

Can Governments Use Get Out the Vote Letters to Solve Europe's Turnout Crisis?

Evidence from a Field Experiment

Bhatti, Yosef; Dahlgaard, Jens Olav ; Hansen, Jonas Hedegaard; Hansen, Kasper Møller

Document Version

Accepted author manuscript

Published in:

West European Politics

DOI:

[10.1080/01402382.2017.1339985](https://doi.org/10.1080/01402382.2017.1339985)

Publication date:

2018

License

Unspecified

Citation for published version (APA):

Bhatti, Y., Dahlgaard, J. O., Hansen, J. H., & Hansen, K. M. (2018). Can Governments Use Get Out the Vote Letters to Solve Europe's Turnout Crisis? Evidence from a Field Experiment. *West European Politics*, 41(1), 240-260. <https://doi.org/10.1080/01402382.2017.1339985>

[Link to publication in CBS Research Portal](#)

General rights

Copyright and moral rights for the publications made accessible in the public portal are retained by the authors and/or other copyright owners and it is a condition of accessing publications that users recognise and abide by the legal requirements associated with these rights.

Take down policy

If you believe that this document breaches copyright please contact us (research.lib@cbs.dk) providing details, and we will remove access to the work immediately and investigate your claim.

Download date: 27. Jun. 2024



Can Governments Use Get Out the Vote Letters to Solve Europe's Turnout Crisis?: Evidence from a Field Experiment

Yosef Bhatti, Jens Olav Dahlgaard, Jonas Hedegaard Hansen, and Kasper Møller Hansen

Journal article (Accepted manuscript*)

Please cite this article as:

Bhatti, Y., Dahlgaard, J. O., Hansen, J. H., & Hansen, K. M. (2018). Can Governments Use Get Out the Vote Letters to Solve Europe's Turnout Crisis? Evidence from a Field Experiment. *West European Politics*, 41(1), 240-260. DOI: 10.1080/01402382.2017.1339985

This is an Accepted Manuscript of an article published by Taylor & Francis in *West European Politics* on 30 Jun 2017, available online:

DOI: <http://www.tandfonline.com/10.1080/01402382.2017.1339985>

* This version of the article has been accepted for publication and undergone full peer review but has not been through the copyediting, typesetting, pagination and proofreading process, which may lead to differences between this version and the publisher's final version AKA Version of Record.

Uploaded to [CBS Research Portal](#): February 2019

Can Governments Use Get Out The Vote Letters to Solve Europe's Turnout Crisis? Evidence from a Field Experiment

ACCEPTED FOR PUBLICATION IN WEST EUROPEAN POLITICS

Preprint manuscript

*Yosef Bhatti***, *Jens Olav Dahlgaard****, *Jonas H. Hansen** & *Kasper M. Hansen*§*

*University of Copenhagen, Department of Political Science,
Øster Farimagsgade 5, DK-1353 Copenhagen K, Denmark

**KORA – Danish Institute for Local and Regional Government Research,
Købmagergade 22, 1150 Copenhagen K, Denmark

***Department of Business and Politics, Copenhagen Business School,
Steen Blichers Vej 22, 2200 Frederiksberg, Denmark

E-mail: yobh@kora.dk, jod@ifs.ku.dk, jhh@ifs.ku.dk & kmh@ifs.ku.dk

Abstract

Declining levels of turnout are a problem in European elections. Are Get Out The Vote campaigns the solution to the problem? While many studies have investigated such campaigns in the US, we know little about their effect in Europe. We present a field experiment in which encouragements to vote in an upcoming Danish election are delivered using direct personal letters to more than 60,000 first-time voters. Eight different letters are designed based on the calculus of voting and prospect theory. The sample is randomly divided into treatment groups or the control group. Using validated turnout, we find small positive effects of receiving a letter on turnout, with little difference across letters. The letters mostly mobilized voters with a low propensity to vote and thus increased equality in participation. In sum, while letters have some effect, they are not likely to be a panacea for solving Europe's turnout challenges.

Keywords Voter turnout; Get Out The Vote; Prospect theory; Inequality; Field Experiments; Calculus of voting

§All four authors contributed equally and are listed alphabetically by their last name.

Can Governments Use Get Out The Vote Letters to Solve Europe's Turnout Crisis? Evidence from a Field Experiment

In the 2014 European Parliament elections, the electoral turnout dropped again, hitting an all-time low level of 42.6 percent (European Parliament 2016). With this turnout, the smallest proportion of Europeans ever expressed their commitment to the representative institution of the European Parliament. This nadir was reached despite the many initiatives taken by the institutions of the European Union and other actors prior to the European election to boost interest and ultimately voter turnout. For instance, the European Parliament ran a large information campaign with the slogan "This time it's different" (European Parliament 2014), the 'Spitzenkandidaten' procedure was introduced (Dinan 2015) to make the election more decisive in terms of electing the president of the European Commission, and the Get Out The Vote (GOTV) campaign video "Happy Voting" was initiated by non-institutional actors and volunteers (European Youth Forum 2014).

Decreasing turnout in Europe is by no means isolated to the European Parliament elections but is also a trend in many countries' national and local elections (International IDEA 2015). It has been shown that the average overall turnout since the 1970s has decreased by 11 percentage points across 31 'full democracies', including many European countries (Vowles 2017). High turnout is the central health indicator for a democracy because it indisputably shows how many citizens have made the effort to provide their opinion on the distribution of seats in the European Parliament, national parliaments, and local government bodies. The mark on the ballot is not just a mark for a party or candidate but a signature on the 'social contract' of representative democracy, thus providing a measure of the democratic legitimacy of the representative body (Rousseau 1762; Beetham 1991; Lijphart 1997). Therefore, the turnout decline is an important societal challenge to address.

One possible remedy is voter mobilization campaigns. Over the last 15 years, we have witnessed significant efforts to mobilize voters in the US, and similar campaigns have recently been

gaining traction among governments, parties, and organizations in European elections. However, while a vast body of literature have evaluated the ability of American campaigns to mobilize voters, we only have modest knowledge about the effect of such campaigns on European voters. Treatment effects could potentially differ across contexts due to cultural differences, institutional differences (e.g. automatic registration and widespread use of PR systems in Europe) and the markedly higher European turnout which may limit the mobilizing potential.

The main contribution of our study is to examine the effects of direct mailings on turnout in Europe. Mailings are a particular interesting mobilization tool due to their scalability. On behalf of the Danish Ministry of Interior, we sent a direct, personal postal letter with an encouragement to vote to more than 60,000 first-time voters. The Danish context is particularly interesting due to its high turnout compared to US elections. The research design allows us to test the overall effects of letters and effects across various messages. Both governmental and non-governmental actors focusing on increasing turnout can easily adopt this type of campaign.

Our second contribution is to add to the part of the literature examining differences in effects due to the content of the treatments. Although the GOTV literature have expanded vastly during the last 15 years, few studies exists that conducts horseshoe tests of different content (Gerber & Green 2016). Specifically, we test eight different treatments based on two established theories: the calculus of voting and prospect theory (Riker & Ordeshook 1968; Tversky & Kahneman 1981). To know which type of content that works best would be important knowledge for governmental and nongovernmental organizations applying GOTV-campaigns.

Our third contribution concerns heterogeneous effects of our treatments. A central goal for many GOTV campaigns is to reduce inequalities in participation (Bedolla & Michelson 2012). However, recently scholars have questioned whether GOTV campaigns actually increase or decrease equality in participation (Arceneaux & Nickerson 2009; Enos et al. 2014; [reference

omitted for review]). For instance, in a recent study, Enos et al. (2014) find that most studies in an American context primarily mobilize high-propensity voters and thereby add to inequality in participation by increasing the gap between high- and low-propensity voters' turnout. Nevertheless, it is likely that the results from a European context differ as it has been suggested that inequalities in effects varies with control group turnout (Enos et al. 2014). Thus, we find it crucial to investigate this growing concern in the mobilization literature in a European context. The Danish context is particularly suited due to the existence of high quality register data which provides us with a long list of sociodemographic variables for each individual. This allows us to estimate reliably how treatment effects varies across individual propensity to vote score.

The main results of our experiments are not encouraging for practitioners hoping for a quick fix of European turnout challenges. The letters have small, insubstantial effects on turnout, although some versions of the letters have promising effects. More encouraging, we find that the small effect is larger among low-propensity voters than among high propensity voters, leading to a decrease in turnout inequality. The study illustrates the importance of rigorously testing mobilization campaigns, something rarely done in a European context. As a field, we need to know both the successes and failures of high quality mobilization campaigns to establish their impact and avoid bias from the publication of only positive results (Franco et al. 2014).

Mobilization through GOTV campaigns

The modern voter mobilization literature can arguably be divided into two categories, with one focusing on the effectiveness of different modes of contact and the other on the effectiveness of different messages (Gerber & Green 2016). Most existing research has been conducted in the US, where more than 150 field experiments have been published since Gerber and Green's study in New Haven in 1998 (Gerber & Green 2000; Green, McGrath & Aronow 2013; Gerber & Green 2016).

Studying different modes of contact, especially direct mailings and door-to-door canvassing, has received much attention, while others have tested phone calls, e-mailing, and text messages (e.g., Gerber & Green 2000; Gerber, Green & Larimer 2008; Nickerson 2007; Nickerson 2008; Dale & Strauss 2009).

The large number of studies using the same basic research design (i.e., random controlled field experiments) has enabled a meaningful meta-analysis from which we can draw some general findings (see Green, McGrath & Aronow 2013 for an extensive meta-review and Gerber and Green 2016 for a recent review of the literature). Personal contact via door-to-door canvassing delivers substantially larger effects than the impersonal modes of contact. Based on data from Green, McGrath and Aronow (2013), the pooled estimate of 70 US-based canvassing experiments is a 2.5 percentage points increase in turnout among those reached by the campaign (see also [reference omitted for review]).

While the effect of door-to-door canvassing potentially is substantial, scalability poses a substantial problem for practitioners. Implementing a door-to-door campaign targeting tens of thousands of voters requires considerable organization and resources, and for a government wanting to increase participation among a large number of citizens, it is not a realistic approach. Furthermore, contact rates are typically well below 50 percent ([reference omitted for review]), and those who are contacted will often be those most likely to vote in the first place (Gerber & Green 2000). Direct mailings could be a better solution because they can easily be mass distributed and do not require the target to be home when someone knocks on the door. Across 85 US experiments applying direct mailings, the pooled effect estimate is 0.76 percentage points (cf. Gerber & Green 2016: table 4). In addition, advocacy mailings seem to have no effect on turnout, non-advocacy mailings increase turnout by 0.52 percentage points on average, and social pressure mailings deliver effect sizes of 2.28 percentage points on average.

The cumulated studies together have taken us a long way in understanding how we can design mobilization campaigns to increase turnout more efficiently. However, the clear majority of these studies are from a non-European context, and it is far from certain that the same type of mobilization campaigns will work the same way in Europe. For instance, a review of nine studies using door-to-door canvassing to increase turnout in Western Europe finds that the best point estimate of the effect is substantially smaller in Europe than in the US ([reference omitted for review]).¹

Previous studies have also examined mailings in Europe to a limited extent. Fieldhouse et al. (2013; 2014) find mailings to increase turnout among British voters in two elections, with intention-to-treat (ITT) effect sizes between 1 and 2.9 percentage points. [reference omitted for review] find that mailing the Danish constitution along with an encouragement to vote to first-time voters increases turnout among young Danes, with an ITT-effect of 1.1 percentage points. In a working paper, Ramiro et al. (2012) finds no effect of partisan leaflets delivered in the city of Murcia in Spain. Foos & John (2016) finds that delivering partisan leaflets on behalf of the Conservatives did not increase overall turnout, but that the campaign does seem to change the electorate's composition in favor of the Conservatives. All in all, initial studies from a European context show some promise with respect to the use of nonpartisan mailings for voter mobilization, while partisan mailings and leaflets seem to have no or little positive effect on turnout.²

Inspiration from political psychology and prospect theory

These results lead us to the second focus of the literature, which is the testing of different messages.

In this strand of research, insights from social psychology and behavioral economics have been

¹ The canvassing studies included in the meta-analysis are John & Brannan (2008); Nyman (2014); Pons & Liegey (2016); Pons (2016); Cantoni & Pons (2016); Foos & John (2016); ([reference omitted for review]).

² Since the focus of this paper is on how mailings can be used to increase turnout, we limit the literature review to studies most relevant for this purpose. For European studies applying phone calls, see John & Brannan (2008); Fieldhouse et al. (2013). For text messages, see ([reference omitted for review]) and for a lottery, see John et al. (2015).

used to formulate GOTV messages (Gerber & Green 2016).³ Whereas social pressure messages have produced substantively large effects, expressing gratitude for voting or asking people to plan how and when they will cast their vote seem to produce minor effects. A central concern with the existing studies of different messages is, in the words of Gerber and Green (2016: 25), that *“because there are relatively few studies that conduct a “horserace” between messages in which scripts vary but the other experimental conditions are held constant, it is possible that some of the observed differences in message effectiveness are due to variation in conditions other than the message (...)*”. Thus there is a need for further studies that tests differences between message content in the same experiment by assigning random sub-groups to different messages.

The theoretical point of departure for our tests is the calculus of voting, where an individual decides to vote if the perceived utility of doing so is positive (Riker & Ordeshook 1968; Downs 1957). Formally,

$$R = P*B - C + D$$

R is the reward or utility of voting for the individual and depends on the following four factors: P is the probability that the individual’s vote will be decisive for the election result; B is the individual gain if the preferred party or candidate wins the election compared to the alternative; C is the cost of voting, such as the cost of gathering information for one's vote choice and the effort involved in casting the vote (time, transport, etc.); D is the feeling of duty towards voting. This includes the satisfaction gained by performing one’s civic duty of voting by, e.g., living up to the democratic norm of participation by voting.

³ Researchers have also tested, e.g., monetary incentives (Panagopoulos 2013) and the relevance of timing (e.g., Nickerson 2006; Panagopoulos 2011), also based on psychological and behavioral science theories.

While we cannot alter the actual values in the calculus in a letter⁴, it might be possible to change the citizens' perception of the different values in a positive direction. This can be done by increasing the perceived saliency of, e.g., the benefits gained from voting in the citizens' mind (Nelson et al. 1997). Regardless of whether the arguments put forward in the treatment are new to the citizens or have been heard before, reading the arguments should increase the accessibility of thoughts that are positively associated with the decision to vote and perhaps even remind the reader of the external pressure to vote. The consequence of receiving a treatment should therefore be a more positive attitude towards voting. The essential question is whether this change in attitude causes a change in behavior. If so, an increase in turnout should result. Some American studies have examined the difference in effects depending on which element of the voting calculus one appeals to; they generally find no differences across treatments (Gerber & Green 2000; Green, McGrath & Aronow 2013).

In addition to the calculus of voting, we draw on prospect theory. The idea behind this theory is to be explicit regarding how to frame the arguments. We do this by putting forward equivalent arguments, with the only difference being whether the argument is framed in terms of potential gains or losses by voting (Chong & Druckman 2007). Prospect theory is often applied in lab or survey experimental designs, in which the dependent variables are attitudes or intentional behaviors. In classic studies tested on small samples of university students, it has been shown that people tend to react more strongly to potential losses than potential gains. This result implies that our loss frame should have the largest effect (Tversky & Kahneman 1979; Tversky & Kahneman 1981). By applying the theory in a field experiment, we are able to evaluate whether the gain/loss framing actually makes a difference for real-world voting behavior. We expect that arguments referring to loss of democratic influence will be more effective in increasing turnout than arguments

⁴ One could, however, identify certain elections in which, e.g., one's chance of casting the decisive vote would be extraordinarily high (Enos & Fowler 2014).

referring to increased democratic influence by voting. One study has examined this before in an American context. In two experiments using personally delivered messages, Arceneaux and Nickerson (2010) find no systematic difference in the effects of positive and negative messages on turnout. In this study, we re-examine the effects in a new context using a non-personal form of delivery.

We also test if the quantity of arguments makes a difference. There might be a synergetic effect from multiple arguments. Individuals may draw inferences about the quality of a position based on the quantity of arguments (Winke, Bless & Biller 1996). Additionally, different types of arguments may convince people. Multiple arguments in one treatment could make it more likely that at least one argument among many will motivate the voter to go the polling station. However, information overload may cause individuals to discard the letter or to miss the most effective argument.

Context and data

The context of our study is the Danish Municipal Elections held on November 19, 2013 across the 98 Danish municipalities.⁵ This is a setting with a salient election, with multiple parties competing in a proportional election in a media setting dominated by a public service broadcast with emphasis on politically balanced news. The Danish municipalities are central in the government service provision, and around one third of the total Danish GDP (2010) is spent at the local government level (Statistics Denmark 2012). The areas of municipal responsibility include, among others, care for the elderly, childcare, and schools.

Danish municipal elections are multiparty races dominated by the main national parties. Allocation of seats to each party is proportional, and there is no official threshold. Individual

⁵ The election for the five regional governing bodies of Denmark were held together with the municipality elections. The Regions plays a quite small part in the public in Denmark. The turnout for the two elections was almost equal. For simplicity we focus on the municipalities in this paper.

members of the municipal boards are mostly selected from open lists. The average turnout across all municipalities in the 2013 elections was 71.9 percent. During recent decades, the turnout has fluctuated approximately 70 percent in municipal elections. The turnout is somewhat lower than in national elections, which have fluctuated around 85 percent. Turnout among 18-21-year-olds, the target group in our experiment, was 61 percent in 2013. The news media pay much attention to the local elections in Denmark, and the national politicians and national parties are very active in the election campaign.

All Danes have a personal ID number linked to their personal address. In the field experiment, we randomize treatment at the household and individual citizen level using the personal ID numbers and addresses. Individual, validated turnout is merged to the personal ID, which gives a measurement of actual turnout for all citizens with practically no measurement error, freeing us of issues with self-selection or overreporting (Karp & Brockington 2005; Bernstein et al. 2001). In addition, we have a large number of variables from high quality government registers available to us through *Statistics Denmark*, the official statistics bureau. Among the variables that we merge with the individual level voting status, address, and treatment are gender, age in days, completed and ongoing education, composition of the household, ethnicity, and employment. These data provide us with strong predictors of turnout, which in turn yield good opportunities to study the heterogeneous effects of the treatments. We use these variables to predict an individual propensity score of voting, and we assess whether this propensity score is linked to our treatment effects.

Field experimental design

Since Gerber and Green's seminal study (2000), the default method for evaluating the effect of GOTV campaigns has been field experiments, sometimes referred to as Random Controlled field Trials. In a field experiment, a target group of voters is randomly assigned with some known probability greater than zero and less than one to either receive a campaign communication or voting encouragement or to be in the control group. Random assignment of subjects to treatment and control implies "*creating two groups that are, in expectation, identical prior to application of the treatment*" (Gerber & Green 2012: 31). While any given random assignment of treatments may produce groups that have different potential outcomes, the randomization process is fair in the sense that it on average does not favor any of the groups. Furthermore, when the sample size is as large as is the case in this study, the chance of obtaining widely different groups due to sampling variability is very small. Overall, the randomization should protect us from omitted variable bias that could potentially produce a spurious effect in our analysis.

To illustrate the advantages of a field experiment with validated turnout, we can contrast it with another approach that researchers may feel tempted to apply when evaluating the effect of a GOTV campaign: a survey. In a survey, researchers could ask a sample if they recall seeing campaign communications and if they turned out to vote. However, such an approach is likely to give a biased estimate (Gerber & Green 2000). Campaigns may target likely voters, and likely voters could pay more attention or better recall campaigns. In addition, likely voters may be more reachable (Arceneaux et al. 2006). Consequently, when subjects are not treated at random, we are unlikely to obtain unbiased estimates of the effect. The field experiment, where voters are randomly assigned to treatment and control, provides a fair test of the effect, free of omitted variables (Gerber & Green 2012: 7).

The virtues of using data with validated turnout are that when turnout is not validated but self-reported, contacted voters may over-report as a consequence of a demand effect (Karp & Brockington 2005; Bernstein et al. 2001). In addition, validated turnout holds potential for greatly improving the power of an experiment. In this paper, we consider turnout for more than 140,000 voters. Obviously, contacting that many voters is out of reach for all normal survey research.

Sample and randomization

The field experiment we present targeted first-time municipality voters aged 18-21 on Election Day in all 98 municipalities in Denmark. This group was selected as targets for a GOTV campaign by the Ministry of Economic Affairs and the Interior, who chose to cooperate with the authors about the design and evaluation of the campaign. Ahead of the election, the researchers obtained a list with all first-time municipality voters in Denmark from the Danish registers. The age group was partitioned into several subgroups, as some of the young voters were also subject to treatment in separate mobilization experiments ([references omitted for review]). Here, we consider only the young voters who were not treated or did not reside with young voters treated in other experiments. Some young voters shared a residence with one or two others in the same age group. For them, we used a double randomization process as recommended by Gerber & Green (2012: 257-58), where we first sampled households for treatment and then sampled one subject within each household for treatment. The subjects in the untreated households were placed in the control group. In the multi-member household, we keep only the directly treated participant in the treatment households. The untreated subjects in the treated households could potentially be subject to spillover effects which makes it inappropriate to include them in the control group (Gerber & Green 2012, [reference omitted for review]). Since the direct treatment effects were too small to meaningfully explore spillovers, we simply removed the indirectly treated subjects, i.e. the untreated subjects living

together with a treated person, from our experiment. The remaining young voters, those who did not share residence with one or two peers, were randomly assigned to either treatment or control.⁶ The sampling leaves us with a total sample size of 141,684 after we remove the few voters missing from the final voter files or with missing values on key background covariates.⁷

Treatment design

We mailed 64,923 personal letters in a closed envelope to potential first-time voters. Treatment individuals were randomly assigned to eight treatment wordings. The control group contained 76,761 voters. The receiver's name was typed on the envelope and on the letter itself. Each letter followed the same template similar to what we described above. This template included the receiver's own full name followed by a short, general text about the election. After this text, a headline appeared in bold with a short argument, defined by the treatment group. Finally, some practical information about the voting process was included before the treatment headline also concluded the text. The letter was signed by The Ministry of Economic Affairs and the Interior and printed on official letterhead paper from the ministry (all letters are shown in a translated version in the supporting information). While this type of research does not require ethical approval from an institutional review board in Denmark, the design and letters were discussed with colleagues, and was approved by the Ministry of Economic Affairs and the Interior whom also paid for the deliverance of the letters. Of the more than 60,000 letters sent, only one percent was returned as failed delivery. We opened the returned letters and confirmed that there was no systematic bias in the returned letters, which indicates that the randomization worked as intended. As an additional

⁶ Some of these voters lived with three or more peers. We considered them as single-young-voter-households based on the assessment that these household sizes are more likely to be like co-opts where the interaction between each person is more limited than in small households with one or two other 18-21-year-olds. These voters only constitute a minor proportion of the sample. In the supporting information, we present the main results without these voters. The divergences are minimal.

⁷ We only lost approximately 1 percent of the population of interest this way. Crucially, assignment to treatment was unlikely to affect who we lost.

check, we also included our names and addresses among the various experimental groups, and the letters were received according to the randomization and at the correct time.⁸

We combined the calculus of voting and prospect theory aspects in eight treatment letters. For three of the four factors in the calculus of voting, benefit, duty, and pivotality, there was an argument framed as either a gain or loss. This testing of content is both theoretically interesting and interesting for practitioners because it can provide guidelines to produce the most effective campaigns. The substance of the arguments within each of the factors was equivalent. For instance, the letter with the potential benefits from voting, framed as a gain, had the following headline: "**Take part in deciding your everyday - vote**". The text stated, "If you vote in the municipal and regional elections, you take part in deciding how the money will be spent in your municipality and region. Make your influence count – vote!" Contrary to this message, the benefit argument framed as a loss had the following headline: "**Do not let others decide your everyday - vote**". The text stated, "If you do not vote in the municipal and regional elections, you do not take part in deciding how the money will be spent in your municipality and region. Don't lose your influence – vote!" For the duty and pivotality arguments, we had similar variations. We only sent one treatment for the cost argument, which focused on reducing the perceived cost associated with voting. Finally, a group received a letter combining all the arguments. The letter consisted of arguments focused on the cost of voting, perceived benefits, the probability of being the pivotal voter, and the civic duty of voting. We direct the reader to the supporting information for a full description of the content of each treatment.

Estimation strategy

⁸ In recent years, pre-registration of research designs has gained traction in political science (for a brief introduction, see Monogan 2015). Unfortunately, our study was not pre-registered with a time-stamp ahead of conducting it. The research design was presented to colleagues at the MPSA conference in Chicago in April, 2013 ([Reference omitted for review]). Although unfortunate, we note that pre-registration was not that common at the time of conducting our study.

We use a linear probability model with robust standard errors to estimate the effect of receiving any treatment at all and the effect of each of the eight distinct treatments. In line with a difference-in-means estimator, a linear probability model is an unbiased estimator of the average treatment effect in an experiment with a dichotomous treatment (Gerber & Green 2012: 102; Angrist et al. 2009). Furthermore, a linear probability model yields standard errors that in large samples as ours tend to match those from more sophisticated techniques as for instance randomization inference (Gerber & Green 2012, p. 115).⁹ An advantage of using a linear probability model compared to a difference-in-means estimator is that it is easy to include background covariates without jeopardizing the unbiasedness. Given our random assignment of units to treatment and control, including background covariates is neither more or less biased, but can reduce the standard errors and increase the precision of the estimate (Gerber & Green 2012, p. 103).¹⁰ Finally, we have to adjust for the fact that the probability of assignment to treatment and control varied across subjects. We weight our regression with weights that are inverse to the probability of being assigned to treatment. This is necessary in an experiment where treatment probabilities vary for subjects and if we did not do this, our estimates would be biased (Gerber & Green 2012, p. 76).¹¹

Results

Table 1 presents the actual turnout for each treatment condition, including the control group, weighted with inverse probability weights. In general, treated voters are more likely to vote. The

⁹ With a limited dependent variable we could also use other regression techniques such as logistic regression. However, since linear probability models give an unbiased estimate of the average treatment effect in units (percentage points) that are easy to interpret and the standard errors are equivalent to more sophisticated techniques in large samples as ours, we opt for simplicity and use a linear probability model.

¹⁰ The point estimate will change too, but only due to sampling variability in the covariates.

¹¹ As previously discussed, we sampled within sub-groups to avoid contamination from other experiments being run in the same elections. Treatment probabilities for the four sub-groups used were 0.497, 0.222, 0.513 and 0.326 with probabilities for being in the control group of one minus the treatment probabilities. The weights used are the inverse of the probabilities.

voters treated with the duty-framed-as-loss message and all arguments are the most likely to vote. In the supporting information, we pool turnout for the loss and gain frames and for the pivotal, duty, and benefit messages. However, there is no systematic variation in the extent to which any of the variants mobilized voters.

***** Table 1 around here *****

In table 2, we present the treatment effects from our linear probability model weighted with inverse probability weights. The first two columns display the effect of receiving any letter at all, both with and without inclusion of background covariates. The background covariates are as follows: an interaction between sex, age in days, and age in days squared; household type; country of birth grouped by type; completed and ongoing education; occupation; income; municipality fixed effects; a four-way interaction between each parent's turnout in the 2009 municipality election and whether or not the voter lived with each parent; and an indicator for whether the young voter was eligible to vote in the Danish 2011 General Election. In the third and fourth column, we display the effect of each individual treatment with and without background covariates.

***** Table 2 around here *****

When we pool all our treatments together without background covariates in column 1, there is an effect of 0.41 percentage points. The effect is slightly larger and it reaches statistical significance, 0.48 percentage points, when we include background covariates in column 2. Based on this result, our best estimate is that the GOTV letters made recipients marginally more likely to

vote. The point estimates are within a credible interval of [0.299; 0.748] from a meta-analysis on U.S. non-advocacy mailings (Gerber & Green 2016). However, they are smaller than what both Fieldhouse et al. (2014) find using letters and phone calls and what [reference omitted for review] find in European door-to-door-canvassing experiments. Thus, based on this study direct personal letters seem substantially less promising as a means to improve Europe's declining turnout than indicated by the early studies in the field.

In columns 3 and 4, we examine the individual treatments with and without background covariates. The differences between the treatments are small, and the effects are not significantly different from one another. If we, cautiously, look at individual treatment effects, the effects for the letter with all arguments and for the letter with the duty framed as a loss argument stand out. Two interpretations of the effect of all arguments may be that more arguments are more convincing or that they increase the likelihood that there is one argument that speaks to the recipient. The effect of the duty framed as loss argument could also be consistent with the literature, in which negative social pressure has been demonstrated to be very effective (Gerber et al. 2008, Green et al. 2013). However, as we already cautioned, the differences are small, and we ran multiple comparisons. The differences could also just be sampling variation, and the extent to which they reflect meaningful differences is a question for future experiments. As described above, we also collapse our treatments from table 3 together by type and framing in the supporting information. We see no significant differences in effects between different message types or framings.

Does the campaign help decrease inequalities in turnout?

The aim of our mobilization effort was not only to decrease the young voters' underrepresentation in the election but also to level out differences within the group of first-time

voters. We next turn to examining if our treatment had a different impact conditional on voters' propensity to vote. Searching for heterogeneous effects of treatment effects is a risky venture because it requires multiple comparisons and induces researcher degrees of freedom (Green & Kern 2012). To minimize these problems, we follow Enos et. al. (2014) and estimate a propensity score for voting. We use our control group to fit a logistic model of turnout from which we obtain predicted probabilities for voting among our control group and our treatment groups. To avoid overfitting, we use only covariates that previous research has shown to predict turnout well among first-time voters at Danish elections ([reference omitted for review]). The covariates are identical to the list of covariates applied in table 2. If our treatment had heterogeneous effects conditioned on the propensity to vote, we should expect an interaction between our treatment and the estimated propensity score.

In the left panel of figure 1, we display the marginal, predicted probabilities from a logistic regression of turnout on propensity to vote interacted with treatment. In the right panel, we display the marginal treatment effect. Each panel includes a rug plot where we have plotted each percentile from the distributions of the propensity to vote, including the maximum and minimum value.¹²

Figure 1: Interaction effect between treatment and propensity to vote

***** Figure 1 around here *****

The figure indicates that the marginal effect is higher for the least likely voters. The effects seem to be only statistically significant for voters who we would predict to have a probability of voting of approximately 50 percent or less, whereas high-propensity voters seem to

¹² In the supporting information, we display the coefficients in a table.

be unaffected. Although the overall effect of the mobilization campaign was small, figure 1 suggests that, if anything, the effort was more efficient in mobilizing low-propensity voters than high-propensity voters. This adds somewhat to the potential for letters in Europe to contribute to marginally greater equality in voting between groups.

We can reflect upon why there is a discrepancy between the findings from the US-based experiments (Enos et al. 2014) and our results in figure 1. A potential explanation is that our experiment is conducted in a context with relatively high turnout compared to the US-based GOTV-experiments. Since GOTV campaigns are a quite brief and simple interaction between a campaign and a potential voter, it seems reasonable to expect that such campaigns primarily can affect voters near the threshold of turning out to vote. The threshold of voting is to a large degree determined by the general interest in the election (Arceneaux & Nickerson 2009). Consequentially, in elections with relatively low general interest and turnout – which many of the US experiments have been conducted in relation to – the voters close to the threshold of voting shares most characteristics with those already voting. When studying elections with relatively high interest and turnout – such as the one studied in this article – we should expect voters around the threshold to share more characteristics with the nonvoters. In fact, Enos et al. (2014) in their online appendix presents analyses suggesting a negative effect of higher turnout propensity when control group turnout is above 50 percent which is the case of our study. Our results here and elsewhere ([reference omitted for review]) are in line with these expectations since the newly mobilized voters indeed does seem to share characteristics with nonvoters instead of those already voting ([reference omitted for review]).

The finding can also be viewed in a broader perspective about seeing voting as a social act (Rolfe 2012; Sinclair 2012). In a context with a high general interest in the upcoming election, the social network of high propensity voters will likely already be talking about the

election and affecting each other to participate. However, we can imagine that the election have not received much attention within the social network of low propensity voters in such context. Thus, the GOTV-mailing potentially makes a larger difference for the low propensity voter than the high propensity voter since the latter are already in a social network in which people vote in the election.

Discussion

Turnout has been declining in both national European elections and elections for the European Parliament. This fact is countered by campaigns to increase voter turnout, many of which have been inspired from lessons from an American context. In this paper, we presented evidence from a large-scale field experiment on the effect of mailed voter encouragements among young Danish voters. We found that the mailings had a limited effect on turnout. Voters who received any postal letter mobilization message had a turnout rate approximately 0.4 percentage points higher than the control group, who did not receive a letter. The effect was statistically significant when pretreatment covariates were used to increase the precision of the estimate. In sum, letters do not seem to be a panacea for solving Europe's turnout challenges.

We mailed eight different versions of the letters to the young voters based on insights from prospect theory and the calculus of voting. Overall, there were only small variations in the effect of each letter. The letter framed as the duty of voting as a potential loss and a letter in which we applied all arguments had the largest effects. However, the effects of these letters were not statistically larger than the other effects. The effect of the duty-framed-as-loss mailer could be seen to align with the finding in the literature that social pressure mailings are among the most effective means of mobilizing voters. The effect of the letter with all arguments lends support to the idea that more arguments can be more convincing than few. Perhaps there is an additive effect of arguments,

or the individual might find one argument among many that influences him or her. Furthermore, the result also suggests that we should not worry about providing multiple arguments.

Although the effect estimates are higher for two of the letters than the others, we highlight that they are not significantly different from each other. Ultimately, we need more research and more statistical power to establish if the variations express meaningful differences or statistical noise. If they are indeed expressions of meaningful differences in the effectiveness of the different messages, future studies could also examine the effects of the number of arguments, including if and when information overload kicks in. When we pooled the letters based on the aspects in the calculus of voting, we found no difference in the treatment effects. Likewise, we found no difference between the treatments that were framed in terms of a potential gain or a loss, as inspired by insights from prospect theory.

Finally, we found some evidence for a curvilinear relationship between the individuals' propensity to vote and the treatment effect. The results suggest that the effect is strongest among low-propensity voters. This is an encouraging finding because it suggests that the mobilization campaign decreased inequalities in turnout. For organizations and politicians worried about inequalities in political participation, our results show that a quite simple campaign intervention perhaps can make the electorate more representative of the general population. Furthermore, we believe that our findings on this points towards a useful avenue for future experimental research that investigates in more details how to mobilize underrepresented citizens to use their vote in elections as well as engage in other types of political participation.

Although the mailings did increase turnout, the effects were small. As such, a mobilization effort as we evaluate here may provide a marginal improvement in turnout, but clearly our message was not a silver bullet that "solved" the problem of low turnout. In addition, sending mailers to young voters is expensive. If the price per mailing is approximately 1.5 €, the price per additional

vote was 312 to 365 €, depending on the method of estimation. This finding leads to a central strength of our experimental design: now we know. Too often policy-makers field initiatives, like the one we have presented here, without properly evaluating their effect (Dahler-Larsen 2011; Vedung 1997). An experiment provides a rigorous and fair test of the causal effect of the initiative and can easily be implemented by practitioners and scholars in diverse countries who want to test theories of how to increase voter turnout.

Because of our design, we can also relate our findings to a central issue that we raised in the beginning of this paper: the degree to which the findings from the primarily US-based studies can be transferred to a distinctly different context. Our study shows that the use of direct, non-partisan mailings seem to work to approximately the same degree as the average American experience with this type of treatment. However, the cost-effectiveness is lower because mailings are expensive in the context under investigation here. In this way, our study suggests that the findings do seem to be transferable. We must stress two important points. First, this is only among the first large-scale field experiment on mobilization letters conducted in Europe. Surely, there is a need for more large-scale field experimental evidence in non-US contexts across countries and elections. Second, this is a test of one specific type of treatment. The extent to which the effects of other types of treatments, such as text messages (SMS) or door-to-door canvassing, can help solve the receding turnout rates in Europe and if their effects from the US transfer to Europe are also relevant questions for research.

References

- Angrist, J. D., Pischke, J.-S. & Pischke, J.-S. 2009. *Mostly harmless econometrics: an empiricist's companion*. Princeton University Press, Princeton.
- Arceneaux, K., Gerber, A. S. & Green, D. P. 2006. Comparing experimental and matching methods using a large-scale voter mobilization experiment. *Political Analysis*, 14(1): 37-62.
- Arceneaux, K. & D. W. Nickerson .2009. 'Who Is Mobilized to Vote? A Re-Analysis of 11 Field Experiments, *American Journal of Political Science* 53(1): 1-16.
- Arceneaux, K. & D. W. Nickerson 2010. Comparing Negative and Positive Campaign Messages: Evidence From Two Field Experiments, *American Politics Research*, 38(1): 54-83.
- Bedolla, L. G. & M. R. Michelson 2012. *Mobilizing inclusion: Transforming the electorate through get-out-the-vote campaigns*. Yale University Press.
- Bernstein, R., Chadha, A. & Montjoy, R. 2001. Overreporting voting – why it happens and why it matters, *Public Opinion Quarterly* 65(1): 22-44.
- Beetham, D. 1991. *The Legitimation of Power*. Basingstoke: Macmillan.
- Cantoni, E. & Pons, V. 2016 Do Interactions With Candidates Increase Voter Support and Participation? Experimental Evidence from Italy. *Harvard Business School Working Paper* 16-080.
- Chong, D. & Druckman, J.N. 2007. Framing theory. *Annual Review of Political Science*, 10, 103-126.
- Dahler-Larsen, P. 2011. *The evaluation society*: Stanford University Press.
- Dale, A. & Strauss, A. 2009. Don't forget to vote: text message reminders as a mobilization tool. *American Journal of Political Science*, 53: 787–804.
- Dinan, D. 2015. Governance and Institutions: The Year of the Spitzenkandidaten. *JCMS: Journal of Common Market Studies*, 53: 93-107.

- Downs, A. 1957. *An Economic Theory of Democracy*, Harper & Row
- Enos, R. D. & A. Fowler 2014. Pivotality and Turnout: Evidence from a Field Experiment in the Aftermath of a Tied Election. *Political Science Research and Methods*.
- Enos, R. D., A. Fowler & L. Vavreck 2014. Increasing Inequality: The Effect of GOTV Mobilization on the Composition of the Electorate. *Journal of Politics*, 76(1): 273-288.
- European Parliament 2014. The 2014 European Elections: This Time It's Different. Available: http://www.europarl.europa.eu/pdfs/news/expert/background/20140210BKG35568/20140210BKG35568_en.pdf [accessed 10-08-2016].
- European Parliament 2016. Results of the 2014 European elections / Turnout. www.europarl.europa.eu/elections2014-results/en/turnout.html [accessed 10-08-2016].
- European Youth Forum 2014. Happy Voting! Available: <http://www.youthforum.org/latest-news/happy-voting/> [accessed 10-08-2016].
- Fieldhouse, E., Cutts, D., John, P. & Widdop, P. 2014. When context matters: assessing geographical heterogeneity of get-out-the-vote treatment effects using a population based field experiment. *Political Behavior*, 36(1): 77-97.
- Fieldhouse, E., Cutts, D., Widdop, P. & John, P. 2013. Do impersonal mobilisation methods work? Evidence from a nationwide Get-Out-the-Vote experiment in England. *Electoral Studies*, 32(1): 113-123.
- Foos, F. & John, P. 2016. Parties Are No Civic Charities: Voter Contact and the Changing Partisan Composition of the Electorate. *Polit. Sci. Res. Methods*.
- Franco, A., Malhotra, N. & Simonovits, G. 2014. Publication bias in the social sciences: Unlocking the file drawer. *Science*, 345(6203): 1502-1505.
- Gerber, A. S. & D. P. Green 2000. The Effects of Canvassing, Telephone Calls, and Direct Mail on Voter Turnout: A Field Experiment, *American Political Science Review*. 94(3): 653-663.

- Gerber, A. S. & D. P. Green 2012. *Field Experiments: Design, Analysis, and Interpretation*, W. W. Norton Limited.
- Gerber, A. S. & D. P. Green 2016. Field Experiments on Voter Mobilization: An Overview of a Burgeoning Literature, manuscript.
- Gerber, A. S.; D. P. Green & C. W. Larimer 2008. Social Pressure and Voter Turnout: Evidence from a Large-Scale Field Experiment, *American Political Science Review*, 102(1): 33-48.
- Green, D.P. & Kern, H. L 2012. Modeling heterogeneous treatment effects in survey experiments with Bayesian additive regression trees. *Public Opinion Quarterly*, 76, 491-511.
- Green, D. P., Mary C. McGrath, & P. M. Aronow 2013. Field experiments and the study of voter turnout." *Journal of Elections, Public Opinion & Parties* 23(1): 27-48.
- International IDEA 2015. Voter Turnout Database. I: Idea, I. (ed.). Stockholm: The International Institute for Democracy and Electoral Assistance (IDEA).
- John, P. & Brannan, T. 2008. How Different Are Telephoning and Canvassing? Results from a 'Get Out the Vote' Field Experiment in the British 2005 General Election. *British Journal of Political Science*, 38(03): 565-574.
- John, P., Macdonald, E. & Sanders, M. 2015. Targeting voter registration with incentives: A randomized controlled trial of a lottery in a London borough. *Electoral Studies*, 40(December): 170-175.
- Karp, J. & Brockington, D. 2005. Social desirability and response validity: A comparative analysis of overreporting voter turnout in five countries, *Journal of Politics* 67(2): 825-840.
- Lijphart, A. 1997. Unequal participation: democracy's unresolved dilemma. *American Political Science Review*, 91(1): 1-14.
- Monogan, J. E. 2015. Research Preregistration in Political Science: The Case, Counterarguments, and a Response to Critiques. *PS: Political Science & Politics*, 48(3): 425-429.

- Nelson, T. E., Oxley, Z. M. & Clawson, R. A. 1997. Toward a Psychology of Framing Effects. *Political Behavior*, 19(3): 221-246.
- Nickerson, D. W. 2006. Forget Me Not? The Importance of Timing in Voter Mobilization. Paper presented at the annual meeting of the American Political Science Association, Marriott, Loews Philadelphia, and the Pennsylvania Convention Center, Philadelphia, PA.
- Nickerson, D. W. (2007): Does email boost turnout? *Quarterly Journal of Political Science*, 2: 369-379.
- Nickerson, D. W. 2008. Is voting contagious? Evidence from two field experiments, *American Political Science Review*, 102(1): 49-57
- Nyman, P. 2014. Door-to-door canvassing in the European elections: Evidence from a Swedish field experiment. Uppsala University.
- Panagopoulos, C. 2011. Timing is Everything? Primacy and Recency Effects in Voter Mobilization Campaigns, *Political Behavior*, 33(1): 79-93.
- Panagopoulos, C. 2013. Extrinsic Rewards, Intrinsic Motivation and Voting, *Journal of Politics*, 75(1): 266-280.
- Pons, V. 2016. Will a Five-Minute Discussion Change Your Mind? A Countrywide Experiment on Voter Choice in France. *Working Paper 16-079*. Cambridge, MA: Harvard Business School.
- Pons, V. & Liegey, G. 2016. Increasing the Electoral Participation of Immigrants: Experimental Evidence from France. *Working Paper No. 16-094*. Cambridge, MA: Harvard Business School.
- Ramiro, L., Morales, L. & Jiménez-Buedo, M. 2012. The Effects of Party Mobilization on Electoral Results. An Experimental Study of the 2011 Spanish Local Elections. *IPSA 2012*.
- Riker, W.H. & Ordeshook, P.C. 1968 A Theory of the Calculus of Voting. *American Political Science Review*, 62(1): 25-42.

- Rolfe, M. 2012. *Voter turnout: a social theory of political participation*: Cambridge University Press.
- Rousseau, J.-J. 1762 *Du contrat social ou Principes du droit politique*.
- Sinclair, B. 2012. *The social citizen: Peer networks and political behavior*: University of Chicago Press.
- Statistics Denmark (2012): *Statistisk Årbog 2012*, Copenhagen: Statistics Denmark.
- Tversky, A. & Kahneman, D. 1979. Prospect Theory: An Analysis of Decision under Risk, *Econometrica*, 47(2): 263-291.
- Tversky, A. & Kahneman, D. .1981. The Framing of Decisions and the Psychology of Choice. *Science*, 211, 453-458.
- Vedung, E. 1997. *Public policy and program evaluation*: Transaction Publishers.
- Vowles, J. 2017, forthcoming. The Big Picture: Turnout at the Macro-Level. I: Fisher, J., Cantijoch, M., Fieldhouse, E., Franklin, M., Gibson, R. & Wlezien, C. (ed.) *The Routledge Handbook of Public Opinion and Voting Behaviour*. London: Routledge.
- Winke, M., H. Bless & B. Biller 1996. Subjective Experience Versus Content of Information in the Construction of Attitude Judgments, *Personality and Social Psychology Bulletin*, 22: 1105-1113.

Table 1: Experimental groups

Treatment	Frame direction	Turnout	N
Control group		0.6189	76,761
Benefit argument	Gain	0.6146	8,093
Benefit argument	Loss	0.6263	8,101
Pivotality argument	Gain	0.6193	8,114
Pivotality argument	Loss	0.6189	8,128
Duty argument	Gain	0.6211	8,111
Duty argument	Loss	0.6320	8,113
Cost argument	-	0.6256	8,136
All arguments included	-	0.6298	8,127

NOTE: Turnout is weighted with inverse probability weights

Table 2: Effect of treatments from a linear probability model with inverse probability weights

	Voted	Voted	Voted	Voted
Any letter	0.0041 (0.0027)	0.0048* (0.0024)		
Benefit, gain			-0.0043 (0.0059)	-0.0043 (0.0054)
Benefit, loss			0.0034 (0.0058)	0.0042 (0.0054)
Pivotality, gain			0.0005 (0.0058)	0.0053 (0.0053)
Pivotality, loss			0.0001 (0.0058)	0.0023 (0.0053)
Duty, gain			0.0022 (0.0059)	0.0060 (0.0054)
Duty, loss			0.0132* (0.0058)	0.0118* (0.0053)
Cost			0.0068 (0.0058)	0.0055 (0.0054)
All arguments			0.0110* (0.0058)	0.0073 (0.0054)
Background covariates	NO	YES	NO	YES
Turnout in control group	0.6189	0.6189	0.6189	0.6189
<i>N</i>	141,684	141,684	141,684	141,684
adj. <i>R</i> ²	0.00	0.15	0.00	0.15

Standard errors in parentheses. * $p < 0.05$ in one-sided test.

Figure 1: Predicted voting and predicted marginal effect over propensity to vote

