate the Complier Average Causal Effect (CACE), where the complier is the individual who was successfully contacted by a canvasser.[[21]](#footnote-21) I find that, at the door, the organization’s GOTV canvass had a 1.7 percentage point impact on voter turnout (CACE, se = 0.0034, p<0.001); CACEs by experiment are also presented in Table 5.1. A meta-analysis by Gerber & Green (2019) finds that among the studies that had a control group turnout rate between 50% and 70% (the control group in my data had a turnout rate of 69.6%), the CACE conditional on opening the door was 1.8 with a 95% confidence interval of [0.4, 3.3] (Green & Gerber 2019, p. 211). We can, therefore, safely conclude that the upstream impact of the doorknock treatment in this paper is typical of door-to-door mobilization efforts.

*5.2: Checking for Weak Instruments*

One threat to the validity of CACE estimates is weak instruments. This is not be confused with the problem of causal identification: e.g., if a canvass was not actually conducted, the canvass experiment would fail to turn out voters to the polls and turnout persistence among those who were successfully turned out is not estimatable.[[22]](#footnote-22) Weak instruments, however, are a statistical problem—if the first stage coefficient (in this case, the proportion of voters successfully turned out by the treatment) is close to zero, our CACE estimates may be distorted (Andrews, Stock, & Sun 2019).

To check for weak instruments, I examine the impact of each treatment on upstream elections by experiment (see Table 5.2). Though some experiments were underpowered on their own, most of the outreach programs had upstream treatment effects that were well within range of one another. The only exception is the 2014 canvass program in Washington, which appears to have had a near-perfect zero impact on turnout in the 2014 general election (p=0.95). The extremely small magnitude of the upstream coefficient is likely to blow up our CACE estimates in this experiment.[[23]](#footnote-23) More formally, a weak instrument test yields a Kleibergen & Paap (2006) F-statistic of 0.003, well below the rule-of-thumb threshold of 10 (Keele & Morgan 2016) and even the most lenient Stock-Yogo (Stock & Yogo, 2005) critical values (i.e., 25% maximal IV size: 5.53). As such, in all subsequent analyses, the Washington data is omitted. Including the Washington data in all subsequent analyses does not substantially change the reported results, since it accounted for just ~6% of the entire sample.[[24]](#footnote-24)

**Table 5.2**: Upstream effects by each experiment

|  |  |  |
| --- | --- | --- |
| ***Experiment*** | ***Intent-to-treat Effect*** | ***CACE (conditional on opening the door)*** |
| *2014 IA* | 0.0140 | 0.0666 |
|  | (0.0034) | (0.0164) |
| *2014 IL* | 0.0016 | 0.0101 |
|  | (0.0010) | (0.0063) |
| *2014 MI* | 0.0022 | 0.0112 |
|  | (0.0012) | (0.0061) |
| *2014 NY* | 0.0019 | 0.0139 |
|  | (0.0018) | (0.0136) |
| *2014 WA* | 0.0002 | 0.0006 |
|  | (0.0030) | (0.0104) |
| *2015 PA SUPREME COURT* | 0.0116 | 0.0652 |
|  | (0.0045) | (0.0254) |
| *2015 PHILLY MAYORAL* | 0.0078 | 0.0249 |
|  | (0.0023) | (0.0074) |
| *2016 NC* | 0.0024 | 0.0532 |
|  | (0.0036) | (0.0817) |
| *2017 VA* | 0.0074 | 0.0497 |
|  | (0.0053) | (0.0354) |

For all analyses reported in the main manuscript, I conduct weak instrument tests and when these tests indicate that the instrument is weak, as an alternative specification to 2SLS regression, I include the weak-instrument-robust test formalized by Anderson & Rubin (1949) and recommended in e.g., Staiger & Stock (1997); Andrews, Stock, & Sun (2019). I present the results here:

*Aggregate Results:* The Kleibergen & Paap (2006) F-statistic=26.0 is well above the rule-of-thumb threshold of 10 (Keele & Morgan 2016) and over the 9.0 Stock-Yogo 15% maximal IV size critical value (Stock & Yogo, 2005). This suggests that the instrument is not weak.

*Prediction 1:* The Kleibergen-Paap F-statistic is 11.9 and above the rule-of-thumb threshold of 10 and over the 9.0 Stock-Yogo 15% maximal IV size critical value. This suggests that the instrument is not weak.

*Prediction 2:* The Kleibergen-Paap F-statistic for the low-salience subgroup is 16.9 and 14.2 for the high-salience subgroup. Both are above the rule-of-thumb threshold of 10 and over the 9.0 Stock-Yogo 15% maximal IV size critical value. This suggests that both instruments are not weak.

*Prediction 3:* The Kleibergen-Paap F-statistic for the Dem Loss subgroup is 11.1, above the rule-of-thumb threshold of 10 and over the 9.0 Stock-Yogo 15% maximal IV size critical value, indicating that this instrument is not weak. However, the Dem Win subgroup may suffer from a weak instrument with a Kleibergen-Paap F-statistic of 6.0. As reported in the main text, I use Anderson-Rubin weak-instrument-robust tests, which do not yield CACEs or comparable coefficients but can identify if the downstream effect is statistically significant. These tests find evidence of turnout persistence in the Dem Win subsample (p =0.001) and no turnout persistence in the Dem Loss subsample (p-value=0.503). This is consistent with the reported CACEs.

*5.3: Additional Predictions*

***Prediction 4:*** *Household income may be associated with turnout persistence formation.* This final prediction is dependent on a number of assumptions that have mixed support in the existing literature.Fujiwara, Meng and Vogl (2016) noted that the cost of voting is relatively higher for individuals with lower household incomes. On the other hand, the opportunity cost of voting for higher income individuals may be higher and thus discourage some higher income individuals from voting. A meta-analysis of 40 published studies found that income was positively correlated with likelihood to turn out, but the relationship failed to be significant in nearly half of the studies included in the meta-analysis (Smets & van Ham 2013). As such, this prediction rests on precarious theoretical ground and the empirical results from this prediction should be viewed with a healthy dose of skepticism.

That said, the proposed mechanism is as follows. There is a documented association between lower levels of income and higher levels of geographic mobility (e.g., Purcell, 2020), which would imply that contextual cues in different geographies are likely to be different and hence detrimental to the formation of heuristic shortcuts. This is consistent with Coppock & Green’s (2016) conclusions that movers were less likely to exhibit turnout persistence. And so, the PMBC framework implies that lower income individuals would be less likely to exhibit turnout persistence.

When I look at the initial upstream effect among participants with lower than the mean income level in our sample ($50,000 - $75,000),[[25]](#footnote-25) the upstream treatment effect is actually higher than the upstream treatment effect for participants who have a higher than mean level of income (0.43 percentage points [p<0.001] vs. 0.20 percentage points [p=.04]). In other words, the GOTV outreach in this study appears to be more effective for lower income individuals in upstream elections. But the downstream CACE for higher income households is twice as large as for lower income households (1.1 [p=.08] vs. 0.45 [p=0.03]). In short, these findings offer some support for the PMBC prediction that turnout persistence is less likely to occur for lower income voters.

*5.4: Robustness Checks*

***Prediction 2:*** *Voters induced to vote in a low-salience election will not exhibit stronger turnout persistence effects.* One could construe the NY 2014 general election as low salience. Even though there was a gubernatorial election on the ballot, Andrew Cuomo was predicted to win handedly (*RCP* 2014) and turnout was only seven percentage points higher than the Philadelphia Mayoral race. When I include NY as one of the three low salience elections (in the second row of Table 5.3), there is, again, no persistence in the 2018 general election among these low salience experiments (the third row cell in light gray), despite an upstream effect that’s nearly twice as large as the upstream effect of all other experiments (0.50 percentage points versus 0.28 percentage points). The other experiments, in contrast, again have a large and statistically significant CACE (0.95, p=0.003) in the 2018 general election. In the remaining two rows, I try two more combinations of low salience races and, in both cases, find null downstream CACEs (the final two cells in light gray).

**Table 5.4**: Turnout persistence is more likely when the GOTV treatment was in a high salience election than in a low salience election

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
|  |  |  |  | 2018 Mid-term General Election |
| Lowest Projected Turnout + No Federal Race on Ballot | Low Salience | (PA, Philly) | 0.0085\*\*\* | -0.0736 |
|  |  | (0.0021) | (0.2241) |
| High Salience | (All others) | 0.0026\*\*\* | 0.8829\*\* |
|  |  | (0.0007) | (0.3068 ) |
| Lowest Projected Turnout | Low Salience | (PA, Philly, NY) | 0.0050\*\*\* | 0.0267 |
|  |  | (0.0014) | (0.2643) |
| High Salience | (All others) | 0.0028\*\*\* | 0.9463\*\* |
|  |  | (0.0007) | (0.3195) |
| No Federal Seats on Ballot | Low Salience | (PA, Philly, VA) | 0.00835\*\*\* | -0.0087 |
|  |  | (0.0019) | (0.2094) |
| High Salience | (All others) | 0.0025\*\*\* | 0.9038\*\* |
|  |  | (0.0007) | (0.3279) |
| No President/Governor/Senate Race on Ballot | Low Salience | (PA, Philly, VA, NY) | 0.0053\*\*\* | 0.0792 |
|  |  | (0.0013) | (0.2411) |
| High Salience | (All others) | 0.0027\*\*\* | 0.9717\*\*\* |
|  |  | (0.0008) | (0.3426) |

*This table examines the impact of voting in the elections described in each row on downstream turnout in the 2018 mid-term election. Robust standard errors in parentheses. \*\*\* = p-value≤0.001, \*\* = p≤0.01*

1. ***Additional Discussion***

Another reasonable concern about the analysis of Prediction 1 is ceiling effects. The maximum possible effect size among compliers is much lower in high turnout downstream elections than in low turnout downstream elections. However, the turnout rates in the control group for 2016 general and the 2018 general elections in my sample are not dramatically different (74% in the 2018 midterm versus 86% in the 2016 general) and the CACE for the 2018 midterm is over four times larger than the CACE for the 2016 general. I conclude the PMBC interpretation is more likely than the ceiling effect interpretation.

It is also worth noting that even though the experiments were conducted in different election contexts in different states, the types of voters included in these samples are nevertheless *not* nationally representative and reflect the labor organization’s strategic targeting criteria (i.e., working-class, Democratic-leaning voters). As such, it is unclear whether these results generalize to a broader population. Additionally, though I assert that the treatment did not include any social pressure interventions (which might credibly change the cost-benefit equation in the short-term), the conversations were freeform and we have no way of knowing if there were large numbers of canvassers who went off-script. This possible limitation is, however, unlikely, as the labor organization employs paid canvassers who receive extensive training.

1. ***Additional Tables and Figures***

**Figure 7.1**: Average treatment effects for upstream and downstream elections by experiment



**Table 7.2**: Upstream intent-to-treat effect and CACEs for downstream elections by experiment

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| ***Experiment*** | ***1st Stage (Upstream Election)*** | ***2016 Primary CACE*** | ***2016 General CACE*** | ***2018 General CACE*** |
| *2014 IA* | 0.0140 | -0.1727 | 0.0701 | 0.2225 |
|  | (0.0034) | (0.2136) | (0.1963) | (0.2333) |
| *2014 IL* | 0.0016 | 1.2897 | 0.6913 | 1.7542 |
|  | (0.0010) | (0.9389) | (0.5664) | (1.1250) |
| *2014 MI* | 0.0022 | 0.4255 | -0.2773 | 1.0485 |
|  | (0.0012) | (0.5964) | (0.4953) | (0.6810) |
| *2014 NY* | 0.0019 | 0.1986 | 0.6927 | 0.4395 |
|  | (0.0018) | (0.9400) | (0.9494) | (0.9907) |
| *2014 WA[[26]](#footnote-26)* | 0.0002 | -11.0919 | 15.2327 | 18.5140 |
|  | (0.0030) | (199.5858) | (259.6603) | (315.9351) |
| *2015 PA SUPREME COURT* | 0.0116 | -0.3215 | 0.1649 | -0.2050 |
|  | (0.0045) | (0.3925) | (0.1970) | (0.3364) |
| *2015 PHILLY MAYORAL* | 0.0078 | 0.2097 | 0.2504 | 0.0239 |
|  | (0.0023) | (0.2676) | (0.2159) | (0.2739) |
| *2016 NC* | 0.0024 | NA | NA | -0.2869 |
|  | (0.0036) | NA | NA | (1.8542) |
| *2017 VA* | 0.0074 | NA | NA | 0.6973 |
|  | (0.0053) | NA | NA | (0.6410) |

**Table 7.3**: Prediction 1: Covariate-adjusted turnout in control and treatment across upstream and downstream elections

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
|  | 2014 General Election | 2016 Presidential Primary Election | 2016 Presidential General Election | 2018 Mid-term General Election |
| Control | 70.15% | 54.43% | 86.14% | 74.22% |
| Treatment | 70.40% | 54.53% | 86.20% | 74.45% |

**Table 7.4**: Prediction 2: Covariate-adjusted turnout in control and treatment across upstream and downstream elections

|  |  |  |  |
| --- | --- | --- | --- |
|  |  | Upstream Election | 2018 Mid-term General Election |
| Low salience | Control | 57.03% | 71.75% |
| Treatment | 57.88% | 71.68% |
| High salience | Control | 70.15% | 72.00% |
| Treatment | 70.41% | 72.23% |

**Table 7.5**: Prediction 3: Covariate-adjusted turnout in control and treatment across upstream and downstream elections

|  |  |  |  |
| --- | --- | --- | --- |
|  |  | Upstream Election | 2018 Mid-term General Election |
| Dem Win | Control | 75.53% | 76.72% |
| Treatment | 75.72% | 76.98% |
| Dem Loss | Control | 68.49% | 55.83% |
| Treatment | 69.37% | 56.01% |

1. ***Anonymized Canvasser Script and Training Documents***

November 4, 2014

Governor: Mark Schauer;

US Senate: Gary Peters

[REDACTED] Persuasion Rap

Introduction

Hi, my name is \_\_\_ with [REDACTED]. Are you [name]? Great! We’re out today talking with folks in [insert community] about the election for Governor and the Michigan House.

QUESTIONS

Question 1

Are you planning to vote in the election?

[*Do not record]*

Question 2 (Voter ID Governor)

Thank you. If you were going to vote today in the election for Governor would you support for Republican Rick Snyder or Democrat Mark Schauer?

[*Record Response*: Snyder-R, Schauer-D, Unsure/Undecided]*]*

MARK SCHAUER

PERSUASION AND ENDORSEMENTS

[REDACTED] is an independent organization that represents 120,000 Michiganders who want an economy that works for working people. We are not part of any political party or campaign and support candidates based on their record.

[IF SCHAUER]

We are also supporting Mark Schauer to be the next Governor. Thanks for your support!

*Hand over lit. Go to Peters* *Endorsement*

[IF UNDECIDED]

[REDACTED] has researched the issues and we have found that Mark Schauer has the strongest record of supporting working families and the issues important to the community. That is why he has the support of tens of thousands of working men and women in Michigan and will work to create good jobs and improve our schools.

So can we count on your vote for Mark Schauer for Governor?

*Do not record response. Hand over lit. Go to Peters* *Endorsement.*

[IF OTHER CANDIDATE]

I understand. How you vote is a personal decision. [REDACTED] has done the research on the issues and we believe that Mark Schauer has the strongest track record of getting things done for working people and will be the strongest leader for Michiganders .

*Hand over lit. Go to Peters* *Endorsement.*

US SENATE:

GARY PETERS

ENDORSEMENT

[REDACTED] has done the research and found that Gary Peters has the strongest record of fighting for Michigan. In Congress he helped deliver on the loans to save the auto industry jobs and start to turn around Michigan’s economy.

*Go to Plan Making.*

PLAN MAKING

[IF SCHAUER or UNDECIDED FOR STATE HOUSE]

Great! Thanks for your support. As you probably know, this election will determine the direction Michigan takes and turnout is going to be high.

The polls open at 7 am and close at 8 pm on Election Day. When do you normally go- on your way to work? When you get off? Just before dinner? [LISTEN FOR RESPONSE & FOLLOW UP IF APPROPRIATE. DRAW RESPONSE OUT OF VOTER.]

[FOR VOTERS WHO COMMIT TO TURNOUT]

Thanks. You know one of the best ways to increase Mark Schauer’s chances of winning is to talk to your family about why this election matters to you. Have you had a chance to talk to your family about this election? [LISTEN FOR RESPONSE & FOLLOW UP IF APPROPRIATE]

*Sample follow up question: “How do they feel about the Governor’s race? Did you share your reasons for voting for Mark Schauer?”*

Thanks, Goodnight.

2014 Michigan General Election For Governor

Mark Schauer

PERSUASION TALKING POINTS

Crumbling Infrastructure: Fixing our roads and bridges is not Governor Snyder’s priority. The delay means an added $1.8 billion in costs for taxpayers, nearly $400 per year in unnecessary auto repairs for the average driver and countless hours of frustration. Experts say that one in four bridges rated as structurally deficient.

Mark Schauer will finally fix our roads and bridges by getting corporations to pay their fair share in taxes for the infrastructure they rely upon to conduct business.

Jobs and the Economy: Rick Snyder promised to focus on job creation, but in the last four years, Michigan has trailed the nation and neighboring states like Ohio in adding new jobs. Instead of creating jobs, Snyder has pursued divisive policies like “right to work” that hurt middle class Michiganders.

Mark Schauer has a record of defending Michigan jobs and workers. In Congress he supported the $80 billion investment to stabilize the auto industry. Since then, the money has been repaid and the Big Three are profitable. Schauer also wants to repeal the divisive Right to Work law, arguing that “Snyder cannot point to one job that’s been created as a result of right to work.”

Failing our Schools: In his first year in office, Governor Snyder cut over $1 billion in state school funding- more than $400 per pupil. As a result, class sizes have increased, learning environments have suffered and more than 50 school districts have experienced a credit down grade.

Mark Schauer will stop the diversion of Michigan School Aid fund dollars away from the classroom- nearly $400 per pupil that has resulted in increased class sizes and poorer learning environments, and threatens our finances.

Training Memo

Please distribute only to trainers

Opening the Rap: Canvassers have found it works better to use the words “upcoming election” with voters versus “out talking about issues”. They found that voters perceive “out talking about issues” as a longer conversation compared to “upcoming election”. Again work with folks on what approach works best for them – if there is push back on “upcoming election” it may be better to approach the voter with “out talking about issues in our community”

Inclusive Language: Reminding the voter that we are part of the same community and share the same interests facilitates a more collaborative conversation. For example, “That is why so many folks here in our community have decided that they are voting for Mark Schauer” assures the voter that we are looking at the issues from the same perspective. This approach assumes support. By contrast, exclusive language like “you” and “they” imply social distance and that you are attempting to convince the person to share the perspective.

Part of Something Bigger and Enthusiasm: It’s common for voters to want to feel like they are part of a larger effort to achieve a shared goal. Using language that conveys the point helps frame their decision to vote as part of a larger effort and makes them more likely to turnout. For example, “We have been talking to a lot of people who tell us that they are just tired of all of the negative attack ads that don’t do anything to fix the economy and other problems we are all facing. I am sure you probably feel the same way, right,” or “This is going to be an important election, and we are expecting high turnout” are effective ways to convey that just like in school, all of the popular students are participating. It is for the same reason that it is important to show the voter that we are enthusiastic about supporting candidates who support our issues.

Clean IDs/Recording Responses: Capturing the initial response to issue and candidate ID questions is an important part of building the longer term political program. For example, know where we had the highest initial rate of support for candidate verses undecided voters informs where we prioritize future passes. While recording an ID after we have given some indication of our preferred candidate has value, the goal is to find those voters who are most likely to be with us.

Plan Making: An important part of plan making is that we walk voters through thinking about when and where they are going to go vote. Using time prompts and inclusive language, and effective plan making conversation helps the voter think about voting in the context of their own lives. For example, “It will be important to have plan on when you go to the polls. I like to go in the morning before work, but a lot of people like to go in the evening before dinner. What about you- when do you plan to head to the polls?” involves emphasizing everyday events that people plan their days around.

Talk To Your Family: A new part of the canvass is asking supportive voters to discuss the election with other voting members of their family. Research studies performed by the [REDACTED] have shown that asking voters talk to family members about voting for Mark Schauer can increase overall vote probability for Schauer. The study found that getting supporters to talk to their family members about the election was over THREE TIMES MORE EFFECTIVE than existing practices of getting uncommitted household members to support a candidate.

1. ***Additional References***

Achen, C., & Bartels, L. M. (2017). *Democracy for realists*. Princeton University Press.

Anderson, T. W., & Rubin, H. (1949). Estimation of the parameters of a single equation in a complete system of stochastic equations. *The Annals of mathematical statistics*, *20*(1), 46-63.

Andrews, I., Stock, J. H., & Sun, L. (2019). Weak instruments in instrumental variables regression: Theory and practice. *Annual Review of Economics*, *11*, 727-753.

Blais, A. (2000). *To vote or not to vote?: The merits and limits of rational choice theory*. University of Pittsburgh Press.

Coffman, L. C., & Niederle, M. (2015). Pre-analysis plans have limited upside, especially where replications are feasible. *Journal of Economic Perspectives*, *29*(3), 81-98.

Coppock, A., & Green, D. P. (2016). Is voting habit forming? New evidence from experiments and regression discontinuities. *American Journal of Political Science*, *60*(4), 1044-1062.

Davenport, T. C., Gerber, A. S., Green, D. P., Larimer, C. W., Mann, C. B., & Panagopoulos, C. (2010). The enduring effects of social pressure: Tracking campaign experiments over a series of elections. *Political Behavior*, *32*(3), 423-430.

DellaVigna, S., List, J. A., Malmendier, U., & Rao, G. (2016). Voting to tell others. *The Review of Economic Studies*, *84*(1), 143-181.

Dinas, E. (2012). The formation of voting habits. *Journal of Elections, Public Opinion & Parties*, *22*(4), 431-456.

Downs, A. (1957). *An economic theory of democracy.* Harper and Row Publishers.

Duckworth, A. L., & Gross, J. J. (2020). Behavior change. *Organizational Behavior and Human Decision Processes*, *161*, 39-49.

Etang, A., Fielding, D., & Knowles, S. Looking for Framing Effects in Expressive Voting Experiments.

Franklin, M. N., & Hobolt, S. B. (2011). The legacy of lethargy: How elections to the European Parliament depress turnout. *Electoral Studies*, *30*(1), 67-76.

Frey, E., & Rogers, T. (2014). Persistence: How treatment effects persist after interventions stop. *Policy Insights from the Behavioral and Brain Sciences*, *1*(1), 172-179.

Fujiwara, T., Meng, K., & Vogl, T. (2016). Habit formation in voting: Evidence from rainy elections. *American Economic Journal: Applied Economics*, *8*(4), 160-88.

Garcia Bedolla, L, & Michelson, M. R. (2012). *Mobilizing Inclusion*. Yale University Press.

Gelman, A. (2007). Struggles with survey weighting and regression modeling. *Statistical Science*, *22*(2), 153-164.

Gerber, AS, and DP Green (2012): *Field Experiments: Design, Analysis, and Interpretation*, WW Norton & Co.

Gerber, A. S., Green, D. P., & Larimer, C. W. (2008). Social pressure and voter turnout: Evidence from a large-scale field experiment. *American political Science review*, *102*(1), 33-48.

Gerber, A. S., Green, D. P., & Shachar, R. (2003). Voting may be habit‐forming: evidence from a randomized field experiment. *American journal of political science*, *47*(3), 540-550.

Gerber, A. S., Huber, G. A., Doherty, D., & Dowling, C. M. (2016). Why people vote: estimating the social returns to voting. *British Journal of Political Science*, *46*(2), 241-264.

Gerber, A. S., Huber, G. A., Doherty, D., & Dowling, C. M. (2016). Why people vote: estimating the social returns to voting. *British Journal of Political Science*, *46*(2), 241-264.

Green, D. P., & Gerber, A. S. (2019). *Get out the vote: How to increase voter turnout*. Brookings Institution Press.

Green, D. P., McGrath, M. C., & Aronow, P. M. (2013). Field experiments and the study of voter turnout. *Journal of Elections, Public Opinion and Parties*, *23*(1), 27-48.

Green, D. P., & Shapiro, I. (1996). Pathologies revisited: Reflections on our critics. *The rational choice controversy*, 235-276.

Gollwitzer, P. M., & Sheeran, P. (2006). Implementation intentions and goal achievement: A meta‐analysis of effects and processes. *Advances in experimental social psychology*, *38*, 69-119.

Hill, S. J., & Kousser, T. (2016). Turning out unlikely voters? A field experiment in the top-two primary. *Political Behavior*, *38*(2), 413-432.

Holbein, J., A Rangel, M., Moore, R., & Croft, M. (2021). Are Voting Treatments Transformative? Expanding Upon and Meta-Analyzing the Evidence. Expanding Upon and Meta-Analyzing the Evidence.

Jakee, K., & Sun, G. Z. (2006). Is compulsory voting more democratic?. *Public Choice*, *129*(1), 61-75.

Just, D. R., & Gabrielyan, G. (2018). Influencing the food choices of SNAP consumers: Lessons from economics, psychology and marketing. *Food Policy*, *79*, 309-317.

Keele, L., & Morgan, J. W. (2016). How strong is strong enough? Strengthening instruments through matching and weak instrument tests. *The Annals of Applied Statistics*, 1086-1106.

Kleibergen, F., & Paap, R. (2006). Generalized reduced rank tests using the singular value decomposition. *Journal of econometrics*, *133*(1), 97-126.

Laitin, D. D. (2013). Fisheries management. *Political Analysis*, *21*(1), 42-47.

Lally, P., & Gardner, B. (2013). Promoting habit formation. *Health psychology review*, *7*(sup1), S137-S158.

Lally, P., Van Jaarsveld, C. H., Potts, H. W., & Wardle, J. (2010). How are habits formed: Modelling habit formation in the real world. *European journal of social psychology*, *40*(6), 998-1009.

Meredith, M. (2009). Persistence in political participation. *Quarterly Journal of Political Science*, *4*(3), 187-209.

Michelson, M. R. (2003). Dos Palos revisited: Testing the lasting effects of voter mobilization. In Annual Meeting of the Midwest Political Science Association.

Nickerson, D. W., & Rogers, T. (2010). Do you have a voting plan? Implementation intentions, voter turnout, and organic plan making. *Psychological Science*, *21*(2), 194-199.

Nosek, B. A., Ebersole, C. R., DeHaven, A. C., & Mellor, D. T. (2018). The preregistration revolution. *Proceedings of the National Academy of Sciences*, *115*(11), 2600-2606.

Ofosu, G. K., & Posner, D. N. (2021). Pre-analysis plans: An early stocktaking. *Perspectives on Politics*, 1-17.

Olken, B. A. (2015). Promises and perils of pre-analysis plans. *Journal of Economic Perspectives*, *29*(3), 61-80.

Purcell, P. J. (2020). Geographic Mobility and Annual Earnings in the United States. *Soc. Sec. Bull.*, *80*, 1.

*RCP.* (2014). 2014 Governor Races. *Real Clear Politics*. <https://www.realclearpolitics.com/epolls/2014/governor/2014_elections_governor_map.html>

Riker, W. H., & Ordeshook, P. C. (1968). A Theory of the Calculus of Voting. *American political science review*, *62*(1), 25-42.

Rogers, T., & Frey, E. (2014). Changing behavior beyond the here and now. <https://papers.ssrn.com/sol3/papers.cfm?abstract_id=2410465>

Rogers, T., Fox, C. R., & Gerber, A. S. (2013). Rethinking why people vote. In *The behavioral foundations of public policy* (pp. 91-107). Princeton University Press.

Rogers, T., Green, D. P., Ternovski, J., & Young, C. F. (2017). Social pressure and voting: A field experiment conducted in a high-salience election. *Electoral Studies*, *46*, 87-100.

Rogers, T., Ternovski, J., & Yoeli, E. (2016). Potential follow-up increases private contributions to public goods. *Proceedings of the national academy of sciences*, *113*(19), 5218-5220.

Rolfe, M. (2012). *Voter turnout: A social theory of political participation*. Cambridge University Press.

Smets, K., & Van Ham, C. (2013). The embarrassment of riches? A meta-analysis of individual-level research on voter turnout. *Electoral studies*, *32*(2), 344-359.

Solvak, M., & Vassil, K. (2018). Could Internet Voting Halt Declining Electoral Turnout? New Evidence That E‐Voting Is Habit Forming. *Policy & Internet*, *10*(1), 4-21.

Staiger, D., & Stock, J. H. (1997). Instrumental Variables Regression with Weak Instruments. *Econometrica*, *65*(3), 557-586.

Stock, J., & Yogo, M. (2005). Asymptotic distributions of instrumental variables statistics with many instruments. *Identification and inference for econometric models: Essays in honor of Thomas Rothenberg*, *6*, 109-120.

1. This variable is also referred to as the “civic duty” benefit of voting (e.g., Gerber et al., 2016), but since this benefit may imply some extrinsic social dimension (i.e., “duty” implies some obligation to others, which may be enforced through social sanctions), I opt to use “warm glow.” [↑](#footnote-ref-1)
2. This intervention exploited the fact that whether or not someone has voted is public record in the United States. [↑](#footnote-ref-2)
3. “90% of respondents on the 1984 GSS claimed voting was a very important civic obligation" (Rolfe, 2012, p. 50) [↑](#footnote-ref-3)
4. Though for conceptual clarity in later discussions, I decomposed their “social image” functional into social cost and social benefit. [↑](#footnote-ref-4)
5. In fact, the exact cost-benefit equation can vary from person to person. In the theoretical framework I propose in this manuscript, the cost-benefit equation is simply one stage in an overarching process of recurring behavior. The cost-benefit equation for any given individual outputs an expected utility; the other stages in my theoretical framework use only the resulting expected utility and do not refer to the specific costs and benefits that produced that expected utility. [↑](#footnote-ref-5)
6. Note that the theory of expressive voting (e.g., Rogers, Fox, and Gerber, 2013) can be incorporated in the “warm glow” variable and the social cost/benefits of voting (e.g., Etang, Fielding, & Knowles, 2016). [↑](#footnote-ref-6)
7. For a more thorough criticism of construing persistent voting as “habitual,” see Dinas (2012; 435-436) [↑](#footnote-ref-7)
8. Studies that use such formal models of voting tend to draw on other theories and empirical research for concrete predictions (e.g., Fujiwara, Meng & Vogl, 2016). [↑](#footnote-ref-8)
9. Heuristic behaviors need not be habitual in the strict psychological sense. An example of heuristic behavior that is not habitual would be entering a new cafeteria and opting for a new healthy food option, because it is most prominently displayed (e.g., Just & Gabrielyan, 2018) [↑](#footnote-ref-9)
10. For a more detailed discussion of how social interactions after voting can exacerbate turnout persistence, see Fujiwara, Meng, and Vogl (2016; 182-3). [↑](#footnote-ref-10)
11. It’s worth noting that there are some scholars who asserted that the warm glow variable from voting is so large that it doesn’t lead to turnout persistence, but is also “transformative” in changing individuals’ overall political engagement (e.g., Jakee and Sun, 2006). A large meta-analysis has found little evidence that voting is transformative (Holbein et al., 2021). [↑](#footnote-ref-11)
12. Of course, a very plausible alternative explanation is that this difference in voting persistence may simply be an artifact of two systematically different types of voters. [↑](#footnote-ref-12)
13. See Achen and Bartels (2017) for prominent examples of voters making misattribution errors. [↑](#footnote-ref-13)
14. Or reducing turnout, as with rain on an Election Day. [↑](#footnote-ref-14)
15. A reduction of the transaction cost of voting could be construed as the informational benefit of voting, but this endogeneity cannot be completely eliminated in any study of in-person voting (for a more detailed discussion of this mechanism see Meredith, 2009). It is also possible that talking to someone in-person about an upcoming election may change the social cost or social benefit of voting, but that effect is likely to be smaller than that of social pressure nudges, because the canvassers’ script did not reference whether or not someone votes is public record—nor did it instruct canvassers to ask about participants’ prior turnout. [↑](#footnote-ref-15)
16. Although they emphasize door-to-door contact, they also employ direct mail, phone and digital outreach. [↑](#footnote-ref-16)
17. Additionally, door-to-door canvassing has proven to have the largest persistence effects in prior research (Garcia Bedolla and Michelson, 2012; 188). [↑](#footnote-ref-17)
18. The impacts on the cost-benefit equation in the Appraisal Stage may still occur. It is possible that the plan-making component may lower the transactional cost of voting by providing implicit information about how much time voting takes, but it should be noted that the vast majority of the individuals in our sample have voted before and already have some information about the voting process. It is therefore less likely that their expected transactional costs of voting will change meaningfully as a result of the conversation. (This would not be the case if the sample was predominantly new voters.) [↑](#footnote-ref-18)
19. The recurring contact is made without knowledge of the treatment assignment in any given election. [↑](#footnote-ref-19)
20. I do not use sampling weights as weighting to a particular population will inevitably reflect arbitrary choices by the researcher (for a more thorough discussion of sampling weights, see Gelman (2007)). [↑](#footnote-ref-20)
21. This does not necessarily mean that the entire script was delivered. “Successfully contacted” denotes that the individual on the canvass list opened the door and affirmatively identified themselves. [↑](#footnote-ref-21)
22. The partner organization uses paid canvassers and had strict oversight measures to ensure the canvass was actually conducted. [↑](#footnote-ref-22)
23. I report all downstream CACEs for all experiments in Table 7.2 and find that this is indeed the case. [↑](#footnote-ref-23)
24. The aggregate CACE’s instrument is not weak with and without the Washington experiment (with Washington: Kleibergen & Paap (2006) F-statistic=25.3 and without Washington, F-statistic=26.0, over the 9.0 Stock-Yogo 15% maximal IV size critical value (Stock & Yogo, 2005)). [↑](#footnote-ref-24)
25. The income variable is modelled by the data vendor, Catalist, by appending commercial consumer data to the voter file. [↑](#footnote-ref-25)
26. Washington was omitted from the main analysis since it is very likely that this analysis suffers from too weak an instrument. The first stage impact on upstream turnout was close to a perfect zero with a p-value 0.95. [↑](#footnote-ref-26)