Review of Economic Studies (2017) **84**, 143–181 doi:10.1093/restud/rdw056

Advance access publication 18 October 2016

Voting to Tell Others

STEFANO DELLAVIGNA
UC Berkeley and NBER
JOHN A. LIST
U Chicago and NBER
ULRIKE MALMENDIER
UC Berkeley and NBER
and
GAUTAM RAO
Harvard University and NBER

First version received May 2015*; final version accepted July* 2016 (*Eds.*)

Why do people vote? We design a field experiment to estimate a model of voting "because others will ask". The expectation of being asked motivates turnout if individuals derive pride from telling others that they voted, or feel shame from admitting that they did not vote, provided that lying is costly. In a door-to-door survey about election turnout, we experimentally vary (1) the informational content and use of a flyer pre-announcing the survey, (2) the duration and payment for the survey, and (3) the incentives to lie about past voting. The experimental results indicate significant social image concerns. For the 2010 Congressional election, we estimate a value of voting "to tell others" of about \$15, contributing 2 percentage points to turnout. Finally, we evaluate a get-out-the-vote intervention in which we tell potential voters that we will ask if they voted.

Key words: Turn out, Behavioral political economy, Get-out-the-vote, Structural behavioral economics

JEL Codes: C93, D03, D72, P48

1. INTRODUCTION

Get-out-the-vote interventions in the spirit of Gerber and Green (2000) have attracted significant attention by researchers and practitioners alike. In the most effective such intervention to date, a letter revealing the recipient's and their neighbour's turnout record increases turnout by 8 percentage points (Gerber *et al.*, 2008). This remarkable effect likely reflects social-image concerns: the explicit threat to make one's voting record public.

We consider a related, but more commonplace, social image motivation for voting. While it is rare for others to confront us with our voting record, it is common for neighbours, friends, and family to ask whether we voted. If individuals care about what others think of them, they may derive pride from telling others that they voted or feel shame from admitting that they did not vote. In addition, they may incur disutility from lying about their voting behaviour.

Such individuals are motivated to vote (in part) because they anticipate that others will ask if they did. If they vote, they can advertise their "good behaviour" when asked. If they do not vote, they face the choice of being truthful but incurring shame, or saying that they voted but incurring the lying cost. This trade-off is reflected in the established fact that 25–50% of non-voters lie when asked about their past turnout (Silver *et al.*, 1986; Belli *et al.*, 1999).[1]

In this article, we estimate this model of voting "to tell others", which follows Harbaugh (1996), using a natural field experiment. Due to the tight link between the model and the experiment, we are able to estimate the value of voting due to this social image motivation. We consider this a significant contribution given the rarity of estimates of the value of voting in the political economy literature.

The main experiment took place in the summer and fall of 2011 in the suburbs of Chicago. We visited households and asked whether they were willing to answer a short survey, including a question on whether they voted in the 2010 congressional election. In some cases, we posted a flyer on the doorknob a day in advance to announce the upcoming survey. Unbeknownst to the households, we used voting records to restrict the sample to households where either all registered members voted in the 2010 elections (henceforth, voting households) or none of the registered members voted in 2010 (non-voting households). We did not visit households with a mixed 2010 voting record.

The field experiment has three main sets of treatments. In the first set, we randomize the information on the flyer. In one group, the flyer informs households that the next day we will visit their home to ask them to complete a survey. In a second group, the flyer specifies that the survey will be about "your voter participation in the 2010 congressional election". Differences in the share of households opening the door and completing the survey between the first and the second group reflect the anticipated value of being asked about voting. An increase in the participation of voting households would indicate the pride of saying that one voted. A decrease among non-voting households would indicate shame from admitting that one did not vote, or a cost of lying and claiming to have voted.[2]

We find that, on average, voters do not sort in. In fact, voting households are slightly less likely to answer the door and do the survey when they are informed about the turnout question. Non-voters sort out significantly, decreasing their survey participation by 20%.[3]

These results may depend on the particular election. The 2010 elections were disappointing for Democrats and positive for Republicans, including in Illinois the loss of President Obama's previous seat in the Senate. The lack of pride among voters may reflect disappointment, given that the neighbourhoods visited were largely Democratic. Indeed, if we restrict the analysis to voters registered for the Republican primaries, we find evidence of sorting in.

The findings on sorting provide prima facie evidence of social-image utility. In order to quantify the utility value, we measure the cost of sorting in and out of answering the survey. To do so, we introduce a second set of (crossed) randomizations, in which we vary the promised payment for the survey ($10 versus $0) and the pre-announced duration (5 minutes versus 10 minutes). We find that the effect of reducing payment by $10 is comparable to the sorting response of non-voters to the election flyer, implying significant social-image (dis)utility.

To estimate the value of voting "to tell others", we need additional counterfactual social-image values, such as the shame that voters would feel were they to say they did not vote. These counterfactuals are not provided by the sorting moments.

1. Memory failures or inaccurate voting records do not appear to explain the discrepancy (Duff *et al.*, 2007).

2. This randomization also includes a group with no flyer, as well as a group with an opt-out box.

3. We also cross-randomize the information provided by the surveyor at the door. For half of the households, they indicated a survey "on your voter participation in the 2010 congressional election". This manipulation did not have a significant effect on survey take-up for either voters or non-voters.

Thus, in a third set of crossed treatments we randomize incentives to provide a different response to the turnout question. We inform half the respondents of the ten-minute survey that the survey will be eight minutes shorter if they state that they did not vote in the 2010 congressional election. For voters, this treatment amounts to an incentive to lie and permits us to quantify the disutility of voters were they to say (untruthfully) that they did not vote. For the 50% of non-voters who lie without such incentives, this treatment provides an incentive to tell the truth. We provide a parallel \$5 incentive in the 5-minute survey to state that one did not vote.

This novel experimental design makes it possible to price out how much respondents care about making a particular statement. This approach has applications to other settings where responses could have social image or signalling motivations, such as in contingent valuation surveys, or surveys of sensitive political and social attitudes.

The results reveal that non-voters are significantly more sensitive to these incentives than voters. When incentivized, the share of non-voters who lie decreases significantly, by 12 percentage points, while the share of voters who lie increases only insignificantly, by 2 percentage points. The results are similar for time and monetary incentives, and reveal a strong preference of voters for saying that they voted.

We combine the moments from the three sets of treatments to estimate the parameters of our model using a simulated minimum-distance estimator. The estimation accounts for the fact that the social-image variables affect not just the predicted response to the experimental treatments, but also whether an individual would choose to vote in the first place. Individuals select into being voters or non-voters as a function of the social-image values and of their other reasons to vote, such as expressive voting or voting costs. Thus, a voter is more likely to care about social image than a non-voter, and also more likely to have, say, low voting costs.

The benchmark estimates provide evidence of significant social-image value in voting. We estimate that individuals assign on average a \$7 value to being seen as a voter rather than a non-voter, when asked once by a surveyor. This social-image value does not come from pride in voting: in fact, individuals on average prefer not to be asked, even when they can say truthfully that they voted. Rather, they assign a quite negative value to admitting to not voting.

We estimate that individuals assign a disutility of \$7 to lying about voting to a surveyor. The combination of social-image utility and sizable lying costs implies that the anticipation of being asked provides a reason to vote.

To quantify the value of voting "to tell others", we combine these estimates with survey-based evidence on the number of times that people report being asked whether they voted, five times on average for the 2010 congressional election. Altogether, we estimate a value of voting "to tell others" of \$18 for voters and \$13 for non-voters.

There are caveats to this estimate. For example, this magnitude likely understates the value of voting "to tell others", since it is based on being asked by a (previously unknown) surveyor. Even taking these caveats into account, our ability to assign a dollar value to voting through the design is a unique contribution to the literature. For example, Coate and Conlin (2004) and Coate *et al.* (2008) estimate, respectively, a group-rule utilitarian model and a pivotal-voting model on alcohol-regulation referenda data. Their estimates for the value of voting are up to a scaling for the voting cost, which is not identified; thus, they do not provide a monetary value of voting due to their model. Levine and Palfrey (2007) estimate a pivotal-voting model, but use laboratory elections where parameters can be controlled. The unique element in our design that makes the difference is that we use monetary inducements—variation in the value of the survey and incentive to lie about voting—to translate the findings into a monetary value of voting.

A second metric to evaluate the model of voting "to tell others" is in terms of the extra turnout that it generates. The baseline turnout in our setting is 60%. How much would turnout change if people stopped asking others whether they voted, for example because of a shift in norms?

What about if conversely the rate of asking doubled, perhaps because campaigns encourage such asking? We predict that eliminating asking about voting would lower turnout by 2 percentage points, while doubling the number of times asked would increase turnout by 2 points. While these impacts may seem small, consider the extraordinary amount of effort that campaigns put into get-out-the-vote efforts, with the average such letter yielding a turnout impact of 0.2 percentage points (Green *et al.*, 2013).

The main field experiment was designed to measure the value of voting without affecting voting itself—a crucial difference from the get-out-the-vote literature. Instead, we rely on sorting, survey completion, and survey responses. This allows us to estimate the magnitudes and signs of the social image utility associated with being asked about voting (a common occurrence). But the model also suggests an obvious intervention to increase turnout: individuals with social-image motives are more likely to vote, the more they expect to be asked. Experimentally increasing this expectation should thus lead to an increase in turnout.

In November of 2010 and of 2012, we did just that. A few days before the election, a flyer on the doorknob of treatment households informed them that "researchers will contact you within three weeks of the election [...] to conduct a survey on your voter participation".[4] A control group received a flyer with a mere reminder of the upcoming election. The results are consistent with the model, though statistically imprecise. In 2010, the turnout of the treatment group is 1.3 percentage points higher than the control group (with a one-sided p-value of 0.06). In the higher-turnout presidential election of 2012, the turnout difference is just 0.1 percentage points (not significant). The results are consistent with the contemporaneous results of Rogers *et al.* (2016), who also inform a treatment group that they may be called after the election about their voting behaviour, and find a positive impact on turnout (of 0.2 percentage points). The much smaller effect sizes than in Gerber *et al.* (2008) are not surprising, since they estimate the effect of informing neighbours about the official turnout record, while we isolate the effect of (at most) one more interaction with a questioner, where lying remains an option.[5]

Are the get-out-the-vote results consistent with the estimates of the value of voting? Using the model estimates, we predict that an announced visit to ask about voting would increase turnout by 0.3 percentage points, well within the point estimates of the estimated effects. Thus, the get-out-the-vote results are qualitatively consistent with the model, if imprecise.

Finally, we would like to mention some caveats and alternative interpretations. First, we address the important concern that the observed "sorting out" among non-voters may reflect a dislike of talking about politics, rather than any stigma from not voting. When we allow for a different taste among voters and non-voters for talking about politics, we lose the ability to estimate one of the social-image parameters. But the net value of voting "to tell others" is still identified and in fact remains unchanged, since it is identified by the lying treatments. Intuitively, while a differing taste for talking about politics could explain the sorting patterns in response to the flyer treatments, it does not explain the lying about voting, nor the differential response of voters and non-voters to the lying incentives.

Second, the results are specific to their time and location—the 2010 congressional elections in Illinois. As we discussed, the lack of estimated pride in voting is possibly related to the disappointing results for Democrats in 2010, and could more generally be a function of the aggregate turnout rates, closeness and importance of the election. It will be interesting to apply this methodology to other elections to test this directly in future work.

4. We follow up with a door-to-door visit, as advertised.

5. In addition, Gerber *et al.* (2008) focus on non-competitive primary elections, where turnout interventions lead to larger effects than for competitive general elections like the ones we study.

Third, in a series of robustness and sensitivity checks, we relax and vary numerous assumptions of the model, including modelling heterogeneity in lying cost, allowing for measurement error in the voting record, and omitting groups of moments. The estimated value of voting "to tell others" remains largely robust to these variations.

In addition to complementing the substantial literature on get-out-the-vote field experiments, summarized in Green and Gerber (2008), this article more broadly contributes to the vast literature on why people vote.[6] Our main contribution is to provide an estimate of the value of voting and a welfare evaluation of a get-out-the-vote campaign, which is rare in the literature. We obtain these estimates by virtue of the design of the field experiment.

The article also relates to the literature on social image. The theoretical papers provide a micro-foundation for social-image concerns as signalling models (Benabou and Tirole, 2006; Andreoni and Bernheim, 2009; Ali and Lin, 2013) and suggest intriguing possibilities for how our estimated social image parameters might vary in different elections with differing degrees of turnout, closeness, and importance. The empirical papers highlight the impact of social image on productivity (Ashraf *et al.*, 2014), contributions to public goods (Ariely *et al.*, 2009; Lacetera and Macis, 2010), campaign contributions (Perez-Truglia and Cruces, 2013), and energy consumption (Allcott, 2011). Our study attempts to bring these literatures closer by providing estimates of the social-image parameters. We hope that future research strengthens the ties, providing estimates of the underlying signalling game.

This article also complements a small but growing literature on behavioural political economy, including Shue and Luttmer (2009), Finan and Schechter (2012), Passarelli and Tabellini (2016), and Bursztyn *et al.* (2014). This article links this literature with the literature on structural behavioural economics (Laibson *et al.*, 2007; Conlin *et al.*, 2007; DellaVigna *et al.*, 2012).

The remainder of the article proceeds as follows. The next section introduces the model. Section 3 summarizes the experimental design. Sections 4 and 5 present, respectively, the reduced-form results and structural estimates for the main experiment. Section 6 introduces the get-out-the-vote experiment and Section 7 concludes.

2. MODEL

2.1. *Voting*

Voting depends on four factors: pivotality, warm glow, cost of voting, and expected social image. Individuals vote if the net expected utility of doing so is positive:

$$pV+g-c+N[\max(s_V,s_N-L)-\max(s_N,s_V-L)]\geq 0. \tag{1}$$

The first three terms in expression (1) capture the standard model of voting. The first term is the expected utility of being pivotal (Downs, 1957), with a pivotality probability p and value V assigned to deciding the election. The second term, g, is the warm glow from voting (as in Riker and Ordeshook (1968)). The third term, $-c$, is the transaction cost of going to the polls. Since our experimental design does not focus on these components, only their sum will matter, which we denote by $\varepsilon=pV+g-c$. We assume ε has c.d.f. H.

The crux of the model is the fourth term, the social-image motivation to vote (in the spirit of Harbaugh (1996)). An individual expects to be asked N times whether she voted, and has to

6. This vast literature includes Downs (1957), Ledyard (1984), and Palfrey and Rosenthal (1983), Palfrey and Rosenthal (1985) on pivotal voting and Riker and Ordeshook (1968), Harsanyi (1977), Blais (2000), and Feddersen and Sandroni (2000) on norm-based voting.

decide whether to be truthful or to lie. Assume first that she has truly voted. In this case, she can (truthfully) state that she voted, which earns her utility s_V; or she can lie and look like a non-voter, which earns her utility s_N minus a psychological lying cost L. Therefore, the utility a voter receives when being asked about her turnout is $z^v \equiv \max(s_V, s_N - L)$. Now assume that she did not vote. In this case, she can either state the truth and obtain the utility from appearing to be a non-voter, s_N, or lie and obtain s_V minus the lying cost L. Hence, the utility of being asked for a non-voter is $z^{nv} \equiv \max(s_N, s_V - L)$. The term in square brackets in (1) is therefore the net utility gain from voting due to being asked once.

The terms s_V and s_N capture how much the individual cares about being seen as a public good contributor (voter), or not, by others. These terms can be understood as reduced-form representations of a signalling model, such as Benabou and Tirole (2006) and Ali and Lin (2013). Experimental evidence suggests that information about whether a person votes affects how favourably they are viewed by others (Gerber *et al.*, 2016).

The term L captures the utility cost of lying. We assume that the cost of lying is non-negative, $L \geq 0$, and additive with respect to the social-image term. The assumption of positive lying costs is motivated by introspection and by experimental evidence documenting that in cheap talk communication games, which are similar to survey questions, a sizeable portion of subjects prefer to tell the truth even when lying is profitable.[7]

We stress four important assumptions. First, the variables s_V, s_N, and L could depend on a variety of factors, such as whether one's party won the election, or the closeness of the race. One could envision a model, for example of social signalling, which would provide a micro-foundation for the values of the variables.[8]

Second, we can relate this model to a more general model. Assume that the utility $U_{k,s}$ of being asked about voting depends on whether one voted or not ($k = V, N$), and on whether one tells the truth or lies ($s = T, L$). Our model can be mapped to such a model: $U_{V,T} = s_V$, $U_{V,L} = s_N - L$, $U_{N,T} = s_N$, and $U_{N,L} = s_V - L$. We impose the restriction that the lying cost in the two cases (V, N) is equal.

Third, we assume that the value of being asked increases linearly in the number of times asked N, an assumption that is untested. Alternatively, the value of being asked may be concave in the number of times asked, with later asks yielding less disutility, say, from lying about voting. If that is the case, we likely are underestimating the value of voting to tell others. By the time we ask voters about their past turnout, months after the election, respondents on average have already been asked multiple times. The value of our marginal ask, which we estimate with our treatments, would understate the value of being asked the previous times.

A final point is about heterogeneity. We assume that s_V, s_N, and ε are stochastic and heterogeneous. As we explain in Section 5, individuals become voters or non-voters following (1) depending on the draws of the variables, thus inducing systematic differences between voters and non-voters in these variables. In the benchmark model, for simplicity we assume that the lying cost L is instead deterministic and identical for voters and non-voters, though we relax this assumption in a robustness check. In another robustness check, we also estimate a model that allows for differences between voters and non-voters in the utility of talking about politics, another form of heterogeneity.

7. The model assumes that the respondents do not have an option of refusing to answer the vote question. This seems justified by the data: out of 1,738 people that agreed to do the survey, the answer to the turnout question is missing in only five cases.

8. Indeed we present evidence in our setting suggesting different values for registered Republicans and Democrats.

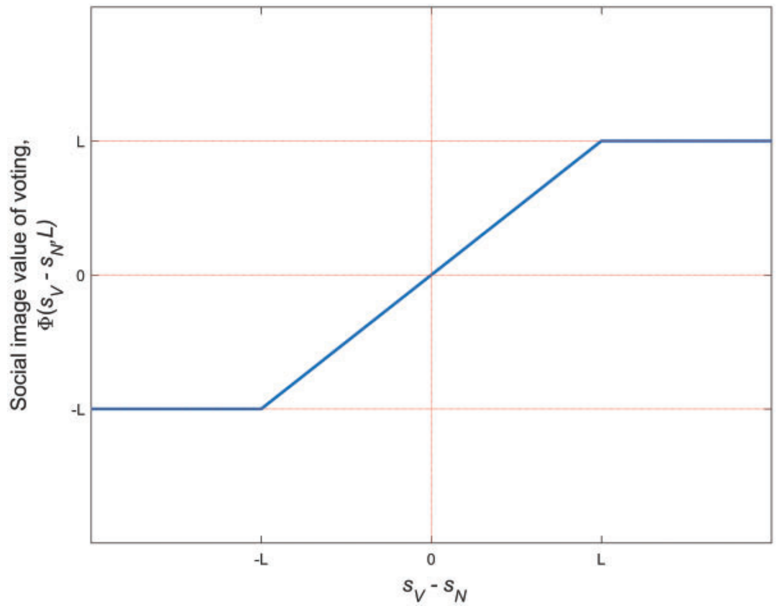


FIGURE 1
Social-image value of voting per interaction

Note: Figure 1 plots the social-image value of voting due to the anticipation of being asked once, as a function of the net social image utility $s_V - s_N$ and the cost of lying L.

Returning to our model, we do not impose any restrictions on s_V and s_N, but we consider two special cases: (1) *Pride in Voting* ($s_V > 0$): individuals care (positively) about stating that they are voters; (2) *Stigma from Not Voting* ($s_N < 0$ and $s_V - L < 0$): individuals dislike both (truthfully) admitting to being non-voters and (untruthfully) saying that they are voters. Notice that both conditions could hold, for $s_V > 0 > s_N$, provided L is large enough.

Using the abbreviated notation ε for the other reasons to vote, we can rewrite the voting condition (1) as $N\Phi(s_V - s_N, L) + \varepsilon \geq 0$, where

$$\Phi(s_V - s_N, L) = \begin{cases} L & \text{if } s_V - s_N \geq L \\ s_V - s_N & \text{if } -L \leq s_V - s_N < L \\ -L & \text{if } s_V - s_N < -L. \end{cases} \tag{2}$$

As expression (2) shows, voting depends on the net social-image value $s_V - s_N$ and on the cost of lying L. Figure 1 displays $\Phi(s_V - s_N, L)$ as a function of $s_V - s_N$ and makes it clear that, in order for social image to contribute to voting, the net utility $s_V - s_N$ must be non-zero and the lying cost L must be positive. If either of these conditions is not met, then the individual either does not care about image, or can always signal the best-case scenario, irrespective of her true actions. Also notice that as long as individuals prefer to signal that they are voters ($s_V - s_N > 0$), the net value of being asked for voting is weakly positive.

2.2. *Door-to-door survey*

To estimate this model, we design a door-to-door survey in which individuals are asked, among other questions, whether they voted. We model the behaviour of an individual whose home is visited by a surveyor. If the visit is pre-announced by a flyer and the person notices the flyer (which occurs with probability $r \in (0,1]$), she can alter her probability of being at home and opening the door. A "survey flyer" (denoted by F) informs the reader when the surveyor will visit, but leaves the content of the survey unspecified. An "election flyer" (denoted by FE) additionally informs the reader that the survey will be about her voter participation in the previous election.

Once the surveyor visits the home, the respondent opens the door with probability h. If she did not notice the flyer (or did not receive one), h is equal to a baseline probability $h_0 \in (0,1)$. If she noticed the flyer, she can optimally adjust the probability to $h \in [0,1]$ at a cost $c(h)$, with $c(h_0)=0$, $c'(h_0)=0$, and $c''(\cdot)>0$. That is, the marginal cost of small adjustments is small, but larger adjustments have an increasingly large cost. We allow for corner solutions at $h=0$ or $h=1$. In the estimation, we assume $c(h)=(h-h_0)^2/2\eta$.

If the individual is at home at the time of the surveyor's visit and opens the door, she must decide whether to complete the survey. Consumers have a baseline utility s of completing a generic 10-minute survey for no monetary payment. The parameter s can be positive or negative to reflect that individuals may find surveys interesting, or they may dislike surveys. In addition, individuals receive utility from a payment m and disutility from the time cost c, for a total utility from survey completion of $s+m-c$. The time cost c equals τv_s, where τ is the duration of the survey in fraction of hours, and v_s is the value of one hour of time. As in DellaVigna *et al.* (2012), the respondent pays a social pressure utility cost $S \geq 0$ for refusing to do the survey when asked in person by the surveyor. There is no social pressure if the individual does not open the door when the surveyor visits. We assume that the respondent is aware of her own preferences and rationally anticipates her response to social pressure. In addition to the baseline utility $s+m-c$ of doing a survey, there is the additional utility from being asked about voting, z^v for voters and z^{nv} for non-voters, as defined above.

We also vary whether the survey content is announced to the respondent when she opens the door with two "announcement" treatments, $a \in \{I, NI\}$. When informed that the survey will ask about her voter participation ($a=I$), an individual will consider the utility of being asked about voting, z^i, while deciding whether to complete the survey. If she is instead not informed at the door ($a=NI$), she will neglect z^i—provided she has not already seen an election flyer. This announcement treatment is in the spirit of the election flyer treatment, but by design can only affect survey completion, not the probability of opening the door.

Finally, in some treatment cells we provide an incentive for the respondents to say that they did not vote; the incentive is either in terms of time—an 8-minute shortening of the survey duration—or money—an extra \$5 for 1 more minute of questions. We denote by I the monetary value of the incentive. By incentivizing the respondent to say she did not vote, a voter is provided an incentive to lie, and will lie if $s_N - L + I \geq s_V$. In contrast, a non-voter is provided an incentive to tell the truth, and will do so if $s_N + I \geq s_V - L$. By comparing the treatments with and without incentive I, we estimate the distribution of $s_V - s_N + L$ for voters and of $s_V - s_N - L$ for non-voters. Note that this treatment is unanticipated, and hence does not appear in the respondent's decision to answer the door or participate in the survey.

2.3. *Solution*

Conditional on opening the door, the respondent agrees to the survey if $s+m-c+z \geq -S$ assuming the respondent knows that the survey is about the election and if $s+m-c \geq -S$ otherwise. Working backwards, consider a respondent who sees a survey flyer (which does not mention the election questions). The decision problem of staying at home and opening the door (conditional on seeing a flyer) is $\max_{h\in[0,1]} h \max(s+m-c, -S) - (h-h_0)^2/2\eta$, leading to the solution $h^* = \max[\min[h_0 + \eta \max(s+m-c, -S), 1], 0]$. An increase in pay m or a decrease in the time cost c will increase the probability of opening the door and completing a survey. The parameter η determines the responsiveness of opening the door to incentives. Alternatively, for a respondent who sees the election flyer the solution is given by $h^* = \max[\min[h_0 + \eta \max(s+m-c+z, -S), 1], 0]$. If $z > 0$, the respondent will open the door with a weakly higher probability with the election flyer, compared to the survey flyer, and vice versa if $z < 0$.

Finally, for both the survey flyer and the election flyer, there is a variant with an opt-out box (denoted by OO and OOE, respectively) which makes avoidance of the surveyor easier. In terms of the model, this is equivalent to the agent being able to costlessly reduce the probability of being at home and opening the door to zero. Formally, $c(0) = 0$ and $c(h)$ is as above for $h > 0$.[9] The optimal probability of being at home and opening the door h^* remains the same as without the opt-out option if there is no social pressure and, hence, no reason to opt out (since the respondent can costlessly refuse to do the survey) or if the agent expects to derive positive utility from completing the survey. In the presence of social pressure, however, the respondent opts out if the interaction with the surveyor lowers utility.

The following Propositions summarize the testable predictions about the impact of the election flyer (Propositions 1 and 2), about the incidence of lies about past turnout (Proposition 3) and about the expected number of times asked, which we manipulate in the get-out-the-vote intervention (Proposition 4).[10]

Proposition 1. (Pride in Voting) *With Pride in Voting, the probability of opening the door $P(H)$ and of survey completion $P(SV)$ for voters is higher under the election flyer than under the survey flyer: $P(H)^v_{FE} \geq P(H)^v_F$ and $P(SV)^v_{FE} \geq P(SV)^v_F$. Parallel results hold for the opt-out flyers: $P(H)^v_{OOE} \geq P(H)^v_{OO}$ and $P(SV)^v_{OOE} \geq P(SV)^v_{OO}$. The probability of survey completion for voters is higher when informed at the door that the survey is about voting: $P(SV)^v_I \geq P(SV)^v_{NI}$.*

Proposition 2. (Stigma from Not Voting) *With Stigma from Not Voting, the probability of opening the door $P(H)$ and of survey completion $P(SV)$ for non-voters is lower under the election flyer than under the survey flyer: $P(H)^{nv}_{FE} \leq P(H)^{nv}_F$ and $P(SV)^{nv}_{FE} \leq P(SV)^{nv}_F$. Parallel results hold for the opt-out flyers: $P(H)^v_{OOE} \leq P(H)^v_{OO}$ and $P(SV)^v_{OOE} \leq P(SV)^v_{OO}$. The probability of survey completion for non-voters is lower when informed at the door that the survey is about voting: $P(SV)^{nv}_I \leq P(SV)^{nv}_{NI}$.*

Proposition 3. (Lying about Voting) *If the net social-image utility is positive, the probability of lying about past voting, $P(L)$, should be zero for voters and larger for non-voters assuming no incentives to lie ($I = 0$): $P(L)^v = 0 \leq P(L)^{nv}$ for $s_V - s_N > 0$. For any social-image utility, the probability of lying is (weakly) increasing in the incentive I for voters and (weakly) decreasing in I for non-voters: $\partial P(L)^v / \partial I \geq 0$ and $\partial P(L)^{nv} / \partial I \leq 0$.*

Proposition 4. (Times Asked) *The probability of voting is increasing in the number of times asked N if the social-image utility is positive and lying costs are positive: $\partial P(V) / \partial N \geq 0$ for $s_V - s_N > 0$ and $L > 0$.*

9. This formalization allows a costless reduction of h to 0 but not to other levels. This is not a restriction because agents who prefer to lower h below h_0 (at a positive cost) will strictly prefer to lower h to 0 at no cost.

10. The proofs are in Supplementary Appendix A.

3. EXPERIMENTAL DESIGN

3.1. *Logistics*

We employed 50 surveyors and many flyer distributors, mostly undergraduates at the University of Chicago, who were paid $10.00 per hour. Most surveyors conducted surveys over multiple weekends.[11] The distribution of flyers took place on Fridays and Saturdays, and the field experiment took place on Saturdays and Sundays between July 2011 and November 2011. The locations are towns around Chicago shown in Appendix Figure 1.[12]

The unit of treatment assignment is a route, which consists of typically 13 households on a street, to be reached within a half-hour. On a day, a surveyor has a workload of 8 routes (10 am–12 pm and 1–3 pm). Every half-hour, the surveyor moves to a different street in the neighbourhood and begins a new route of 13 homes, typically entering a different treatment in the next route. Surveyors do not know whether a treatment involves a flyer, though they can presumably learn that information from observing flyers on doors.

To determine the households in the sample, we obtain voting records from the Election unit of the Cook County Clerk's office in January 2011. We begin with the full sample of addresses with at least one adult registered to vote. We then reduce the sample to households with homogeneous voting records in the congressional elections of November 2010[13]: either every registered voter at the address voted in 2010, or no one did. Next, we randomize these households to a treatment at the surveyor-route level. Houses are grouped into surveyor-routes, which are then randomized to treatments. The treatment is a combination of four crossed interventions: (1) flyer treatments, (2) payment and duration of the survey, (3) survey content announcement at the door, and (4) incentives to claim non-voter status.

3.2. *Treatments*

Each household was randomized into five flyer treatments with equal weights: *No Flyer, Survey Flyer, Election Flyer, Opt-Out Flyer*, and *Election Opt-Out Flyer*. Households in the *No Flyer* treatment receive no flyer. Households in the *Survey Flyer* treatment receive a flyer on the doorknob announcing that a surveyor would approach the home the next day within a specified hour (*e.g.* 3 pm–4 pm, see top left example in Figure 2). Households in the *Election Flyer* treatment receive a similar flyer, with the added information that the survey will be about "your voter participation in the 2010 congressional election" (second flyer from left in Figure 2). Households in the *Opt-Out Flyer* treatment receive a flyer as in the *Survey Flyer* treatment, except for an added check-box which the household can mark if it does not wish to be disturbed (third flyer from left in Figure 2). Similarly, the flyer in the *Election Opt-Out Flyer* treatment has an added opt-out check box. The flyers were professionally produced.

11. Additional details about the experiment, including the recruitment process, are in the Supplementary Appendix.

12. Arlington Heights, Elk Grove Village, Evanston, Glenview, Hoffman Estates, Lincolnwood, Mount Prospect, Northbrook, Oak Park, Park Ridge, Schaumburg, Skokie, Streamwood, Wilmette, and Winnetka. On almost all days, we visited one or two towns on a given day.

13. The ballot for the 2010 Congressional elections included, in addition to House and Senate races, ballots for governor and lieutenant governor, attorney general, secretary of state, comptroller and treasurer, and state senators. There are also Cook county-specific candidates and a proposed amendment to the Illinois constitution, and two town-specific referenda. The items on the ballot would largely be the same in the towns surveyed. Any differences, such as in the local referenda, would be controlled by town fixed effects (with the minor exception of towns spanning different congressional districts).

FIGURE 2

Flyer samples

Note: The top three flyers are for (5-min., $0) surveys in treatments Flyer (left), Flyer election (centre), and Opt Out (right). The bottom three flyers are for Flyer treatments (5-min., $0), (10-min., $10), and (5-min., $10).

A second crossed randomization involves the duration of the survey as well as the compensation offered (if any) for completing the survey. The bottom row of Figure 2 displays flyers for the three treatments: (5-Minutes, No Payment), (10-Minutes, $10 Payment), and (5-Minutes, $10 Payment), each sampled with equal probability. In each of these treatments, we reiterated the compensation and duration at the door.

The third set of crossed treatments involves how the surveyors described the survey once, after a knock on the door, a household member answered. The respondents were told "We are conducting confidential _ _ _ minute surveys in _ _ _ today. [You would be paid $ _ _ _ for your participation.]", with the empty fields filled depending on the payment and duration treatments and the assigned town. The *No Information* group was then simply asked "Do you think you might be interested?". The *Information* group was instead told "The survey is about your voter participation in the 2010 congressional election. Do you think you might be interested?". Hence, the Information treatment provides information about the content of the survey in a similar way to the Election Flyer treatment. Respondents in the *Election Flyer* or *Election Opt-out Flyer* already knew about the content, provided they read the flyer. The top part of Figure 3 summarizes this first set of crossed treatments.

The fourth set of crossed treatments, summarized at the bottom of Figure 3, involves incentives to affect the response to a turnout question. In control surveys, individuals are simply asked whether they voted in the 2010 congressional election. For a subject in a 10-minute, $10 survey in the treatment group, we offer an 8-minute incentive to the respondent to state that he or she did not vote. After the first question in the survey, the surveyor reads aloud: "We have 10 minutes of questions about your voter participation in the 2010 congressional election, but if you say that you did not vote then we only have 2 minutes of questions. Either way you answer you will be paid $10. That is, we have 10 minutes of questions, but if you tell us no to the question 'did you vote in the 2010 congressional election' then we only have 2 minutes of questions to ask. Regardless of your answer you will earn $10." The surveyor then points to where the survey ends if the respondent answers "no", in which case the survey is indeed much shorter.

For respondents assigned to a 5-minute survey, we did not assign a time discount which could only have been a modest 3-minute reduction. Instead, we provide a monetary incentive to the treatment group as follows (with the material in brackets applying only to the (5-Minutes, $10 Payment) conditions): "We have 5 minutes of questions about your voter participation in the 2010 congressional election, but if you say that you did not vote then we have 1 extra minute of questions and we will pay you an extra $5 for answering these additional questions [.IF PAID: for a total of $15]. If you say that you voted then we will just ask you the original 5 minutes of questions. [.IF PAID: and pay you $10 as promised.] That is, we have 5 minutes of questions, but if tell us no to the question 'did you vote in the 2010 congressional election' then we have 1 extra minute of questions and you will earn an additional $5 for answering these questions." Conditional on a 5-minute or a 10-minute survey, we determined the incentive or no-incentive treatment with equal weights.[14]

Overall, we conducted sixty treatments: five flyer types crossed with three survey lengths, whether or not the survey content was announced at the door, and two types of lying incentives.

Finally, we followed the promises made: we pay the individuals as promised, and we conducted a longer survey when the survey was advertised as lasting 10 minutes rather than 5 minutes. Further, in the treatments with a lying incentive, if the subject responded "no" to the turnout question, the survey duration and payment were altered as promised. The complete survey transcript is in the Supplementary Appendix.

3.3. *Sample*

We reached a total of 14,475 households. From this initial sample, we exclude 1,278 observations in which the households displayed a no-solicitor sign (in which case the surveyor did not contact

14. To keep a parallel structure, in the treatments with no incentive to lie about voting, there is instead an incentive to lie about the year of house purchase.

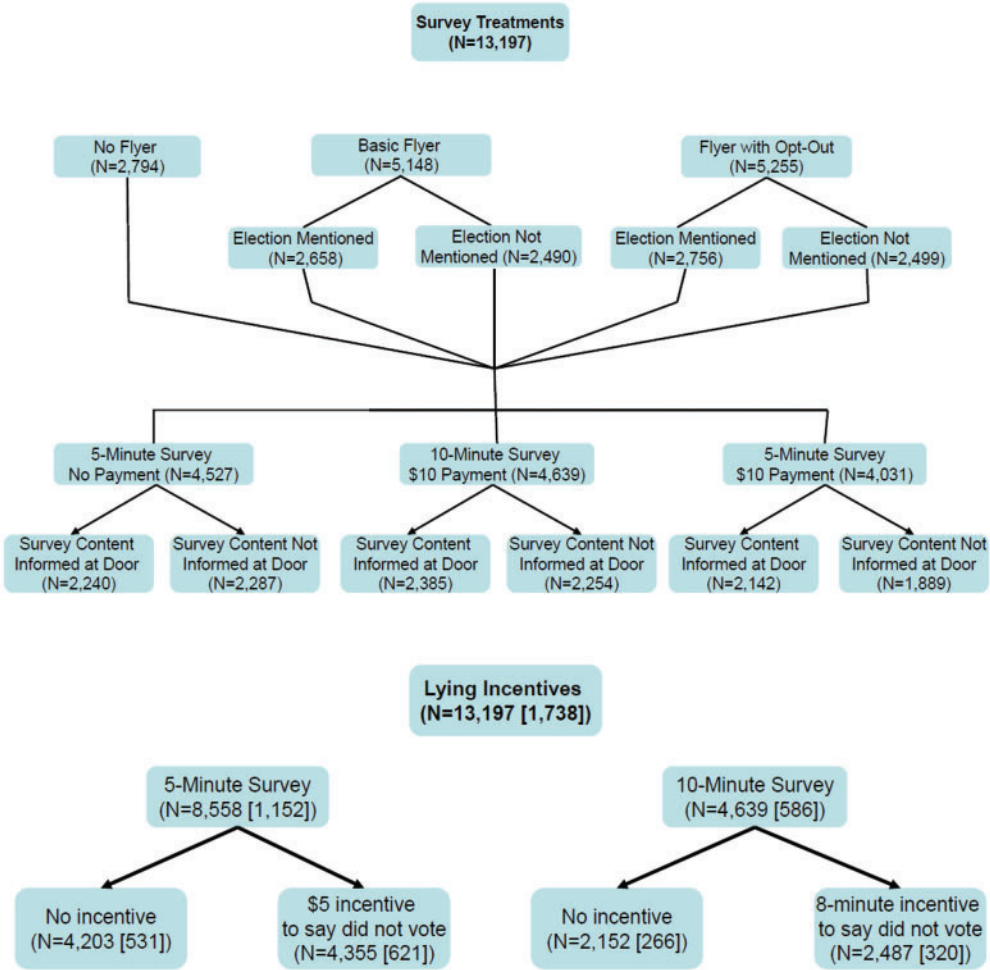


FIGURE 3

Experimental treatments

Note: Figure 3 presents the crossed experimental randomizations, with sample sizes in parentheses. On top are the five arms of the flyer treatment, crossed with whether respondents at the door are informed that the survey is about participation in the 2010 congressional election, crossed with survey duration and payment. At the bottom are the arms of the lying incentives, indicating both the initial sample size and [in square brackets] the sample size among individuals who responded to the survey. All arms are equally weighted and crossed.

the household) or the surveyor was not able to contact the household for other reasons (*e.g.* a lack of access to the front door or a dog blocking the entrance).[15] The final sample includes 13,197 households.

15. The rate at which the subjects are dropped is comparable across the different flyer treatments, but is higher in the no-flyer treatments (14% versus 8%). The reason is that households with a no-solicitor sign in the flyer treatments are excluded altogether from the sample when flyerers find the no-solicitor sign on their flyering visit; these houses are not visited the next day, to save time. This does not happen in the no-flyer treatment since there is no flyering visit. Thus, the no-flyering treatments include in the sample more no-solicitor households (previous to us dropping them). This being said, this difference plays a minimal role since the no-flyer treatments only help to identify auxiliary parameters. As we show in Column 5 of Supplementary Appendix Table 5, the results are similar if we do not drop any observations.

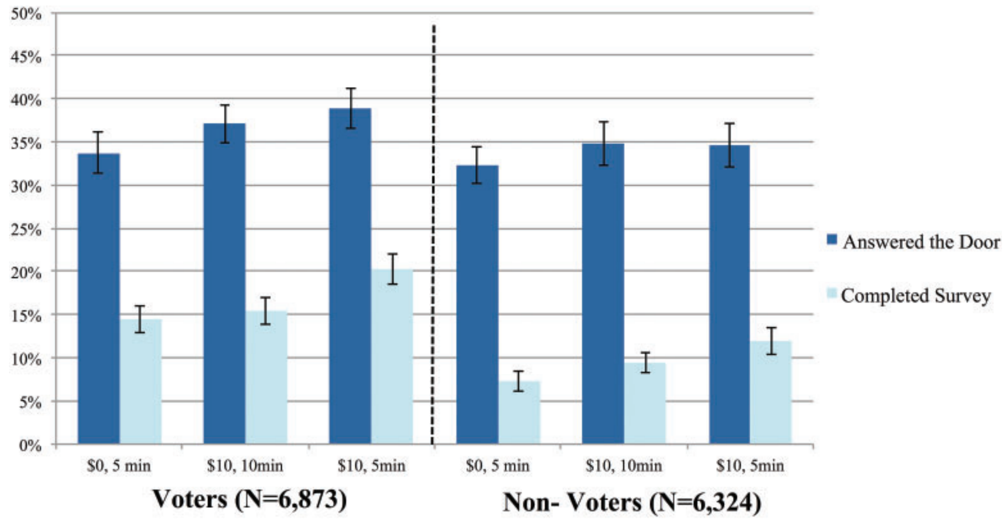


FIGURE 4

Response to survey duration and payment

Note: Figure 4 presents the share of households answering the door and the (unconditional) share completing the survey across the three different combinations of payment and duration, separately for voting households and non-voting households. The averages are pooled across the different flyer treatments featured in Figure 3. Standard errors are clustered at the solicitor-date level.

4. REDUCED-FORM ESTIMATES

4.1. *Opening the door and survey completion*

We present graphical evidence in Figure 4 on the share of households opening the door and completing the survey as a function of the survey details, pooling across the 5 flyer treatments. Voters are very responsive to incentives, going from 33% opening the door for a \$0, 5-minute survey to 39% for the \$10, 5-minute survey. Hence, a \$10 incentive induces a 6 percentage point (20%) increase in the share opening the door. The effect is similarly large for the share completing the survey, a 6 percentage points (45%) increase. The responsiveness to incentives of non-voters is smaller with regards to opening the door, but is large with respect to survey completion: 5 percentage points (62%).

Having established that households are responsive to the survey incentives, we turn to the key flyer treatment—whether the flyer informs the household about the election question. Figure 5a plots the results for voters, pooling across the different survey durations and payment incentives. We do not observe much difference for voters in the share opening the door, or the share completing the survey, between the Survey Flyer and the Flyer Election treatments. In the Opt-out treatments, we observe a *decrease* in the share of voters opening the door and in the share completing a survey when the flyer informs about the survey content compared to when it does not. Thus, there is no evidence of pride from voting, and it appears that voters prefer not to be asked whether they voted.

For non-voters (Figure 5b), the difference between the Flyer and the Flyer Election treatments is large: there is a 6 percentage point drop (20%) in the probability of opening the door. The size of this effect is comparable to the effect of a \$10 incentive to complete the survey. There is a similar 3 percentage point (25%) decrease in the share completing a survey when the flyer announces the election question. The impacts are consistent but smaller in the opt-out treatments, with a 1.5 percentage point (15%) decrease in the share opening the door when the flyer mentions elections.

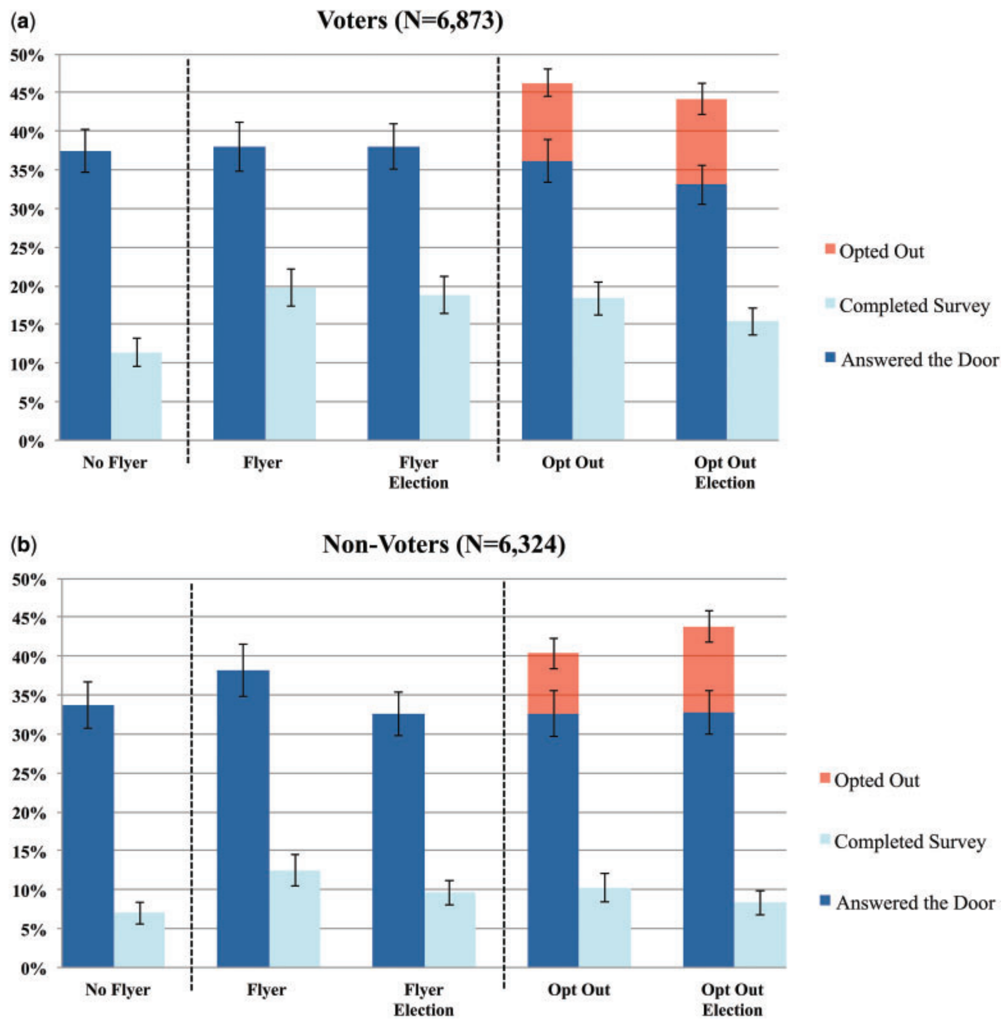


FIGURE 5

Response to information about election in flyer

Note: Figure 5 presents the share of households answering the door, the (unconditional) share completing the survey, and (when applicable) the share opting out, separately for each of the five flyer treatments and separately for voting households and non-voting households. The averages are pooled across the three different payment and duration treatments featured in Figure 3. Standard errors are clustered at the solicitor-date level.

These results indicate strong avoidance of non-voters, pointing to shame from admitting to not voting and disutility from lying.

These findings may depend on the context. The results of the 2010 congressional elections were very disappointing for Democrats, including in Illinois the loss of President Obama's seat in the Senate, and correspondingly positive for Republicans. The lack of evidence for pride among voters may well be due to disappointment, given that the neighbourhoods visited were largely Democratic. While our results are from a single election, we can differentiate the response based on the primary registration. In Figure 6, we present separate results for households with voters who participated in Republican primaries (left panel) versus households with voters registered in

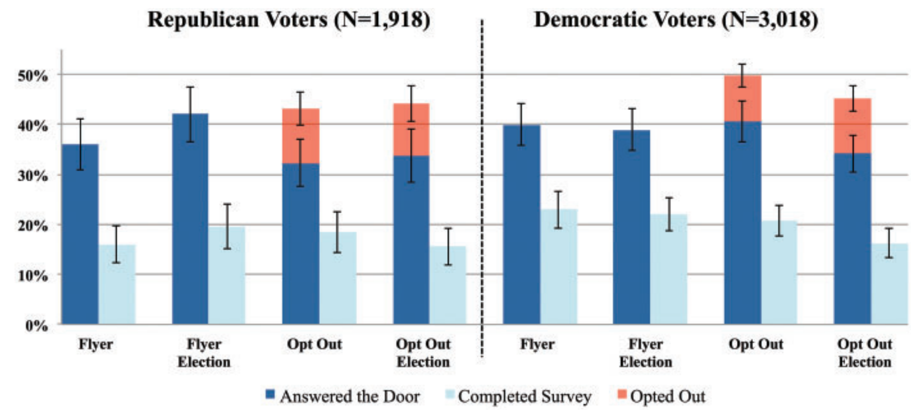


FIGURE 6

Response to flyer by party registration (*for voters*)

Note: Figure 6 presents the data from Figure 5 for *voting* households (omitting for space reasons the no-flyer treatment) split into two groups. In the left group, at least one household member voted at a Republican primary between 2004 and 2010. In the right group, at least one member voted at a Democratic primary between 2004 and 2010. Households with neither, or with voters participating in different party primaries, are not included. Standard errors are clustered at the solicitor-date level.

Democratic primaries (right panel).[16] Indeed, we detect sizeable sorting in by Republican voters in response to the election flyer, indicative of pride in voting in an election with positive results for the party. Among Democratic voters, instead, we observe sorting out as in the overall results, consistent with disappointment about the election. Among voters who did not participate in a primary (not shown), we also detect sorting out.

We now examine the effects of announcing the survey content at the door. Figure 7 plots survey completion rates by the door announcement type (Informed or Not Informed), pooling across all the flyer treatments. For voters, the effects of the door announcements are similar to those of the flyer announcements: there is no increase in survey completion from being informed about the voting question, and thus no evidence of pride. But non-voters also show essentially no effect on survey completion from being informed at the door. This is in contrast to the flyer treatments, where the election flyer leads to a sharp drop in opening the door and in survey completion by non-voters. We speculate that the difference (not captured in the model) could be that the flyer gives individuals time to think through the decision problem, while they must respond immediately when warned only at the door.

In Table 1, we present the regression analysis underlying Figures 4, 5a and b, 6, and 7 both with no controls and with fixed effects for surveyor i, day-town t, and hour-of-day h. We estimate, separately for voters and non-voters, the OLS regression:

$$y_{i,j,t,h} = \alpha + \Gamma T_{i,t,h} + \eta_i + \lambda_t + \zeta_h + \varepsilon_{i,j,t,h}, \tag{3}$$

where the dependent variable $y_{i,j,t,h}$ is, alternatively, an indicator for whether individual j opened the door (y^H) or agreed to complete the survey (y^S). The vector $T_{i,t,h}$ contains indicators for the various survey treatments, with the baseline No-Flyer treatment for a \$0, 5 minute survey as the

16. We record the most recent participation in primary elections by any registered member of the household. We define as "households with registered Republican voters" households where at least one voter has voted in a Republican primary, and no voter has voted in a Democratic primary. Vice versa for the definition of households with registered Democrats.

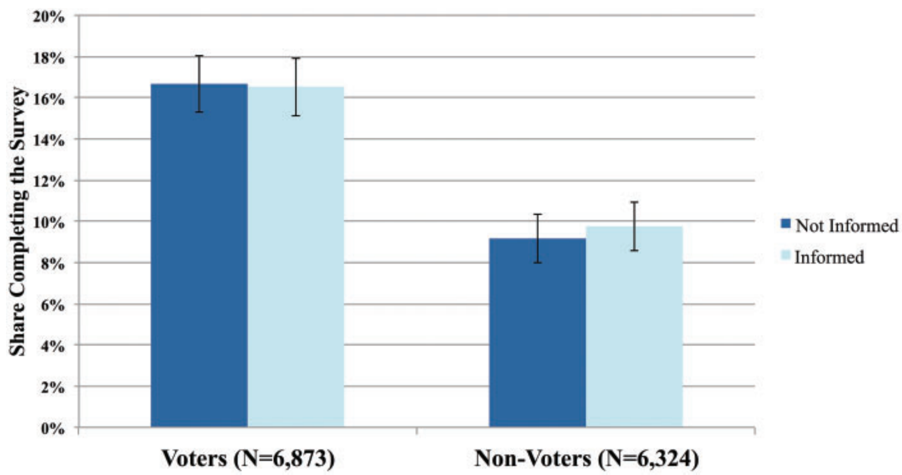


FIGURE 7

Response to announcement of survey content at door

Note: Figure 7 presents the (unconditional) share of households completing the survey, separately for voting and non-voting households. The households in the *Not-Informed* treatment are not informed ex ante about the survey content at the door. The households in the *Informed* treatment are told at the door that the survey will be about their voter participation in the 2010 congressional election. The averages are pooled across the different flyer treatments featured in Figure 3. Standard errors are clustered at the solicitor-date level.

omitted group. We cluster the standard errors at the surveyor×date level.[17] Table 1 shows that the results shown in the previous figures are robust to the inclusion of the surveyor, date-location, and hour fixed effects. In the Supplementary Appendix we present two sets of robustness results. In Supplementary Appendix Table 1 we allow for different effects of a surveyor on different dates and location by including surveyor*date*location fixed effects; the results are unaffected. In Supplementary Appendix Table 2 we present separate estimates for the first two months of the experiment (July and August 2011) and the next two months (October and November 2011); the results are comparable.[18] Finally, in Supplementary Appendix Table 3 we present the results split by political registration, as in Figure 6.

4.2. *Lying about voting*

Next, we estimate the rates at which voters and non-voters misrepresent their voting behaviour, and how these lies respond to the randomized incentives to lie (for voters) or to tell the truth (for non-voters). For the sample of individuals who completed the survey, we estimate the OLS regression

$$y_{i,j} = \alpha + \Gamma T_{i,j} + \eta_i + \varepsilon_{i,j}, \tag{4}$$

where $y_{i,j} = 1$ if individual j lied about her voting behaviour to surveyor i, and 0 otherwise, and $T_{i,j}$ is an indicator for whether respondent j is provided an incentive to say she did not vote. Due to the smaller sample, only location-day fixed effects η_i are included in regressions.[19]

17. For space reasons, the specification in Table 1 assumes an additive effect between the flyer treatments, the payment and duration treatments and the door information treatments. The empirical moments used for the estimation, listed in Appendix Table A1, are more disaggregated.

18. We did not run the experiment in September 2011.

19. In Supplementary Appendix Table 4, we show that the results are not sensitive to adding the full set of fixed effects.

TABLE 1
Results for survey treatments

Specification	OLS regressions							
Dependent variable	Indicator for answering the door				Indicator for completing survey			
Group	Voters		Non-voters		Voters		Non-voters	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Constant	0.3458***		0.3206***		0.0909***		0.0457***	
	(0.017)		(0.017)		(0.012)		(0.010)	
$10/10 min treatment	0.0337**	0.0364**	0.0251	0.0243	0.0109	0.0132	0.0226***	0.0231***
	(0.016)	(0.015)	(0.016)	(0.015)	(0.010)	(0.010)	(0.008)	(0.009)
$10/5 min treatment	0.0515***	0.0596***	0.0227	0.0204	0.0602***	0.0683***	0.0465***	0.0467***
	(0.017)	(0.017)	(0.015)	(0.015)	(0.013)	(0.013)	(0.009)	(0.009)
Simple flyer treatments	0.0145	0.0128	0.0306	0.0286	0.0907***	0.0960***	0.0522***	0.0496***
	(0.018)	(0.018)	(0.019)	(0.018)	(0.013)	(0.013)	(0.010)	(0.010)
Flyer treatments with opt-out	−0.0195	−0.0232	0.0055	0.0052	0.0673***	0.0695***	0.0354***	0.0325***
	(0.019)	(0.019)	(0.018)	(0.018)	(0.013)	(0.013)	(0.010)	(0.010)
Mention of election in flyer	−0.0158	−0.0143	−0.0276**	−0.0278**	−0.0200*	−0.0194*	−0.0236***	−0.0238***
	(0.013)	(0.013)	(0.014)	(0.014)	(0.011)	(0.011)	(0.008)	(0.008)
Voters informed at door of election topic					−0.0024	0.0001	0.0042	0.0047
					(0.009)	(0.009)	(0.008)	(0.008)
Omitted treatment	No Flyer, $0/5 min treatment				No Flyer, $0/5 min, not informed treatment			
Fixed effects for solicitor, date-location, and hour		X		X		X		X
R^2	0.0032	0.0279	0.0018	0.0338	0.0116	0.0350	0.0080	0.0269
N	6,873	6,873	6,324	6,324	6,873	6,873	6,324	6,324

Notes: Estimates for a linear probability model with standard errors, clustered by solicitor-date, in parentheses. The table summarizes the result of three crossed treatments. The first is on duration and payment of the survey: $0/5min (the omited category), $10/10min, and $10/5min. The second is the flyer content: No flyer (the omitted category), Simple Flyer, Flyer with Opt-out. Each of the two flyer treatments is randomized into containing a mention of the election question in the flyer, or not. Hence, the coefficient on "Mention of Election in Flyer" captures the differential effect of the mention, compared to a simple flyer or a flyer with opt-out. Finally, we randomize whether at the door we announce the election question in the survey, with the omitted category being no mention. The regressions include fixed effects for the solicitor, for the date-town combination, and for the hour of day whenever indicated.
* Significant at 10%; ** Significant at 5%; *** Significant at 1%.

TABLE 2
Incentives to change reporting of voting status

Specification	OLS regressions			
Dependent variable	Indicator for lie (stated voting does not match official voting record)			
Group	Voters		Non-voters	
	(1)	(2)	(3)	(4)
Panel A. All survey respondents				
Constant	0.1210***		0.4677***	
	(0.014)		(0.031)	
Time or monetary incentive to say did not vote	0.0273	0.0225	−0.1204***	−0.1190***
	(0.020)	(0.019)	(0.040)	(0.040)
N	1,136	1,136	597	597
Panel B. \$0,5 min. respondents				
Constant	0.1479***		0.3971***	
	(0.028)		(0.061)	
5-Dollar incentive to say did not vote	−0.0302	−0.0394	−0.0918	−0.1105
	(0.034)	(0.037)	(0.075)	(0.076)
N	329	329	163	163
Panel C. \$10,5 min. respondents				
Constant	0.1280***		0.4623***	
	(0.024)		(0.046)	
5-Dollar incentive to say did not vote	0.0480	0.0550	−0.1452**	−0.1313**
	(0.036)	(0.034)	(0.059)	(0.065)
N	427	427	229	229
Panel D. \$10,10 min. respondents				
Constant	0.0909***		0.5281***	
	(0.022)		(0.053)	
8-Minute incentive to say did not vote	0.0561*	0.0487	−0.1143	−0.0864
	(0.031)	(0.034)	(0.069)	(0.071)
N	380	380	205	205
Omitted treatment	No incentive to say did not vote			
Fixed effects for location-day		X		X

Notes: Estimates for a linear probability model with standard errors, clustered by solicitor-date, in parentheses. The regressions include fixed effects for location-day in Columns 2 and 4.
* Significant at 10%; ** Significant at 5%; *** Significant at 1%.

In Table 2 and Figure 8, we present the results from these estimations. Recall that the incentive was always to say that one did not vote. Thus, we expect voters in the treatment condition to lie more than in the control, and non-voters to lie less. In Panel A of Table 2, to maximize power we pool across all survey treatments and across the 8-minute and \$5 incentive. Note first that non-voters, in the absence of any lying incentive, lie about 46% of the time about past turnout. This rate is within the range of previous results using the American National Election Studies and validated voter records (Silver *et al.*, 1986), and indicates that non-voters care about the social image that they convey. We also observe a 12% lying rate for voters, which could be explained by measurement error in the match to the voting records, or by a genuine preference among some voters to look like a non-voter.[20]

20. Notice that non-registered voters do not appear in our voting records. Hence, some of the households which we classify as "voting households" may include some non-voters, accounting for some of the lying rate for these households. In the Structural Estimates, we present results which allow for measurement error.

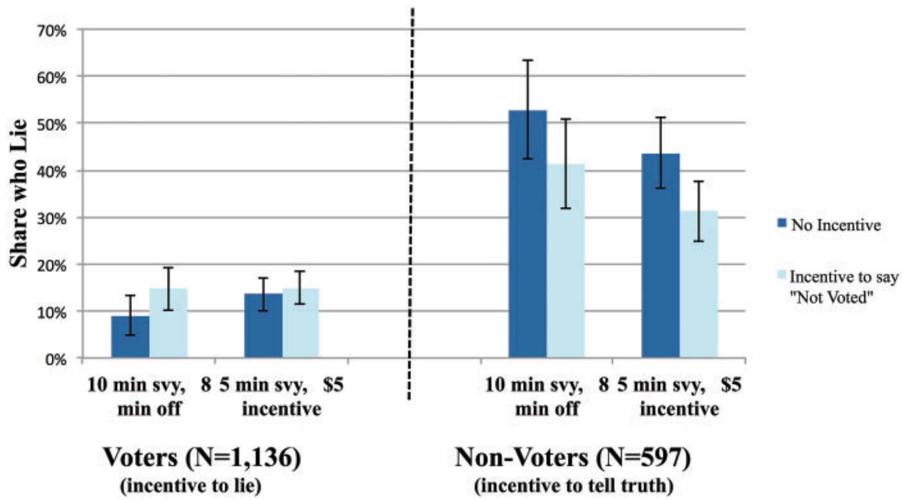


FIGURE 8

Response to lying incentives

Note: Figure 8 presents the share of households completing the survey whose response to the question "Did you vote in the 2010 congressional election?" differs from the official voting record, denoted as "Share who Lie". The shares lying are compared across treatments with an incentive to say that one did not vote to the treatments with no such incentive. The incentives are designed to induce voters to lie and non-voters to tell the truth. The averages are pooled across the different flyer and payment treatments featured in Figure 3. The sample sizes refer to the subsamples who answered the survey including the voting question. Standard errors are clustered at the solicitor-date level.

Turning to the effect of the incentives, the treatments have a small effect on voters: they lie 2.7 percentage points more when incentivized to do so, which is not statistically significant at conventional levels. For non-voters, in contrast, the effect is a highly significant 12 percentage point (25%) decrease in lying rates. Thus, voters appear to greatly dislike lying and claiming to be non-voters (relative to telling the truth), while non-voters are more easily moved between telling the truth and falsely claiming to be voters.

Do the results differ for the 8-minute time discount versus the $5 incentive? Figure 8 shows that the results are very similar for the two types of incentives, especially for non-voters, suggesting an implied value of time of about $35 per hour. Panels B–D in Table 2 further show that the results are similar whether the 5-minute survey was paid or unpaid.

4.3. *Summary*

To summarize, among voters we find little sorting on average into opening the door in the election flyer treatment, and therefore little evidence of pride in voting on average (though there is evidence among Republicans). But this does not imply that social image does not motivate their voting behaviour. In fact, even with substantial incentives of $5 earned or 8 minutes saved, over 85% of voters refuse to say they did not vote. This indicates that voters have a high lying cost L, a low social-image value of being a non-voter s^v_N, or both. Both these factors induce a high social-image value of voting. For non-voters, we find substantial sorting out in the election flyer treatment, indicating that non-voters experience stigma on average from not-voting. Further, close to half of non-voters lie and claim to be voters when asked. This implies that on average they are indifferent between the options: $s^{nv}_V - s^{nv}_N = L^{nv}$. A $5 incentive reduces lying by 25%, indicating that a substantial share of non-voters are close to the margin in their decision to tell the truth or lie. In the next section, we utilize all the experimental treatments to estimate the social-image value of voting.

5. STRUCTURAL ESTIMATES

5.1. *Set-up*

The key variables of interest—the social image variables s_V and s_N and the lying cost L—affect not only the decision whether to answer the survey and the reporting of whether one voted, but also the turnout decision itself. For example, individuals with high social-image utility (high $s_V - s_N$) are more likely to state that they voted, even if they did not. These same individuals are also more likely to choose to vote in the first place, as equation (1) clarifies.

The model estimation procedure acknowledges this dual role of social image, first in determining the voting status, and then in affecting the behaviour in the experiment. *Concretely*, we draw a simulated population of potential voters with values for social image, lying costs, and other parameters. Then, the individuals in the population choose to become voters or non-voters, depending on the draw of the variables. Accounting for the selection into voters and non-voters, we simulate the behaviour of both groups in the various experimental conditions: the share opening the door and doing the survey in a particular condition, the share lying conditional on answering the survey, etc. Hence, the distribution of the key variables of interest, such as social image, will differ for voters and non-voters in ways predicted by the selection model, even though they are initially drawn from the same population distribution.

An alternative approach to estimation, which we adopted in earlier versions, is to assume different distributions of the variables for voters and non-voters *without* modelling the selection into voting. This second approach has two disadvantages. First, this approach ignores a key prediction of the model, that voters and non-voters will differ in a particular way, for example with voters displaying higher social-image concerns. Second, given that it does not model the selection into voting, there is less power to identify the parameters and it is not possible to pin down either the lying cost L or the other reasons to vote ε. Thus, we adopt the estimation with selection as the main set of estimates, but also report below the estimates of the model with exogenous assignment into voters and non-voters.

The key variables in the model of "voting to tell others" are the social-image value of saying one voted, s_V, the social-image value of saying one did not vote, s_N, and the lying cost, L. We assume that the social-image variables s_V and s_N are independently normally distributed across individuals, with differing means μ_V and μ_N but the same standard deviation, $\sigma_V = \sigma_N$, which we denote by σ_{SI}. The distribution of s_V and s_N allows for individuals who prefer the social image associated with not voting ($s_V < s_N$). The lying cost L for tractability reason is assumed to be deterministic, though we relax this assumption below in a robustness check.

Together with the residual value of voting ε, these variables determine the decision to vote according to equation (1). The residual value of voting ε is normally distributed with mean μ_ε and standard deviation σ_ε, and independently drawn from the signalling utility variable s_V and s_N. Thus, there are six parameters of interest: (1) μ_V, the mean social-image utility from saying that one voted; (2) μ_N, the mean social-image utility of saying one did not vote; (3) σ_{SI}, the standard deviation of the social-image utilities; (4) L, the lying cost; (5) μ_ε, the mean of other reasons to vote, and (6) σ_ε, the standard deviation of the other reasons to vote. These parameters suffice to determine the conditions for selection into voters and non-voters.

While these are the relevant parameters for the voting decision, additional auxiliary parameters affect the decision to open the door and answer the survey, and whether to lie about voting. We assume a utility s of completing an unpaid 10-minute survey, distributed normally with parameters

μ_s and σ_s; the utility s is independent from s_V, s_N, and ε. We also assume a quadratic cost of changing plans to be at home, $c(h)=(h-h_0)^2/2\eta$ in case the person sees the flyer (which occurs with probability r). We thus have the following auxiliary parameters: (1) h_0, the baseline probability of opening the door; (2) r, the probability of observing (and remembering) the flyer; (3) η, the responsiveness of the probability of opening the door to the desirability of being at home; (4) μ_s and σ_s, the mean and standard deviation of the baseline utility of doing a survey; (5) v^s, the value of one hour of time[21]; (6) S_s, the social pressure cost associated with saying no to the survey request.

In our preferred specification, we allow voters to differ from non-voters in these auxiliary parameters.[22] In this specification, the total number of parameters is 20, with 6 key parameters and 7*2 auxiliary parameters. In an alternative, more parsimonious specification we force the auxiliary parameters to be constant across voters and non-voters, for a total of 13 parameters.

To estimate the model, we use a simulated method of moments estimator. Denote by $m_N(\xi)$ the vector of simulated moments as a function of the parameters ξ, and by $\hat{m}$ the vector of observed moments. The estimator chooses the parameters $\hat{\xi}$ that minimize the distance $\left(m_N(\xi)-\hat{m}\right)' W \left(m_N(\xi)-\hat{m}\right)$. As a weighting matrix W, we use the diagonal of the inverse of the variance–covariance matrix. Hence, the estimator minimizes the sum of squared distances, weighted by the inverse variance of each moment. For each run, we use a simulated population of at least 500,000 individuals. We discuss further details in Supplementary Appendix B.

To list the moments $m_N(\xi)$, we introduce the following indices: $i \in \{v, nv\}$ indicates voters and non-voters, $k \in \{NF, F, FE, OO, OOE\}$ indicates the flyer treatments, m indexes the payment and duration treatments, $m \in \{\$0, 5min; \$10, 10min; \$10, 5min\}$, a indicates the treatments on survey information at the door, $a \in \{I, NI\}$, and l indexes incentives to lie, $l \in \{NoInc, 8\text{min}, \$5\}$. The moments $m_N(\xi)$ are as follows: (1) the probability opening the door in treatments k, m, $P(H)^i_{k,m}$; (2) the probability of completing the survey in treatments k, m, $P(SV)^i_{k,m}$; (3) the probability of checking the opt-out box in the Opt-Out treatments, $P(OO)^i_{k,m}$ for $k \in \{OO, OOE\}$; (4) the probability of completing the survey in the survey content treatments, given the flyer treatments: $P(H)^i_{a,k}$; (5) the probability of lying about past turnout conditional on completing the survey, given incentive l, $P(L)^i_l$; and (6) the average turnout $P(V)$.[23] The empirical moments $\hat{m}$, 101 in total, are estimated in a first-stage model using the same controls as in the main regressions, and are listed in Appendix Table A1.[24]

5.2. *Identification*

Regarding the main sources of identification, consider first the social-image and lying parameters, μ_V, μ_N, σ_{SI}, and L. The difference in home presence and survey completion between the Flyer and Flyer-election, and between the Opt-Out and Opt-Out election treatments, plays an important role. For voters, they pin down the mean social-image utility $E^v[s_V]$, the expected value of s_V

21. Notice that the estimated value of time refers to minutes of doing a survey, which may differ from the value of time involved in deciding whether to vote.

22. This can be thought of as an (unmodeled) correlation between the auxiliary variables and ε.

23. The turnout moment is the turnout in the control group of the 2010 get-out-the-vote intervention, discussed in the next Section.

24. We present pooled moments across some of the treatments for two reasons. In some cases we do not expect any impact of the treatment on the relevant moment, such as of the lying incentives on the probability of opening the door or completing the survey. In other cases, we pool to keep the list of moments readable and to guarantee a sizeable sample in each cell, when the model does not imply important differences across the pooled treatments; for example, we do not consider the impact of the survey content treatment separately as a function of the survey duration and payment.

for individuals who select to become voters. For non-voters, given that on average half of non-voters lie in our sample (absent incentives to do otherwise), the average social-image utility from admitting to not voting, $E^{nv}[s_N]$, must approximately equal the utility from lying, $E^{nv}[s_V]-L$. A similar role is played by the difference in survey completion between the Information and No Information treatments.

Notice that, due to selection, the conditional expectations above do not equal the population mean. For example, individuals with high social-image utility of saying one voted, s_V, are more likely to choose to be voters; thus, $E^v[s_V]>\mu_V>E^{nv}[s_V]$. Thus the sorting moments, which pin down the conditional expectations, do not suffice to identify μ_V, μ_N, and L. The missing element is the extent of selection into voters and non-voters with respect to the social-image variables, which determines deviations such as $E^v[s_V]-\mu_V$. We return to the selection below.

The response to the lying incentives is crucial for identifying the heterogeneity in social image σ_{SI} and the average utility difference between answering truthfully and lying. For example, an 8-minute incentive reduces the share of non-voters lying by 12 percentage points (Table 2, Panel D), implying $\Pr\left(s_N<s_V-L<s_N+(8/60)v_s^{nv}|nv\right)=0.12$.

As we mentioned above, a critical piece for identification is the extent of selection on social-image variables s_V and s_N, versus on the residual voting variable ε. This selection will depend crucially on the comparison between the variance σ_ε^2 in the residual reasons to vote, versus the variance σ_{SI}^2. If the variance σ_ε^2 of the residual reasons to vote ε is large relative to the variance of social image terms, σ_{SI}^2, the selection into voters and non-voters will mostly depend on other reasons to vote ε, as opposed to variation in social image. In this case, indeed $E^v[s_V]\approx E^{nv}[s_V]\approx\mu_V$ and thus the parameters μ_V, μ_N, and L are identified by the sorting moments alone. Conversely, consider the case in which the variance σ_ε^2 is small relative to the variance σ_{SI}^2 in social-image utility. In this case, there will be large differences in social image s_V and s_N between voters and non-voters. Thus, the estimate of σ_ε^2 plays a critical role in tying together the parameters estimated for voters and non-voters.

This leads us to the identification of the parameters for residual voting ε, μ_ε and σ_ε. A critical moment is the baseline turnout rate, which is precisely measured at 0.60 in this population. For any given value of σ_ε, the value of μ_ε is identified by this moment (taking as given the average value of voting "to tell others"). What identifies then the second parameter, σ_ε? The lying moments play an important role in determining the relative variance of signal value σ_{SI}^2 versus the variance of the other reasons to vote σ_ε^2. With small σ_ε, selection into voting is primarily on the signal value, and therefore lying about voting by non-voters should be very limited. If all of the variance is in social-image, the only individuals who will become non-voters are the ones that actually prefer to say that they did not vote (*i.e.* have $s_V<s_N$) and thus would never lie about not voting. With large σ_ε, there is more selection on ε and therefore we expect to see more lying among voters and non-voters.

Given that lying about voting by non-voters is an established fact, this implies that σ_ε^2 cannot be too small. That being said, the estimates will have a hard time separating large values of σ_ε from very large values of σ_ε. Thus, an important robustness check is one in which we fix σ_ε (Table 5, Columns 6 and 7). As we see, the lying cost is tightly linked to the value of σ_ε: with higher σ_ε, the model requires a higher lying cost to match the lying moments.

As for the auxiliary parameters, the mean and standard deviation of the value of completing a survey, μ_s and σ_s, are identified from the survey completion rates for different monetary incentives. The value of time v_s is identified from the comparison between payment increases (from \$0 to \$10) and duration decreases (from 10 to 5 minutes), and partly also by the response to the 8 minute time saving offered in the lying incentive. The baseline probability of answering the door, h_0, is pinned down by the share opening the door in the no-flyer treatments, and less directly by the share opting out in the opt-out treatments, since respondents are predicted to opt out only if they expect to be home in the first place. The probability of observing and remembering the flyer, r, is mainly identified by the fraction of households checking the opt-out box in the Opt-out treatment (10–13%), which equals $rh_0F_s(c-m)$, and by the fraction opening the door in these treatments. The responsiveness of opening the door η with respect to incentives, and the social pressure S_s, are related to the share opening the door in the different survey treatments.[25]

5.3. *Benchmark estimates*

In Table 3, we present the estimates under two assumptions. In the first column, our benchmark specification, we allow for different auxiliary parameters for voters and non-voters, attaining a Weighted Sum of Squared Errors (SSE) of 160.3. In the second column, we force voters and non-voters to have the same auxiliary parameters, leading to a substantially worse fit, with an SSE of 355.7.

Under either set of estimates, the average individual does not take *pride* in saying that they voted: $\mu_V = -3.9$ (SE 1.5) in the benchmark and $\mu_V = -4.6$ (SE 1.7) in the alternative specification. This reflects the fact that even voters do not sort in to answer the election survey, compared to the non-election survey. We find strong evidence of *stigma* from not voting on average: $\mu_N = -11.3$ (SE 1.8) in the benchmark and $\mu_N = -17.2$ (SE 2.4) in the restrictive model. This reflects the fact that non-voters strongly avoid the election survey.

The benchmark estimates imply that individuals place an average value of over \$7 ($\mu_V - \mu_N = 7.4$) on being seen as a voter, as opposed to as a non-voter, each time they are asked. There is substantial heterogeneity in these social-image values ($\sigma_{SI} = 9.5$ in the benchmark), implying that 34% of individuals would in fact take pride in saying they voted.

Turning to the second key model component, we estimate a cost of lying about voter status of \$7.6 (SE 1.2) in the benchmark and \$13.3 (SE 1.3) in the more restrictive model. To put this in perspective, we estimate a lying cost of \$7 in a representative cheap talk laboratory experiment (Erat and Gneezy, 2012, see Supplementary Appendix C). In both settings, a person is lying to a stranger; this cost is likely to be larger in the context of repeated interactions. This significant lying cost, together with the sizable social-image utility, implies that being asked about voting becomes a reason to turnout. We return to this point below.

The distribution of the other reasons to vote ε is estimated to have a large standard deviation (albeit imprecisely estimated): $\sigma_\varepsilon = 318.7$ (SE 691.4) in the benchmark model.

The next panel shows the average value of the key variables—ε, s_V, and s_N—among the individuals who *select* to become voters and non-voters. (Since L is deterministic, it does not differ for the two groups.) The voters differ most from non-voters in the other reasons to vote, ε, with a mean of \$269 for voters and −\$242 for non-voters. Given the high σ_ε, most of the selection into becoming a voter is driven by factors other than social image.

Voters have a more positive social-image value of voting s_V and a more negative social-image value of non-voting, s_N, compared to non-voters. Thus, as expected, voters have larger social-image utility $s_V - s_N$, though the difference is relatively small: \$8.1 for voters versus \$6.2 for non-voters. Given these values, on average non-voters are close to indifferent between admitting they did not vote and claiming they voted ($E^{nv}[s_N] = -10.7$ versus $E^{nv}[s_V] - L = -12.1$), as implied by a lying rate among non-voters of about 50% in the control group.

25. Consider a respondent of type i who dislikes answering a survey and hence will say no and incur the social pressure cost S_s. In the flyer treatment F, she will choose to be at home with probability $h_0^i - \eta^i r^i S_s^i$ (barring corner solutions for h).

TABLE 3
Simulated minimum-distance estimates, benchmark results

	Voters and non-voters have different auxiliary parameters		Voters and non-voters have same auxiliary parameters	
Voting parameters	(1)		(2)	
Mean social image value of saying voted (μ_V)	−3.9 (1.47)		−4.6 (1.66)	
Mean social image value of saying did not vote (μ_N)	−11.3 (1.77)		−17.2 (2.37)	
Std. dev. of social image values (σ_{SI})	9.5 (1.29)		15.8 (2.34)	
Lying cost in $ (L)	7.6 (1.21)		13.3 (1.31)	
Mean value of other reasons to vote (μ_ε)	64.1 (167.90)		95.6 (109.26)	
Std. dev. of other reasons to vote (σ_ε)	318.7 (691.37)		499.9 (457.74)	
Value of parameters after selection into voting	Voter	Non-voter	Voter	Non-voter
Mean value of other reasons to vote (μ_ε)	268.6	−242.0	417.4	−382.3
Mean social image value of saying voted (μ_V)	−3.6	−4.5	−3.9	−5.6
Mean social image value of saying did not vote (μ_N)	−11.7	−10.7	−17.9	−16.2
Auxiliary parameters	Voter	Non-voter	Voter = non-noter	
Mean utility (in $) of doing 10-minute survey (μ_s)	−22.6 (2.82)	−27.7 (4.07)	−21.5 (1.51)	
Std. dev. of utility of doing survey (σ_s)	26.9 (5.22)	24.7 (4.87)	18.3 (2.18)	
Value of time of one-hour survey (v_s)	42.7 (8.50)	23.9 (12.22)	52.2 (8.99)	
Social pressure cost (in $) of declining survey ($S_s$)	1.6 (1.16)	1.2 (1.45)	0.5 (0.72)	
Responsiveness of probability of opening door (η)	0.14 (0.11)	0.16 (0.20)	0.42 (0.65)	
Probability of seeing the flyer (r)	0.38 (0.02)	0.30 (0.02)	0.34 (0.01)	
Baseline probability of opening door (h_0)	0.38 (0.01)	0.36 (0.01)	0.37 (0.01)	
SSE	160.3		355.7	

Notes: Estimates from simulated minimum-distance estimator using the moments in Appendix Table A1 with weights given by the inverse of the diagonal of the variance–covariance matrix. The sample consists of 6,873 voting households and 6,324 non-voting households. A [non-]voting household is a household in which all registered voters did [not] vote in the 2010 congressional election. Standard errors are in parentheses. SSE reports the Weighted Sum of Squared Errors.

Turning to the auxiliary parameters (bottom panel), we estimate that on average voters display a higher willingness to complete the survey compared to non-voters. (Voters are likely public good providers generally). The estimated time value is $43 per hour for voters and $24 for non-voters, a difference consistent with the strong positive correlation between income and turnout (Leighley and Nagler, 1992).[26] Voters and non-voters incur similar social pressure costs from declining to participate in the survey and have a similar responsiveness of the probability of opening the door.

26. The sizeable estimated value of time is consistent with the relatively high median household income of $92,000 on average across the towns visited.

5.4. *Value of voting to tell others*

Using the estimates, we compute the value of voting *to tell others*, averaged over the population: $N\int\Phi(s_V - s_N, L)\,dF(s_V, s_N)$. As Figure 1 shows, for positive values of social-image utility ($s_V - s_N > 0$), the value of voting *to tell others* Φ equals $\min(s_V - s_N, L)$, since non-voters can get the social-image utility $s_V - s_N$ by lying (and paying cost L). Thus, the value increases in both the social-image utility and the lying cost.

A key parameter for the value of voting *to tell others* is the expected number of times asked, N, given the assumption that the social image utility cumulates with each additional ask. We measure this parameter with survey questions on how often the survey respondents has been asked whether they voted in the 2010 congressional election by friends, relative, coworkers, and other people. The Supplementary Appendix Figure 1 displays the c.d.f. of the total number of times asked: 61% of respondents report being asked at least once, and 15% report being asked more than ten times. On average, respondents report being asked around 5.4 times for the 2010 congressional election, with similar magnitudes for voters and non-voters ($N^v = 5.12$ and $N^{nv} = 6.01$).[27] The figure also reports the number of times people report being asked for the 2008 presidential election: the average is about twice as high, with 38% of people reporting to be asked at least ten times. This number is consistent with the corresponding figures in the Cooperative Congressional Election Study as reported in Gerber *et al.* (2016).

The top row in Table 4 shows the implied value of voting for voters and non-voters for the 2010 congressional election. The magnitudes are sizable: in the benchmark model, we estimate a total value of voting to tell others of \$18 (SE 4.6) for voters and \$13 (SE 3.3) for non-voters, with larger estimates in the alternative specification. These magnitudes follow directly from the sizable estimates for the social-image utility and for the lying cost, as well as the fact that people expect to be asked on average five times.

We should point out three important caveats to this measure of the value of voting. First, as in any structural paper, the value depends on the parametric and distributional assumptions; below, we explore how this value varies for a wide set of sensitivity checks. Second, we estimate the social-image utility when asked by a stranger. If social-image concerns or lying costs when interacting with friends, family and colleagues are higher, our estimates are likely to be lower bounds of the social-image value of voting. Third, the estimates assume that the value of being asked is linear in the times asked. If utility is instead concave in times asked, we could be underestimating the value, since we conduct our experiment well after the election, when respondents report already having been asked a number of times.

Even taking these caveats into account, our ability to assign a dollar value to voting through the design is a unique contribution to the literature. For example, Coate and Conlin (2004) and Coate *et al.* (2008) estimate, respectively, a group-rule utilitarian model and a pivotal-voting model on alcohol-regulation referenda data. Their estimates for the value of voting are up to a scaling for the voting cost, which is not identified; thus, they do not provide a monetary value of voting due to their model. The unique element in our design that makes the difference is that we use monetary inducements—variation in the value of the survey and incentive to lie about voting—to translate the findings into a monetary value of voting.

A second metric to evaluate the model of voting "to tell others" is in terms of the extra turnout that it generates, relative to the baseline turnout, 60% in our setting. How would turnout change

27. Respondents were asked to report the number of times asked about the 2010 election by friends, relatives, coworkers, and other people. Total number of times asked is the sum across these categories, each capped at twenty times asked. Similarly, number of times asked about the 2008 election is the number of times asked by friends and relatives, each capped at twenty times asked.

TABLE 4
Implied value of voting and welfare effects of get-out-the-vote intervention

	Voters and non-voters have different auxiliary parameters		Voters and non-voters have same auxiliary parameters	
	(1)		(2)	
Implications for value of voting to tell others	Voter	Non-voter	Voter	Non-voter
Implied value of voting "to tell others" ($N=5.4$) (in \$)	18.3 (4.6)	13.3 (3.3)	33.1 (4.1)	23.3 (2.6)
Baseline observed turnout in sample	0.599 (0.011)		0.599 (0.011)	
Implied change in turnout if never asked about voting ($N=0$)	−0.019 (0.0031)		−0.022 (0.0178)	
Implied change in turnout if asked about voting twice as often ($N=10.8$)	+0.018 (0.0079)		+0.021 (0.0038)	
Implications for GOTV	Voter	Non-voter	Voter	Non-voter
Utility from being asked once whether one voted (in \$)	−2.8 (1.2)	−5.9 (1.5)	−2.7 (1.4)	−8.3 (1.9)
Implied GOTV effect from asking one more time ($N+1$)	+0.003 (0.0005)		+0.004 (0.0004)	
Implied number of subjects targeted with GOTV intervention to get one additional vote ($N+1$)	295 (84.9)		264 (80.6)	
Utility cost to get one additional vote ($N+1$) (in \$)	1189 (2684.4)		1304 (406.8)	

Notes: Derived from the benchmark model estimates. The implied number of GOTV subjects to get one additional vote assumes that there is no targeting in the intervention (*i.e.* voters and non-voters are included). Standard errors are in parentheses.

if people stopped asking others whether they voted (*i.e.* $N=0$)? What if conversely the rate of asking doubled (*i.e.* $2N$ asks)? We can address these questions because we are able to estimate the distribution of other reasons to vote ε.

Table 4 and Figure 9 display the results. With the benchmark estimates, a shift in norms that were to make it undesirable to ask others about voting would lower turnout by 1.9 percentage points. A doubling in the number of asks, perhaps encouraged by political campaigns, would increase turnout by 1.8 points.[28] While these impacts may seem small, consider the extraordinary effort that campaigns put into get-out-the-vote (GOTV) efforts, with the average such letter yielding a turnout impact of 0.2 points (Green *et al.*, 2013).

5.5. *Get-out-the-vote and welfare*

The estimates allow us also to calculate the implied turnout and welfare effects of a GOTV intervention based on informing potential voters that they will be asked whether they voted. This is a GOTV intervention which we designed ourselves and which we evaluate in Section 6. In addition, a related GOTV message, independently designed, was used in the 2012 presidential

28. The number of times asked can also differ by type of election: in fact, our survey evidence indicates that in the 2008 presidential election people were asked twice as much. Our calculation would imply that just the additional asking (holding the other parameters constant) could contribute 1.8 to 2.1 percentage points.

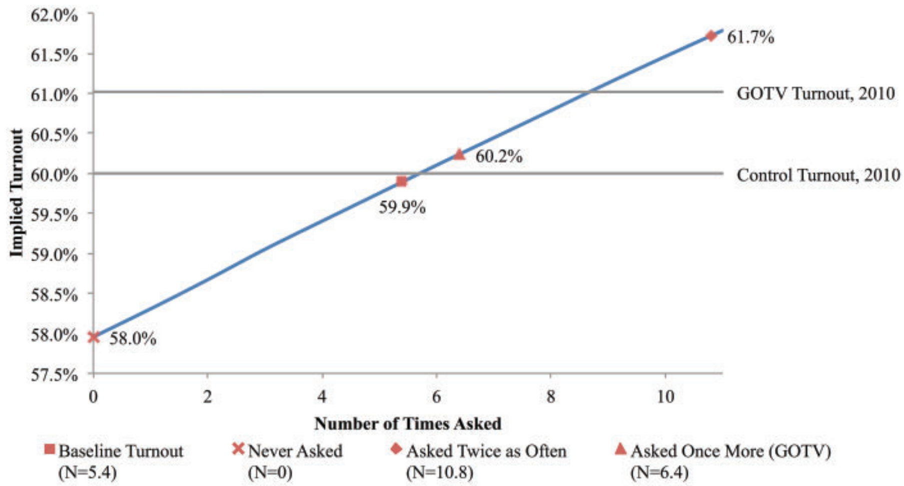


FIGURE 9

Voter turnout as function of times asked about voting

Note: Figure 9 plots the implied voter turnout at the benchmark estimates of the parameters, as a function of the number of times potential voters expect to be asked about voting. The horizontal lines represent actual turnout in the control and get-out-the-vote treatment groups in the Congressional election of November 2010.

election and is evaluated in Rogers *et al.* (2016). We assume that this intervention increases the number of times asked to $N+1$.[29]

First, we calculate the welfare effect of this GOTV intervention. For voters, the average value of being asked about voting, $z^v = \max\left(s_V^v, s_N^v - L^v\right)$, is estimated to be -2.8 (SE 1.2) in the benchmark model. Non-voters are estimated to have a more negative utility from being asked, -5.9 (SE 1.5). These sizable negative welfare effects occur because on average individuals do not derive pride from voting, and they strongly dislike admitting that they did not vote.

Second, we predict the effectiveness of the intervention. Given the distribution of the other reasons to vote ε, this GOTV intervention is predicted to increase turnout by 0.3 percentage points (Table 4 and Figure 9). Thus, to get one extra vote with this intervention, 295 people would have to be contacted.

Taking these two numbers together, we estimate that this GOTV intervention would result in a utility cost of \$1189 *per additional vote yielded*. This estimated magnitude is similar in the restrictive model, and implies a disadvantageous cost-benefit analysis, once one takes into account the voter welfare. This welfare disutility dwarfs the cost estimates in the current literature, which typically include just the postal costs of sending the GOTV material (*e.g.* Rogers *et al.*, 2016). As far as we know, ours is the first welfare evaluation of a get-out-the-vote intervention, an area of vast growth in the political science literature.

5.6. *Robustness*

In Table 5, we explore the robustness of the parameter estimates to alternative assumptions. We maintain the assumption as in the benchmark estimates (reproduced in Column 1) that the auxiliary

29. The impact could be smaller if individuals expect to not be reachable at this follow-up time, or if they miss the flyer (or mailer).

parameters can differ for voters and non-voters. Supplementary Appendix Table 6 reports the parallel results under the assumption of identical auxiliary parameters.

Before we go into the details, we note some common features. A variety of alternative specifications (with one exception) yield an essentially identical fit of the model (measured as SSE), implying that the alternative models do not affect the ability of the model to explain the data. (This is not true any more if we force the auxiliary parameters to be the same for voters and non-voters, see Supplementary Appendix Table 6.) Also, while most parameter estimates are quite stable, the standard deviation of the other reasons to vote, σ_ε, varies substantially. This is not surprising given that this is the least precisely estimated parameter. This does not much impact the dollar value of voting "to tell others", but it affects the implied percentage point impact on turnout, which is therefore more variable across estimates.

First, we remove the assumption, maintained so far, that lying costs L are deterministic and identical for voters and non-voters. We allow lying costs to have an exponential distribution; this one-parameter distribution ensures that lying costs are always positive. In this specification (Column 2), individuals select into voters and non-voters partly based on the realized lying cost. We estimate the same mean lying cost of \$8 as in the benchmark, with little selection into voting by lying cost (mean lying cost of \$8.4 for voters and \$7.3 for non-voters). We estimate a lower standard deviation of ε than in the benchmark model; as a consequence, there is more selection into voting based on signalling value and a larger implied effect on turnout if potential voters expected to be asked twice as often.

Next, we consider an alternative explanation of the results (Column 3): the sorting out of non-voters may be due to a dislike of talking about politics, rather than any stigma from admitting to not voting. We allow for a utility of talking about politics which is independent of whether one voted or not. With this extra parameter, we lose the ability to estimate a social-image parameter, so we fix μ_V to zero. The estimates of the key parameters of the value of voting are largely unaffected, reflecting the key identifying role of the lying moments, which are unaffected by the introduction of a utility of talking about politics.

We then consider measurement error in the voter records. Notice that the voting records do not include information about non-registered adults in a household. Since these individuals are necessarily non-voters, the person answering the door in an apparent voting household may actually be a non-voter. (This would explain why 10% of voting households appear to lie about voting even absent incentives to lie.) In Column 4, we assume that 10% of respondents in voting households are actually non-voters. Allowing for measurement error leads to similar results as in the benchmark case, with larger turnout impacts of doubling the number of asks. The results are similar if we assume a higher measurement rate of 20% (Column 2 of Supplementary Appendix Table 5) or we allow for symmetric measurement error, and assume that 10% of respondents in a voting (or respectively, non-voting) household are non-voters (respectively, voters) (Column 3 of Supplementary Appendix Table 5).

We also test the robustness of the estimates to the assumed number of times potential voters expect to be asked about voting (Column 5). In case the survey results overstate the number of times asked, we re-estimate the model assuming half as many asks ($N=2.7$). The model estimates are similar, with, not surprisingly, a lower monetary value of voting "to tell others" (since people are asked less). We find similar results if we assume voters expect to be asked twice as often (Supplementary Appendix Table 5, Column 4).

Finally, we explore the role of an important parameter that is not precisely estimated, the standard deviation of the other reasons to vote, σ_ε. To demonstrate its role, we fix it to a low level ($\sigma_\varepsilon=10$, Column 6) and to a high level ($\sigma_\varepsilon=1000$, Column 7). For low σ_ε, the model has a harder time matching the moments and the model SSE increases from 160 to 172. With low σ_ε, the selection is mostly on the social-image terms, making it harder to match the lying rate.

TABLE 5
Simulated minimum-distance estimates, robustness results

	Benchmark	Heterogeneous lying cost	Utility from talking about politics	10% Voters mismeasured	Assume asked about voting half as often	Low std. dev. of other reasons to vote	High std. dev. of other reasons to vote	Fixed std. dev. of other reasons to vote
Voting parameters	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Mean social image value of saying voted (μ_V)	−3.9	−4.3	0.0	−3.4	−3.7	−7.3	−3.5	−3.7
	(1.47)	(1.43)	(assumed)	(1.50)	(1.50)	(2.00)	(1.26)	(1.04)
Mean social image value of saying did not vote (μ_N)	−11.3	−11.3	−7.0	−11.3	−11.1	−15.7	−10.9	−11.0
	(1.77)	(1.77)	(1.17)	(1.61)	(2.37)	(3.50)	(1.74)	(1.46)
Std. dev. of social image values (σ_{SI})	9.5	8.5	8.8	8.7	9.5	11.2	9.7	9.5
	(1.29)	(1.23)	(1.94)	(1.03)	(2.35)	(3.20)	(1.76)	(1.59)
Lying cost in \$ (L)	7.6		6.7	7.3	7.4	1.3	8.4	7.7
	(1.21)		(1.27)	(0.50)	(1.08)	(0.07)	(1.35)	(1.12)
Mean value of other reasons to vote (μ_ε)	64.1	10.4	32.8	24.9	22.7	−0.3	235.8	64.4
	(167.90)	(11.29)	(27.28)	(143.47)	(23.26)	(0.45)	(28.71)	(9.19)
Std. dev. of other reasons to vote (σ_ε)	318.7	93.0	184.7	166.7	121.2	10.0	1000.0	318.7
	(691.37)	(50.75)	(110.03)	(520.21)	(103.75)	(assumed)	(assumed)	(assumed)
Mean lying cost L (in \$)		8.0						
		(1.63)						
Utility from talking about politics for voters			−4.7					
			(1.82)					
Utility from talking about politics for non-voters			−2.3					
			(1.80)					

(*continued*)

TABLE 5
Continued

	Benchmark		Heterogeneous lying cost		Utility from talking about politics		10% Voters mismeasured		Assume asked about voting half as often		Low std. dev. of other reasons to vote		High std. dev. of other reasons to vote		Fixed std. dev. of other reasons to vote	
Voting parameters	(1)		(2)		(3)		(4)		(5)		(6)		(7)		(8)	
Implications for value of voting and GOTV	Voter	Non-voter	Voter	Non-voter	Voter	Non-voter	Voter	Non-voter	Voter	Non-voter	Voter	Non-voter	Voter	Non-voter	Voter	Non-voter
Implied value of voting "to tell others" (benchmark $N=5.4$)	18.3 (4.6)	13.3 (3.3)	22.1 (2.8)	4.3 (4.5)	17.5 (3.6)	10.7 (3.6)	21.1 (8.4)	13.0 (18.1)	9.3 (2.3)	6.1 (1.7)	5.4 (0.4)	−0.8 (0.3)	18.6 (3.4)	16.7 (2.9)	18.4 (3.1)	13.3 (1.9)
Implied change in turnout if asked about voting twice as often	+0.018 (0.008)		+0.036 (0.007)		+0.027 (0.005)		+0.037 (0.006)		+0.023 (0.004)		+0.069 (0.010)		+0.007 (0.002)		+0.018 (0.004)	
Utility from being asked about voting once	−2.8 (1.2)	−5.9 (1.5)	−2.7 (1.3)	−5.2 (1.5)	−3.5 (1.8)	−4.0 (2.0)	−2.3 (1.6)	−5.8 (1.9)	−2.4 (1.3)	−5.5 (1.8)	−3.9 (1.3)	−5.7 (2.8)	−2.6 (1.2)	−5.9 (1.5)	−2.5 (1.0)	−5.6 (1.3)
Implied GOTV effect ($N+1$)	+0.003 (0.001)		+0.008 (0.004)		+0.005 (0.002)		+0.007 (0.003)		+0.009 (0.004)		+0.016 (0.002)		+0.001 (0.0004)		+0.003 (0.001)	
Implied number of GOTV subjects to get one additional vote ($N+1$)	295 (84.9)		123 (63.8)		184 (110.9)		135 (182.9)		113 (103.0)		62 (8.0)		818 (79.9)		298 (17.1)	
Utility cost to get one additional vote ($N+1$)	1189 (2684.4)		452 (166.6)		679 (439.4)		503 (1855.6)		412 (436.0)		285 (129.1)		3178 (566.6)		1123 (109.6)	
SSE	160.3		159.1		159.6		158.9		160.0		171.7		160.9		160.3	

Notes: Estimates from simulated minimum-distance estimator using the moments in Appendix Table A1 with weights given by the inverse of the diagonal of the variance–covariance matrix. The sample consists of 6,873 voting households and 6,324 non-voting households. A [non-]voting household is a household in which all registered voters did [not] vote in the 2010 congressional election. Standard errors are in parentheses. SSE reports the Weighted Sum of Squared Errors.

For high σ_ε, instead, the estimates are similar to the benchmark, with a lower implied turnout impact of a doubling in the asks, since fewer people are on the margin of the turnout decision. This highlights the one-sided nature of identification of σ_ε: the data reject small values of σ_ε, but has a hard time telling apart sizable to large values of σ_ε.

Because σ_ε is not precisely estimated, estimating this parameter introduces uncertainty that propagates through to many of the implications for the value of voting. For this reason, we also estimate the model fixing $\sigma_\varepsilon = 318.7$, the value estimated in the benchmark model (Column 8). Fixing σ_ε at the benchmark value, we estimate nearly identical parameters and model implications, but smaller standard errors. Resolving the uncertainty around σ_ε considerably lowers the standard errors on the number of people needed to get one more vote through a get-out-the-vote intervention and the associated utility cost.

We consider a number of other robustness checks in Supplementary Appendix Table 5, some of which we discussed above. We obtain similar estimates if we do not drop the observations with no-solicitor sign or other issues (Column 5) or if we include as an additional moment the turnout of the treated population in the get-out-the-vote experiment (Column 6).[30] Estimates excluding the moments split by whether households are informed at the door about the election topic (Column 7) indicate a larger social-image utility, not surprisingly, since the excluded moments reveal no differential impact of the social-image manipulation. Estimating the model without the lying incentive moments (Column 8), the heterogeneity in signal parameters is not identified, so we fix $\sigma_{SI} = 10$, as in the benchmark estimates. The estimated signalling parameters μ_V and μ_N are similar to the benchmark, but the lying cost is much less precisely estimated, revealing the key role of the lying incentive treatments.

5.7. *Estimation with exogenous voter status*

In Supplementary Appendix Table 7, we present the results of an alternative estimation method used in previous versions of the paper (DellaVigna *et al.*, 2014), which does not model the decision to become a voter or non-voter and instead allows for separate parameters for voters and non-voters ($\mu_V^i, \mu_N^i, \sigma_{SI}^i$ for $i \in \{v, n\}$). Without modelling the turnout decision, we are not able to estimate the distribution of other reasons to vote, ε, and we cannot separately identify the lying cost L. Instead, we estimate of the mean value of saying one voted ($\mu_v = -6$ for voters and $\mu_v - L = -8$ for non-voters) and saying that one did not vote ($\mu_n - L = -28$ for voters and $\mu_n = -7$ for non-voters). The value of voting to tell others is only defined as a function of a given (assumed) value of lying cost. For a lying cost of \$5, this value is \$12 for voters and \$18 for non-voters.

6. GET-OUT-THE-VOTE EXPERIMENT

The experiments described are designed to measure the value of voting without affecting voting itself. Yet, the model suggests a natural treatment to increase voter turnout. As Proposition 4 states, individuals with social-image motives are more likely to vote the more frequently they expect to be asked about voting, an expectation which we can manipulate experimentally.

30. To calculate the simulated moment, we assume that potential voters in the GOTV treatment group expect to be asked on average h_0^i times more than they would absent the treatment, where $i \in \{V, NV\}$ is determined by voting behaviour absent the GOTV treatment.

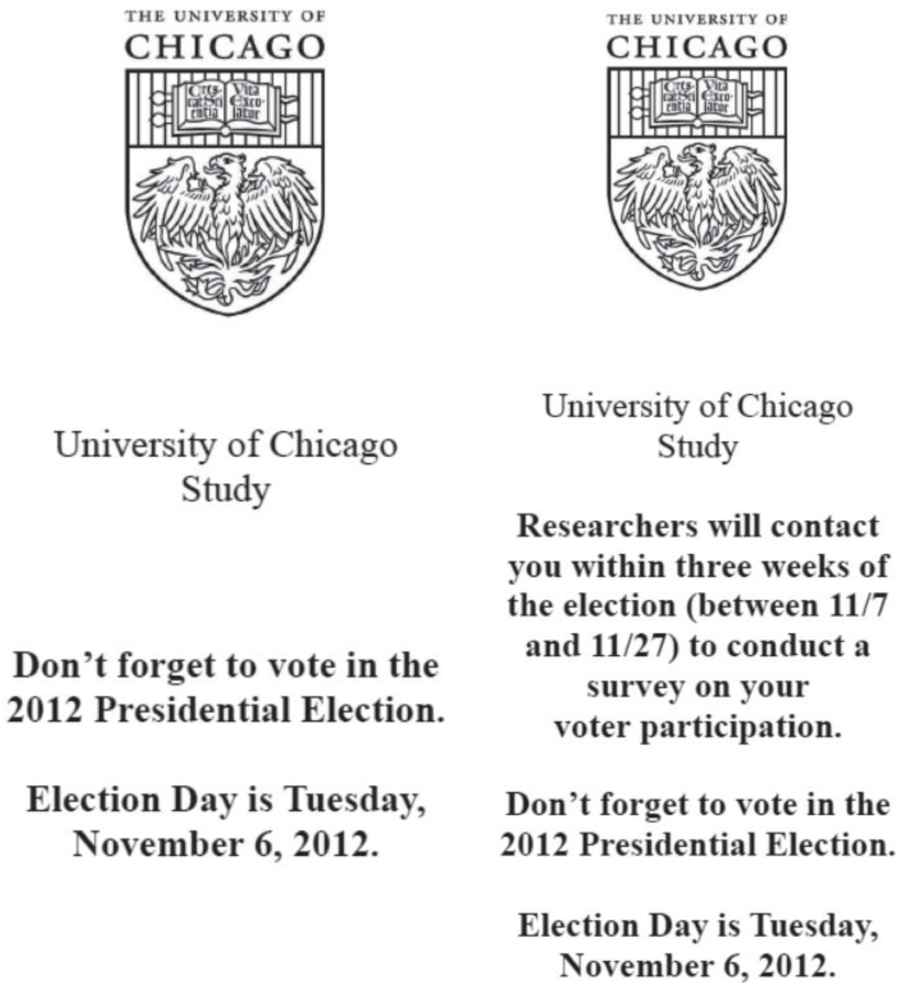


FIGURE 10

Flyer samples for GOTV treatment

Note: Figure 10 shows the door-knob flyers used in the Get-Out-The-Vote treatments in the days before the 2012 presidential election. The left flyer is for the treatment with Voting Reminder, the right flyer is for the treatment with Announcement Will Ask About Voting. Flyers for the 2010 election are similarly styled.

In November of 2010 and of 2012, we did just that in the suburbs of Chicago. In the 5 days before the election date, we posted a flyer on the doorknob of households in the treatment group informing them that "researchers will contact you within three weeks of the election [...] to conduct a survey on your voter participation". Figure 10 shows the flyer for the 2012 election. (After the election, we follow up with a door-to-door visit, as advertised). Since this flyer could also impact turnout through a reminder effect, we compare this group to a group which received a flyer with a mere reminder of the upcoming election, also displayed in Figure 10. A control group received no flyer. After the election, we obtain the voting record for all individuals residing at the addresses targeted in this experiment.

Table 6 reports the results for both the November 2010 and the November 2012 intervention using an OLS specification: the dependent variable is an indicator for whether the individual

TABLE 6
Results for get-out-the-vote treatments

Specification	OLS regressions			
Dependent variable	Indicator for voting in election in year t			
Election	Congressional elections in Nov. 2010		Presidential elections in Nov. 2012	
	(1)	(2)	(3)	(4)
Constant	0.6000***		0.7312***	
	(0.0109)		(0.0033)	
Flyer with voting reminder	−0.0020	−0.0031	0.0060	0.0046
	(0.0152)	(0.0083)	(0.0056)	(0.0034)
Flyer with announcement will ask about voting	0.0120	0.0102	0.0023	0.0056
	(0.0157)	(0.0084)	(0.0056)	(0.0034)
Omitted treatment	No Flyer		No Flyer	
Control for past voting since 2004		X		X
Difference (flyer will ask - flyer reminder)	0.0140	0.0133	−0.0037	0.0010
p-value for test of equality, 2-sided	$p=0.365$	$p=0.120$	$p=0.561$	$p=0.811$
p-value for test of equality, 1-sided	$p=0.182$	$p=0.060$*		$p=0.405$
R^2	0.0001	0.4024	0.0000	0.3251
N	$N=31,306$	$N=31,304$	$N=93,805$	$N=93,805$

Notes: Estimates for a linear probability model with standard errors, clustered by flyering route, in parentheses. The omitted treatment is the No-Flyer condition. The regressions in Columns 2 and 4 include indicators for participation in the following elections: March 2004 (primary), Nov. 2004, Feb. 2005, March 2006 (primary), Nov. 2006, April 2007, Feb. 2008 (primary), Nov. 2008, April 2009, and Feb. 2010 (primary). In addition, Column 4 includes indicators for participation in the Nov. 2010, April 2011, and March 2012 (primary) elections.
* Significant at 10%; ** Significant at 5%; *** Significant at 1%.

voted in the specific election. Note that there may be multiple individuals at one address, each of which is a separate observation. The November 2010 experiment has a sample size of 31,306 individuals targeted (11,462 received no flyer, 10,805 received a control flyer, and 9,039 received a treatment flyer). The turnout in the control group (which received no flyers) is 60.0 percentage points. Compared to this control group, the mere reminder had no effect, leading to an estimated decrease of 0.2 percentage points. Compared to the flyer with a mere reminder, the flyer with announcement of future question about voting raises turnout by 1.4 percentage points, a sizeable effect, albeit statistically insignificant. In Column 2, we add controls for the full history of voting of the households in all elections between 2004 and the election in question. Adding controls in a randomized experiment should not affect the point estimates if the experiment is conducted properly, but can reduce the residual variance, and hence increase precision. Indeed, the controls have very little impact on the point estimates, but they nearly halve the standard errors since past voting is highly predictive of future voting (the R^2 increases from 0.00 to 0.40). In this specification, the estimated effect of the flyer with announcement of future asking is an extra 1.3 percentage points in turnout, with a two-sided p-value of 0.12 (one-sided p-value of 0.06). While not quite statistically significant, the sizeable effect is certainly consistent with the predictions of the model.

Columns 3 and 4 display the estimates for the November 2012 election. In this later election, we were able to deploy a larger flyering team, guaranteeing a sample size of 93,805 individuals (46,868 received no flyer, 23,501 received a control flyer, and 23,436 received a treatment flyer). Given the different nature of the election (presidential versus congressional), the baseline turnout in the control group is higher, at 73.1 percentage points. We find suggestive evidence that the reminder flyer itself may have increased turnout, with little evidence of a differential effect of the flyer with announcement of the future visit. In the specification with controls (Column 4), the differential effect is estimated to be 0.1 percentage points, not significant. The smaller effect in this second election is consistent with the fact that in this higher-stake election our intervention is competing with a high number of campaign materials; in addition, the higher baseline turnout leaves a smaller share of non-voters to be potentially convinced.

An important question is whether these estimated effects are consistent with the estimated value of voting. As described in Section 5, we can compute the predicted increase in turnout due to an increase in N, the number of times asked about voting, for the estimated parameters. Assuming the treatment flyer increases the expected number of asks by 1, the benchmark model predicts an increase in turnout of 0.3 percentage points (Table 4 and Figure 9). This prediction matches the order of magnitude of the findings: it is smaller than the observed effect for the 2010 election (1.3 percentage points), but larger than the effect for the 2012 Presidential elections (0.1 percentage points), and within the confidence interval of both estimates.[31]

The results are consistent with the contemporaneous and independent results of Rogers *et al.* (2016) who similarly inform a treatment group that they may be called after the election about their voting behaviour. They also find a positive impact on turnout, of similar magnitude (0.2 percentage points). Importantly, it is intuitive that the effect sizes are much smaller than those in Gerber *et al.* (2008). Their intervention is conducted in a non-competitive primary election, and explicitly threatens to truthfully reveal one's voting record to the entire neighbourhood, while also providing information about social norms (other's voting records) and a strong civic duty exhortation. In contrast, we operate in competitive elections, do not provide civic duty messaging or information on neighbours, and warn of at most one additional question about voting, from a researcher who can be avoided or lied to.

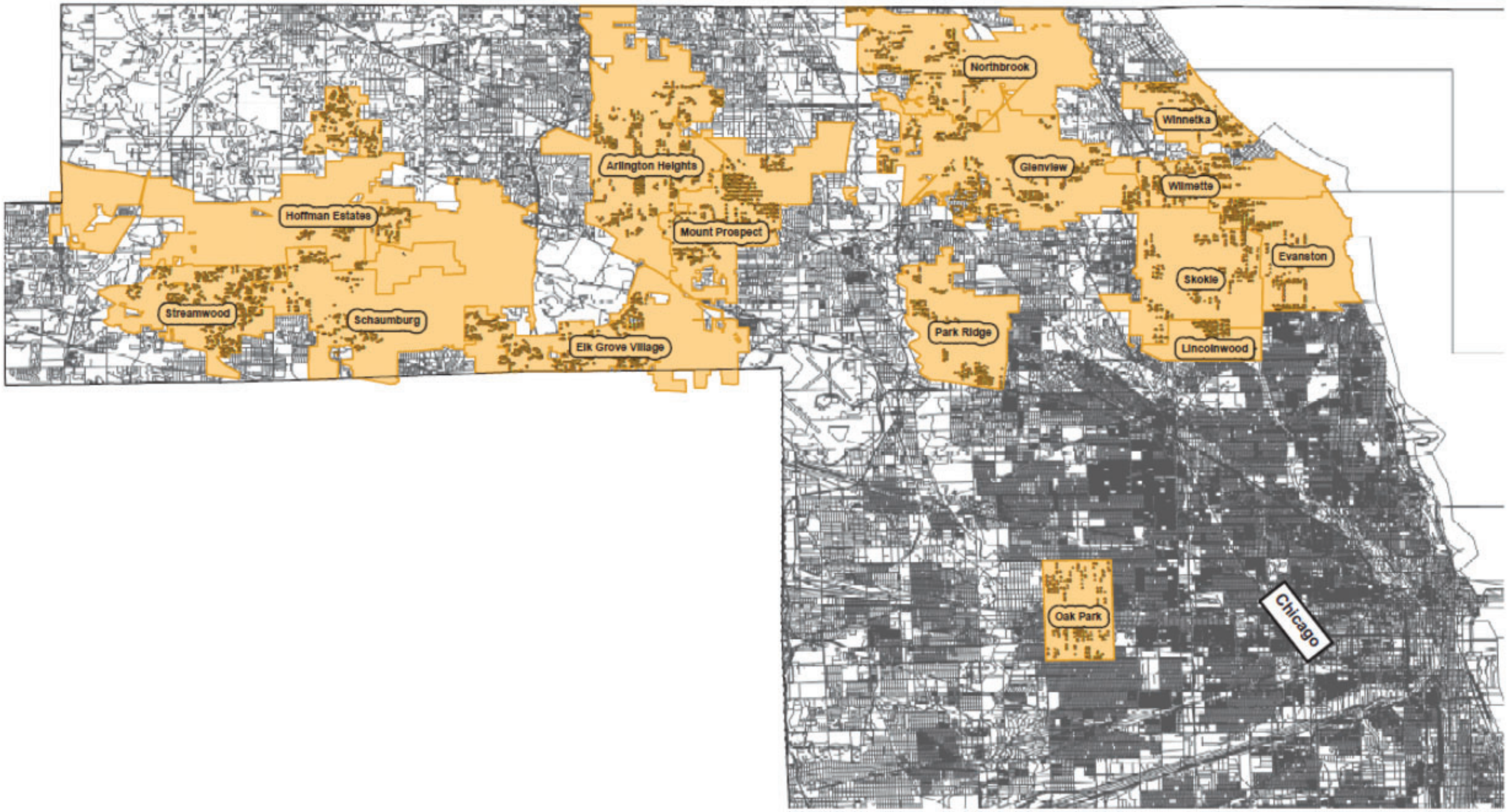
7. CONCLUSION

We have presented evidence from a natural field experiment designed to estimate a social-image model of voting: individuals vote because they expect to be asked, and they anticipate the disutility associated with admitting to not voting, or with lying about voting. The results document substantial shame from admitting to not voting, though little evidence of pride from conversely claiming to vote.

The experimental design allows us to estimate the key social image and lying parameters. We estimate a value of voting simply due to being asked of \$18 for voters and \$13 for non-voters, a sizeable magnitude for a congressional election. These constitute among the first estimates of the private value of voting. We also use the estimates to do a welfare evaluation of a get-out-the-vote intervention, and find that the disutility associated with the intervention is likely to dwarf the administrative cost. This underscores the importance of model-based evaluations of GOTV interventions.

A methodological ingredient of this article is the tight link between a simple model and the experimental design. As such, this article attempts to bridge a gap between two thriving, but largely separate literatures: the theoretical literature on voting and on social image, and the reduced-form field experiments on get-out-the-vote and turnout. We hope that methodologies similar to the ones in this article will be useful in providing further insights.

31. The prediction for the 2012 results should be taken with extra caution given that the social-image and other parameters are likely to vary by election.



APPENDIX FIGURE A1

Area surveyed

Appendix Figure A1 displays in colour (light grey in black/white) the towns visited in Cook County, Ill., as part of the door-to-door field experiment. The locations of the households visited within the towns are displayed in darker colour. For reference, Chicago, which we did not visit, is located on the bottom-right corner.

Appendix Table A1

Empirical moments and predicted moments at benchmark estimates

Voters						Non-voters					
	Empirical Mom.	Predicted Mom.		Empirical Mom.	Predicted Mom.		Empirical Mom.	Predicted Mom.		Empirical Mom.	Predicted Mom.
Flyers: Answering door	P(Answer)		Flyers: Survey Compl.	P(Do Survey)		Flyers: Answering door	P(Answer)		Flyers: Survey Compl.	P(Do Survey)	
No Flyer, $0, 5 min	0.350	0.378	No Flyer, $0, 5 min	0.115	0.094	No Flyer, $0, 5 min	0.367	0.357	No Flyer, $0, 5 min	0.085	0.050
No Flyer, $10, 10 min	0.408	0.378	No Flyer, $10, 10 min	0.100	0.123	No Flyer, $10,10 min	0.330	0.357	No Flyer, $10, 10 min	0.048	0.079
No Flyer, $10, 5 min	0.383	0.378	No Flyer, $10, 5 min	0.129	0.141	No Flyer, $10, 5 min	0.305	0.357	No Flyer, $10, 5 min	0.081	0.088
Survey Flyer, $0, 5 min	0.349	0.368	Survey Flyer, $0, 5 min	0.163	0.139	Survey Flyer, $0, 5 min	0.369	0.335	Survey Flyer, $0, 5 min	0.092	0.072
Survey Flyer, $10, 10 min	0.389	0.392	Survey Flyer, $10, 10 min	0.181	0.186	Survey Flyer, $10, 10 min	0.353	0.355	Survey Flyer, $10, 10 min	0.117	0.115
Survey Flyer, $10, 5 min	0.407	0.407	Survey Flyer, $10, 5 min	0.256	0.215	Survey Flyer, $10, 5 min	0.413	0.361	Survey Flyer, $10, 5 min	0.172	0.128
Election Flyer, $0, 5 min	0.369	0.362	Election Flyer, $0, 5 min	0.168	0.138	Election Flyer, $0, 5 min	0.305	0.326	Election Flyer, $0, 5 min	0.072	0.067
Election Flyer, $10, 10 min	0.370	0.384	Election Flyer, $10, 10 min	0.171	0.184	Election Flyer, $10, 10 min	0.333	0.342	Election Flyer, $10, 10 min	0.106	0.107
Election Flyer, $10, 5 min	0.423	0.398	Election Flyer, $10, 5 min	0.233	0.211	Election Flyer, $10, 5 min	0.325	0.347	Election Flyer, $10, 5 min	0.117	0.119
Survey Opt Out, $0, 5 min	0.359	0.320	Survey Opt Out, $0, 5 min	0.156	0.137	Survey Opt Out, $0, 5 min	0.287	0.291	Survey Opt Out, $0, 5 min	0.065	0.071
Survey Opt Out, $10, 10 min	0.355	0.349	Survey Opt Out, $10, 10 min	0.172	0.184	Survey Opt Out, $10, 10 min	0.357	0.316	Survey Opt Out, $10, 10 min	0.115	0.114
Survey Opt Out, $10, 5 min	0.382	0.367	Survey Opt Out, $10, 5 min	0.226	0.212	Survey Opt Out, $10, 5 min	0.321	0.323	Survey Opt Out, $10, 5 min	0.125	0.127
Election Opt Out, $0, 5 min	0.264	0.313	Election Opt Out, $0, 5 min	0.109	0.136	Election Opt Out, $0, 5 min	0.278	0.281	Election Opt Out, $0, 5 min	0.043	0.066
Election Opt Out, $10, 10 min	0.346	0.339	Election Opt Out, $10, 10 min	0.152	0.182	Election Opt Out, $10, 10 min	0.351	0.300	Election Opt Out, $10, 10 min	0.095	0.106
Election Opt Out, $10, 5 min	0.392	0.356	Election Opt Out, $10, 5 min	0.203	0.209	Election Opt Out, $10, 5 min	0.346	0.306	Election Opt Out, $10, 5 min	0.121	0.118
Opting out	P(Opt Out)		Informed at door of survey content	P(Do Survey)		Opting Out	P(Opt Out)		Informed at door of survey content	P(Do Survey)	
Opt Out, $0, 5 min	0.133	0.109	Not Informed, No Flyer	0.110	0.125	Opt Out, $0, 5 min	0.106	0.091	Not Informed, No Flyer	0.075	0.082
Opt Out, $10, 10 min	0.087	0.098	Informed, No Flyer	0.117	0.114	Opt Out, $10, 10 min	0.063	0.082	Informed, No Flyer	0.065	0.063
Opt Out, $10, 5 min	0.068	0.091	Not Informed, Survey Flyer	0.197	0.190	Opt Out, $10,5 min	0.056	0.079	Not Informed, Survey Flyer	0.129	0.119
Election Opt Out, $0, 5 min	0.136	0.112	Informed, Survey Flyer	0.203	0.170	Election Opt Out, $0, 5 min	0.146	0.095	Informed, Survey Flyer	0.124	0.091
Election Opt Out, $10, 10 min	0.104	0.102	Not Informed, Election Flyer	0.188	0.181	Election Opt Out, $10, 10 min	0.097	0.087	Not Informed, Election Flyer	0.087	0.104
Election Opt Out, $10, 5 min	0.086	0.095	Informed, Election Flyer	0.190	0.174	Election Opt Out, $10, 5 min	0.079	0.085	Informed, Election Flyer	0.106	0.091
			Not Informed, Opt Out	0.183	0.188				Not Informed, Opt Out	0.086	0.118
Lying	P(Lie | Survey)		Informed, Opt Out	0.177	0.167	Lying	P(Lie |Survey)		Informed, Opt Out	0.113	0.090
5 min survey, control	0.139	0.102	Not Informed, Election Opt Out	0.158	0.179	5 min survey, control	0.433	0.450	Not Informed, Election Opt Out	0.084	0.103
5 min survey, $5 incentive	0.156	0.164	Informed, Election Opt Out	0.150	0.173	5 min survey, $5 incentive	0.315	0.342	Informed, Election Opt Out	0.084	0.090
10 min survey, control	0.102	0.103				10 min survey, control	0.522	0.451			
10 min survey, 8 min incentive	0.148	0.190				10 min survey, 8 min incentive	0.409	0.375			
Turnout	0.600	0.599									

Notes: The Table presents the empirical moments and the predicted moments from a minimum-distance estimator. The empirical moments are obtained as regression estimates after controlling for the randomization fixed effects. The minimum-distance estimates are in Table 3, shown are the predicted moments at those parameter values.

Acknowledgments. We thank Nageeb Ali, Leonardo Bursztyn, David Card, Raj Chetty, Justin Fox, Andrew Gelman, Alan Gerber, Yosh Halberstam, Mitch Hoffman, Greg Huber, Rebecca Morton, Tom Palfrey, Ricardo Perez-Truglia, Todd Rogers, Glen Weyl, Noam Yuchtman, audiences at Berkeley, Boston University, Harvard, Microsoft Research, RAND, Stanford, Toulouse, UCSD, and at the 2011 ESA World Congress (Chicago), the 2012 North-American ESA Tuscon Conference, the 2012 NBER Political Economy Summer Institute, the 2013 Yale CSAP summer workshop, the 2013 SITE Psychology and Economics conference and the 2015 NYU-Abu Dhabi conference in Behavioral Political Economy for helpful comments. We also thank Alec Brandon, Tristan Gagnon-Bartsch, Sheng Li, David Novgorodsky, Jessica Shui, Alex Steiny, and Vera te Velde for excellent research assistance.

Supplementary Data

Supplementary data are available at *Review of Economic Studies* online.

REFERENCES

ALLCOTT, H. (2011), "Social Norms and Energy Conservation", *Journal of Public Economics,* **95**, 1082–1095.
ALI, N. and LIN, C. (2013), "Why People Vote: Ethical Motives and Social Incentives", *American Economic Journal: Microeconomics*, **5**, 73–98.
ANDREONI, J. and BERNHEIM, D. B. (2009), "Social Image and the 50–50 Norm: A Theoretical and Experimental Analysis of Audience Effects", *Econometrica*, **77**, 1607–1636.
ARIELY, D., BRACHA, A. and MEIER, S. (2009), "Doing Good or Doing Well? Image Motivation and Monetary Incentives in Behaving Prosocially", *American Economic Review*, **99**, 544–555.
ASHRAF, N., BANDIERA, O. and JACK, K. (2014), "No Margin, No mission? A field experiment on incentives for pro-social tasks", *Journal of Public Economics*, **120**, 1–17.
BELLI, R. F., TRAUGOTT, M. W., YOUNG, M. *et al.* (1999), "Reducing Vote Overreporting in Surveys: Social Desirability, Memory Failure, and Source Monitoring", *Public Opinion Quarterly*, **63**, 90–108.
BENABOU, R. and TIROLE, J. (2006), "Incentives and Prosocial Behavior", *American Economic Review*, **96**, 1652–1678.
Blais, A. (2000), *To Vote or Not to Vote: The Merits and Limits of Rational Choice Theory* (Pittsburgh: University of Pittsburgh Press).
BURSZTYN, L., CALLEN, M., FERMAN, B. *et al.* (2014), "A Revealed Preference Approach to the Elicitation of Political Attitudes: Experimental Evidence on Anti-Americanism in Pakistan" (Working Paper).
COATE, S. and CONLIN, M. (2004), "A Group Rule-Utilitarian Approach to Voter Turnout: Theory and Evidence", *American Economic Review*, **94**, 1476–1504.
COATE, S., CONLIN, M., and MORO, A. (2008), "The Performance of Pivotal-Voter Models in Small-Scale Elections: Evidence from Texas Liquor Referenda", *Journal of Public Economics*, **92**, 582–596.
CONLIN, M., O'DONOGHUE, T. and VOGELSANG, T. J. (2007), "Projection Bias in Catalog Orders", *American Economic Review*, **97**, 1217–1249.
DELLAVIGNA, S., LIST, J. A. and MALMENDIER, U. (2012), "Testing for Altruism and Social Pressure in Charitable Giving", *Quarterly Journal of Economics*, **127**, 1–56.
DELLAVIGNA, S., LIST, J. A., MALMENDIER, U. *et al.* (2014), "Voting to Tell Others" (NBER Working Paper No. w19832).
DOWNS, A. (1957), *An Economic Theory of Democracy* (New York: Harper).
DUFF, B., HANMER, M. J., PARK, W.-H. *et al.* (2007), "Good Excuses: Understanding Who Votes with an Improved Turnout Question", *Public Opinion Quarterly*, **71**, 67–90.
ERAT, S. and GNEEZY, U. (2012), "White Lies", *Management Science*, **58**, 723–733.
FEDDERSEN, T. and SANDRONI, A. (2006), "A Theory of Participation in Elections", *American Economic Review*, **96**, 1271–1282.
FINAN, F. and SCHECHTER, L. (2012), "Vote-Buying and Reciprocity", *Econometrica*, **80**, 863–882.
GERBER, A. S. and GREEN, D. P. (2000), "The Effects of Canvassing, Direct Mail, and Telephone Contact on Voter Turnout: A Field Experiment" *American Political Science Review*, **94**, 653–63.
GERBER, A. S., GREEN, D. P. and LARIMER, C. W. (2008), "Social Pressure and Voter Turnout: Evidence from a Large-scale Field Experiment", *American Political Science Review*, **102**, 33–48.
GERBER, A. S., HUBER, G. A., DOHERTY, D. *et al*. (2016), "Why People Vote: Estimating the Social Returns to Voting", *British Journal of Political Science*, **46**, 241–264.
GREEN, D. P. and GERBER, A. S. (2008), *Get Out the Vote: How to Increase Voter Turnout* (Washington, D.C.: Brookings Institution Press).
GREEN, D. P., MCGRATH, M. C. and ARONOW, P. M. (2013), "Field Experiments and the Study of Voter Turnout", *Journal of Elections, Public Opinion and Parties*, **23**, 27–48.
HARBAUGH, W. T. (1996) "If People Vote Because They Like to, Then Why Do So Many of Them Lie?", *Public Choice*, **89**, 63–76.
HARSANYI, J. (1977), "Morality and the Theory of Rational Behavior", *Social Research*, **44**, 623–56.
LACETERA, N. and MACIS, M. (2010), "Social Image Concerns and Prosocial Behavior: Field Evidence from a Nonlinear Incentive Scheme", *Journal of Economic Behavior & Organization*, **76**, 225–237.

LAIBSON, D., REPETTO, A. and TOBACMAN, J. (2007), "Estimating Discount Functions with Consumption Choices over the Lifecycle" (Working Paper).

LEDYARD, J. O. (1984), "The Pure Theory of Large Two-Candidate Elections", *Public Choice*, **44**, 7–41.

LEIGHLEY, J. E. and NAGLER, J. (1992), "Socioeconomic Class Bias in Turnout, 1964–1988: The Voters Remain the Same", *American Political Science Review*, **86**, 725–736.

LEVINE, D. K. and PALFREY, T. R. (2007), "The Paradox of Voter Participation? A laboratory study", *American Political Science Review* , **101**, 143–158.

PALFREY, T. and ROSENTHAL, H. (1983), "A Strategic Calculus of Voting", *Public Choice*, **41**, 7–53.

PALFREY, T. and ROSENTHAL, H. (1985), "Voter Participation and Strategic Uncertainty", *American Political Science Review*, **79**, 62–78.

PASSARELLI, F. and TABELLINI, G. (2016), "Emotions and Political Unrest", *Journal of Political Economy*, forthcoming.

PEREZ-TRUGLIA, R. and CRUCES, G. (2013), "Social Incentives in Contributions: Field Experiment Evidence from the 2012 U.S. Presidential Campaigns", unpublished manuscript.

RIKER, W.H. and ORDESHOOK, P. C. (1968), "A Theory of the Calculus of Voting", *American Political Science Review*, **62**, 25–42.

ROGERS, T., TERNOVSKI, J. and YOELI, E. (2016), "Potential Follow-up Increases Private Contributions to Public Good", *Proceedings of the National Academy of Sciences*. Advance online publication. doi:10.1073/pnas.1524899113.

SHUE, K. and LUTTMER, E. (2009), "Who Misvotes? The Effect of Differential Cognition Costs on Election Outcomes", *American Economic Journal: Economic Policy*, **1**, 229–257.

SILVER, B. D., ANDERSON, B. A. and ABRAMSON, P. R. (1986), "Who Overreports Voting?", *American Political Science Review*, **80**, 613–24.