The calculus of get-out-the-vote in a high turnout setting

Paper prepared for the panel *"23-15 Relationships and Voter Turnout"* at the MPSA Conference, Chicago 3-6. April 2014.

*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

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

This project is primarily funded by the by the Danish Council for Independent Research (grant no. 12-124983). The project has also received funding from the Danish Ministry for Economic Affairs and the Interior.

Abstract

Through a get-out-the-vote experiment we study the effect of eight different statements related to 'the calculus of voting' and prospect theory on voter turnout in a high salience election with proportional representation of multiple parties. The treatments are randomly assigned to more than 60,000 first-time-voters and the treatments are distributed via a personal postal letter in a closed envelope. We find an overall effect of the letter to be 0.22 percentage points (std.error: 0.30), close to the findings from an U.S. context. The most effective of the treatments seems to be a letter including several arguments (ITT: 1.3 percentage points, std.error. 0.62). This supports the idea that citizens are more convinced when receiving multiple arguments. There is no evidence of difference between versions of the letters emphasizing individual elements of the calculus of voting or whether the decision to vote was framed as a gain or a loss. We find limited heterogeneous effects though there is some tendency for the treatment effect to be somewhat larger for voters with low- to middle turnout propensities.

Introduction

Why some people vote, while others abstain from voting has been a central question for both academics and practitioners for decades. If turnout gets too low or the inequalities in participation get too high, it is a threat to democracy (Lijphart 1997). Over the last 15 years we have witnessed a large focus on how to mobilize voters. Scholars have been studying a wide range of aspects such as the effect of campaign form, e.g. direct postal letters, door-to-door canvassing, telephone calls, e-mails and text messages (e.g., Gerber and Green 2000; Gerber, Green & Larimer 2008; Nickerson 2007; Nickerson 2008; Dale and Strauss 2009), if the sender is a partisan or not (e.g., McNulty 2005; Panagopoulos 2006), content (e.g., Gerber and Green 2000; Gerber, Green & Larimer 2008), timing (e.g., Nickerson 2006; Panagopoulos 2011) and even monetary incentives (Panagopoulos 2013) (see Green, McGrath and Aronow 2013 for an extensive meta-analysis). These studies together have taken us a long way in understanding how we can design mobilization campaigns to increase turnout more efficient.

While the body of literature is now extensive and there is a large variation in the treatments investigated, substantially less work has been done to extend the results to contexts outside the U.S. (Green, McGrath and Aronow 2013: 12). It therefore still remains an open question to which degree the American experiences are transferable to other contexts and political cultures. In this paper we conduct a large scale direct mail experiment with more than 60,000 individually treated first-time voters during the Danish Municipal Elections in November 2013. This is a context with high turnout, a proportional election system with multiple parties and automatic registration. Our experimental study is the first ever conducted in a Danish context, and one of the first studies in a European context outside the UK. The election under investigation was highly salient and had an overall turnout of 71.9 percent, whereas most studies are conducted in low to medium salient

elections in the U.S. We thereby have another turnout level, another culture of political participation, and a different institutional context.

Besides investigating the overall effect of direct mail, our large sample size allows us to vary the treatments in eight sub-groups. We take departure in the well-known calculus of voting. In this framework citizens decide whether to vote from the cost, benefits and duty of voting as well as the probability of casting the decisive vote (Downs 1957; Riker & Ordeshook 1968). Many existing studies take departure in a single or several elements of the calculus of voting (Gerber & Green 2000; Gerber, Green & Larimer 2008), though most studies excludes at least one of the elements (usually cost). In this study we include a treatment for all four elements of the calculus of voting. Applying the calculus of voting in our case is important as the effects of the elements may vary across contexts. For instance, races are often closer than in a multiparty system compared to a plurality system. This may make appeals to the pivotality of the vote more credible and thus more effective (see Enos and Fowler 2014 for an experiment in another context with a close race).

In addition to assigning treatments according to the calculus of voting, we varied our treatments in two seldom studied ways. First, we varied the tone using insights from prospect theory. According to this line of work individuals value losses more than gains (Tversky and Kahneman 1981; Arceneaux and Nickerson 2010). This might cause a letter framed as a lost opportunity to be more efficient than one that focuses on a possible gain. Specifically, we sub-divide the benefit, probability and duty treatments in a gain and a loss framing (we did not split the cost treatment as it is difficult to vary the tone of a cost). The experiment thus allows us to evaluate the effect of manipulating the perceived salience of the different aspects of the calculus of voting in a positive or negative manner. Second, findings from experiments within psychology of judgment suggest that subjects are often more convinced when making judgments on the basis many arguments compared to few (Winke, Bless and Biller 1996). This implies that the number of

arguments may matter. We thus include a letter with all of the arguments to investigate whether this increases the effect compared to any of the individual treatments. In total this gives us eight different treatment groups (benefit gain, benefit loss, probability gain, probability loss, duty gain, duty loss, costs and all) and a control group.

Finally, we consider heterogeneous effects of our treatments. A central goal for many GOTV-campaigns is to reduce the inequalities in participation (Bedolla and Michelson 2012). This can potentially be done by getting the underrepresented – e.g. young citizens or immigrants – to vote. Recently, scholars have questioned whether the GOTV-campaigns actually increase or decrease equality in participation (Arceneaux and Nickerson 2009; Enos, Fowler and Vavreck 2014). The basis for this consideration is that the treatment effects are likely to be heterogeneous. It might be that the treatment has a smaller effect on the underrepresented (or low propensity voters) than the overrepresented (middle or high propensity voters). 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. To build on these experiences we test if this GOTV-campaign increases or decreases the inequalities in turnout.

The remaining of this paper proceeds as follows. Next, we present the context of the field experiment and the data used in the study. Thereafter, we discuss the theoretical background for the calculus of voting, how the saliency of the different parameters can be manipulated to increase turnout and how prospect theory can be a useful contribution to this line of thought. Thereafter, we present our research design before we present the results of the study. Finally, we discuss the findings.

Context and data

The context of our study is the Danish Municipal Elections held on November 19, 2013 across the 98 different Danish municipalities[1]. This is a setting with a high salient election, a context of multiple parties competing in a proportional election in a media setting dominated by a public service broadcast with much emphasis on politically balanced news. The 98 Danish municipalities are central in the government service provision and more than 29% of the total Danish GDP (2010) is spend at this government level (Statistic 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. Individuals member of the municipal boards are mostly selected from open lists. The turnout average across all 98 municipalities in the 2013 elections was 71.9 percent (non-compulsory voting). This corresponds to an increase in turnout on 6.1 percentage points from the latest elections in 2009. During the last decades, the turnout has fluctuated around 70 percent municipal elections, which is somewhat lower than national elections that have fluctuated around 85 percent. Turnout among 18-21 years old citizens, which are in focus in our experiment, was 61 percent. The news media pay much attention to the local elections in Denmark, and the national politicians and national party are very active in the election campaign. In sum, it is fair to say that the context of the study is a highly salient election.

All Danes have a personal ID number (social security no.). This ID number is linked to the personal address of all Danes. We use this information to randomize individuals into our treatments and control group allowing us to have a high level of control over the process. In the experiment, we register treatment and turnout at the individual citizen level using the personal ID

[1] The election for the five regional governing bodies of Denmark where 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.

numbers from the official Danish registers. Thus, we have accurate and precise information on each voter's treatment status (address and household information) and turnout where measurement error is extremely low without any self-selection or over reporting (Karp & Brockington 2005; Bernstein et al. 2001). In addition, we have a large number of variables from high quality government registers at our availability through *Statistics Denmark*, the official statistic bureau. Among the variables that we are able to merge with the individual level voting and treatment status are: Gender, age (in days), completed and ongoing education, composition of the household, ethnicity, employment, where the person live (municipality) and previously national election eligibility. This provides us with strong predictors of turnout, which in turn yields good opportunities to study heterogeneous effects of the treatments. We use these variables to predict an individual propensity score of voting and assess whether this propensity score is linked to our treatment effects.

Theoretical background

As stated above, our theoretical point of departure is the calculus of voting, where an individual decides to vote if the perceived utility is positive (Riker and 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.). Finally, D is the feeling of duty towards voting. This includes

the satisfaction gained by performing ones civic duty of voting by e.g. living up to the democratic norm of participation by voting.

While it is not possible for researchers to alter the actual values in the calculus[2], 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). Whether the arguments put forward in the treatment are new to the citizens or already heard before, reading the arguments in itself 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 then be a more positive attitude towards voting. The essential question is whether this causes a change in behavior. If this is the case, it should lead to an increase in turnout. Several studies have examined the difference in effects depending on which element of the voting calculus one appeals to and generally find no differences across treatments (Gerber & Green 2000; Green, McGrath and Aronow 2013).

In addition to the calculus of voting, we draw on prospect theory. The idea behind this is firstly 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 and Druckman, 2007). Prospect theory is often applied in a lab or survey experimental design, where the dependent variables are attitudes or intentional behavior. In the classic studies, it is shown that people tend to value losses more than gains, which imply that our loss frame should have the largest effect (Tversky and Kahneman 1979; Tversky and Kahneman 1981). By applying the theory in a field experiment, we are able to evaluate whether the

[2] One could however identify certain elections where e.g. one's chance of casting the decisive vote would be extraordinarily high (Enos & Fowler 2014).

gain/loss framing actually makes a difference for real world voting behavior. In other words, we expect that arguments referring to loss of democratic influence are more effective in increasing turnout than arguments referring to increased democratic influence by voting. One study has looked at this before in an American context. Arceneaux and Nickerson (2010) in two experiments using personally delivered message find no systematic difference in the effect between positive and negative messages on turnout. In this study we re-examine the effects in the new context using a non-personal form of delivery (direct mail).

We also test if the quantity of arguments in itself makes a difference. The idea with this is, simply, that there might by a synergetic effect from multiple arguments. Individuals may draw inferences about the quality of a position based on the quantity of arguments (Winke, Bless and Biller 1996). In other words, the number of arguments can serve as a cue for the strength of a recommendation (Petty and Cacioppo 1984). Additionally, people may be convinced by different types of arguments. Multiple arguments in one treatment could make it more likely that at least one argument among many will motivate the potentially first-time voter to go the polling station. On the other hand, information overload may lead individuals to not read the letter or to miss the most effective argument due to the large number of arguments (drowning out effect).

Finally, we will examine potential heterogeneous treatment effect. One of the main motivations behind many GOTV-studies, including ours, is a desire to decrease inequalities in turnout (e.g. Bedolla & Michelson 2012). Enos et.al. (2014) find a significant interaction effect between propensity to vote and treatment in a range of GOTV studies in an American context. Their finding suggests that mobilization efforts to some extent have a larger impact on high propensity voters. In another meta-analysis Arceneaux and Nickerson (2009) show that it is habitual voters who most likely are mobilized in low salience elections, whereas the low propensity voters to a

higher degree can be mobilized by campaigns in high salience elections. In middle salience elections the authors concludes that it is the occasional voters who are most likely to be mobilized.

Our experiment targeted first-time voters whom we know from previous research have a substantially lower turnout rate than the rest of the population in several countries including Denmark (Bhatti, Hansen & Wass 2012). Even within the group of first-time voters, we expect substantial variation in turnout. E.g. the predicted probability of voting in the last election for a female first-time voter who lived with both her parents was only 14 percent if none of her parents voted compared to a predicted probability of voting of 73 percent if both her parents voted (Bhatti & Hansen 2012: 393). In addition to that, we have a long list of other social demographics predictors of turnout used as control variables. If a covariate predicts turnout it implies that there is an inequality in turnout based on that covariate. By using to control group we will estimate a propensity turnout score based on a number of citizen characteristics. The propensity score will then be interacted with the treatments. One major advantage for our study is the high quality data at our availability, which makes it possible to estimate a very reliable propensity score. Furthermore, the study takes place in a high salience election, which is an important addition to the existing literature. All of the GOTV-campaigns that Enos et al. (2014) show to increase inequality in participation is from American elections with much lower turnout than at our elections, which imply that the distribution of propensity to vote is substantially different in our setting. By examining the interaction between treatment and propensity to vote, we hope to build on the existing knowledge of who are mobilized.

Research design and treatment

The target group for the field experiment is young first-time municipality voters aged 18-21 on Election Day in all 98 municipalities in Denmark. We randomly excluded a part of this age group for use in other experiments. The composition of the individual household can influence the effect of various GOTV-treatments. Nickerson (2008), for instance, shows how voting can be contagious within two person households, and Fowler (2005) finds that turnout can spread through social networks in the form of a "turnout cascade," as a single person's turnout decision affects the decisions of at least four other people as to whether or not to vote (Christakis & Fowler 2009).

In order not to contaminate our experiment with these effects, we next split our sample into two samples as suggested by Geber et al. (2012:24). Sample 1 consists of all voters who did not live with other 18-21 year olds. For these individuals there is no risk of intra-household contamination and we therefore just use simple random sampling. Sample 1 also included households with 3 or more other 18-21 years old on Election Day. The choice of keeping households with 3 or more voters in sample 1 was based on the assessment that these household sizes are more likely to be like co-opts where the interaction between each person are more limited than in small households with 2 18-21 year olds (often siblings or roommates). These voters only constitute a minor proportion of the sample. We tried to run our analysis below where we restricted sample 1 to households with only one first-time voter and it did not affect the results substantially. Sample 2 consists of voters who lived with 1 or 2 other 18-21 year old voters three weeks before Election Day. If we just used simple random sampling here, we would risk intra-household contamination. Thus, in sample 2 we treated half the household and within these, we only treated one randomly selected person. In the treated household, we removed the untreated individuals from the sample. Finally, we removed individuals who lived in households with individuals that we treated in other experiments. This leaves us with a total sample size of 104,912 in sample 1, 20,610

in sample 2 and 125,522 combined. Due to the different sampling scheme we analyze the two samples separately.

To answer the question in this study we sent out 60,546 personal letters in a closed envelope to potential first-time voters. The receiver's name was typed on the as well as in the letter. The voters received the letter either four or five days before Election Day. Our letters are best termed as conventional, nonpartisan letters with the Ministry of Economic and Interior official letterhead. Such letters have been shown to have a small, albeit, positive treatment effect with a point estimate of 0.194 percentage point and a 95% confidence interval of (0.106; 0.282) (Green et.al. 2013: 31). The experiment was divided into the two groups described above and we conducted the randomization of treatment independently in the two groups and assigned identical treatments. This gives us an opportunity to assess whether findings in one group are robust in another.

Each letter followed the same template. This included the receiver's own full name followed by a short, general text about the election. After this a headline in **bold letters** with a short argument, defined by the treatment group, appeared. Finally some practical information about the voting process was included, before the treatment headline 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 appendix A).

There were eight different treatments and a control group in the experiment, as presented in table 1:

Table 1: Experimental design

Random group number	Treatment focused on	Frame direction
Group 1a	Civic Duty argument	Gain
Group 1b	Civic Duty argument	Loss
Group 2a	Benefit argument	Gain
Group 2b	Benefit argument	Loss
Group 3a	Probability argument	Gain
Group 3b	Probability argument	Loss
Group 4	Cost argument	-
Group 5	All arguments included	-
Group 6	Control group – no treatment	

There were two treatment groups for three of the four factors in the calculus of voting. We framed each of these arguments as either a gain or loss. The substance of the arguments within each of the factors where equivalent. For instance, group 2A received an argument focused on the potential benefits from voting framed in terms of potential gains. The headline in this letter said "**Take part in deciding your everyday - vote**". Furthermore, the text stated that "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, group 2b received the benefit argument framed in terms of potential loss. The headline in this letter was "**Do not let others decide your everyday - vote**". Inside the text it was noted "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. Avoid losing your influence – vote!"

Regarding the cost factor in the calculus of voting only one treatment group received this type of argument. The reason for this is that it seemed to be rather nonsense to frame the cost as a gain/loss. The cost letter was focused on reducing the perceived costs of voting. The opposite argument would then focus on increasing the perceived costs of voting, which indeed would be

destructive with regards to the goal of increasing turnout. Therefore the division between gain and loss framing was not used in relation to the cost argument.

Finally, group 5 received all the arguments in combination. Concretely the letter consisted of arguments focused on reducing the costs of voting, increasing the perceived benefits and probability of being decisive in the election and raising the feeling of performing ones civic duty by voting. Group 6 was used as a control group and they did not receive any letters (see the appendix for the full version of all the letters).

Of the more than 60,000 letters sent only one percent was returned as failed delivery. We opened the returned letters and could confirm that there was no systematicity in which letters were returned, an indication that randomization worked accordingly to our intentions. 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.

Results

Table 1 presents the actual turnout for each experimental group in the two different samples. Turnout in the control group of sample 1 is 3 percentage points higher than in the sample 2 control group.[3] This is not surprising since a larger proportion live with their parents in sample 1 than in sample 2, and we know from previous research that living with one's parents is a strong predictor of 18-21 year olds turnout (Bhatti and Hansen 2012). The eight individual treatments are displayed individually in row 2-9 in table 1. In row 10 they are all merged together, in row 11-13 the different framings are merged together by message content for the cost, duty and benefit arguments, and in

[3] A two-sample test of proportions show that this difference is significant at $p<0.0001$.

row 14-15 they are merged together by the gain/loss framing direction irrespective of message content.

Table 2: Turnout across experimental groups (percent, n)

	Sample 1 One first-time-voter in household	**Sample 2 Two or three first-time-voters in household**
Control group	61.1 (51,084)	58.1 (13,892)
Benefit. Gain	60 (6,717)	59.7 (841)
Benefit. Loss	61.2 (6,723)	60 (837)
Probability. Gain	61.1 (6,723)	56.6 (838)
Probability. Loss	60.8 (6,730)	59.1 (845)
Duty. Gain	61.5 (6,719)	59 (843)
Duty. Loss	61.7 (6,729)	59.5 (837)
Cost	61.7 (6,747)	58.7 (837)
All arguments	62.4 (6,742)	59.0 (840)
Any treatment (merged)	61.3 (53,828)	58.9 (6,718)
Benefit (merged)	60.6 (13,438)	59.8 (1,678)
Probability (merged)	61 (13,453)	57.9 (1,683)
Duty (merged)	61.6 (13,448)	59.2 (1,680)
Gain (merged)	60.9 (20,159)	58,4 (2,522)
Loss (merged)	61.3 (20,180)	59.5 (2,519)

In table 3 we present our main findings. The coefficients in the table are marginal effects from a logistic regression of turnout on treatment. The first two columns display the effects in the two samples when we pool all our treatments together. Column three to six displays the

treatment effects of the individual treatments in the two samples where we control for a list of covariates in column 5 and 6. These control variables include gender, age (in days), country of birth, completed and ongoing education, composition of the household, occupation, income, municipality residence, a four-way interaction between each parent's turnout at the 2009 municipality election and whether or not the voter lived with her parents, and an indicator for whether the young voter was eligible to vote at the Danish 2011 General Election.

Table 3: Marginal effect of treatments (Logistic Regression Models)

	Pooled treatment effects		Effects of different treatments		Effects of different treatments	
	Sample 1	Sample 2	Sample 1	Sample 2	Sample 1	Sample 2
Any treatment	0.00222	0.00863	-	-	-	-
	(0.00301)	(0.00733)	-	-	-	-
Benefit, gain	-	-	-0.0115	0.0161	-0.0106	-0.00452
	-	-	(0.00636)	(0.0174)	(0.00561)	(0.0156)
Benefit, loss	-	-	0.00149	0.0189	0.00571	0.0198
	-	-	(0.00632)	(0.0174)	(0.00556)	(0.0154)
Probability, gain	-	-	0.000270	-0.0152	0.00107	-0.00615
	-	-	(0.00632)	(0.0176)	(0.00557)	(0.0158)
Probability, loss	-	-	-0.00289	0.0109	-0.000321	0.00260
	-	-	(0.00633)	(0.0174)	(0.00558)	(0.0156)
Duty, gain	-	-	0.00450	0.00872	0.00653	0.0216
	-	-	(0.00631)	(0.0175)	(0.00555)	(0.0154)
Duty, loss	-	-	0.00626	0.0141	0.00858	0.0158
	-	-	(0.00631)	(0.0175)	(0.00555)	(0.0156)
Cost	-	-	0.00684	0.00578	0.00217	0.00739
	-	-	(0.00630)	(0.0175)	(0.00555)	(0.0156)
All arguments	-	-	0.013*	0.010	0.0119*	0.0130
	-	-	(0.00628)	(0.0175)	(0.00554)	(0.0156)
Control variables	No	No	No	No	Yes	Yes
n	104,912	20,610	104,912	20,610	104,912	20,610
Pseudo R^2	0.0000	0.0000	0.0001	0.0001	0.1148	0.1307

NOTE: Entries are average marginal effects. Standard errors in parenthesis. *p<0.05, two-sided test compared to the control group. I.e. the control group is reference category.

When we pool all our treatments together in column 1 and 2 there is no significant effect in either sample. The pooled treatment effects in sample 1 were 0.22 percentage points and 0.86 percentage points in sample 2. The first point estimate is within the confidence interval based on previous research (0.106-0.282 percentage points), the second is above but with substantial uncertainty in the estimate (Green et.al. 2013: 31). Thus, regardless of the difference in institutional characteristics and culture of participation, the results are of similar magnitude as in the U.S.

In model 3 and 4 we examine the individual treatments. In the large sample only the letter with all arguments had an individually statistical, significant impact on turnout. In sample 2 the individual treatment groups were smaller, which makes the study somewhat underpowered with respect to examining individual treatments. The results in model 3 suggests that there is a synergetic effect of receiving a longer letter with several types of arguments instead of a short letter with just one argument. Recipients of this letter had a 1.3 percentage point higher turnout rate than the control group. However, we should note that the effect is just below a 0.05 threshold for significance. If we look at sample 2, we do however find a coefficient of similar size, though it is far from significant due to the substantially lower number of treated in sample 2 compared to sample 1.

We also collapsed our treatments from table 3 together by type (duty, probability etc.) and framing (gain/loss). We saw no significant differences in effects between different message contents: p=0.35 in sample 1, the highest point estimate compared to the control group is 0.68 percentage points for the costs treatment and the lowest is -0.50 percentage points for the benefit treatments – in sample two, p=0.62 with the highest being the benefit treatment and the lowest the probability treatment. We did not find significant differences for frame tone either: the effect estimate in sample 1 is 0.39 percentage points higher (p=0.43) for the negative framing compared to

the positive framing and in sample 2 the difference is 1.14 percentage points (p=0.41). These insignificant results are consistent with the conclusions from existing studies.

Finally, when we include our list of covariates we do not see any significant changes to our results. Since we randomized treatment, the effects should be unbiased even without control variables and the introduction of covariates should not give systematic change in the parameters but only tend to lower standard errors and cause more efficient estimates (Gerber and Green, 2012: 104). This is in line with what we see here and all differences between the model with and without controls variables are very small and insignificant. At the same time, it is still only the effect of the letter with all arguments in the larger sample 1 that is statistical significant.

Is there an interaction between propensity to vote and treatment?

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. To see if our experiment followed the finding from an American context where high propensity voters are more likely to be mobilized and to investigate if we had any success in leveling out differences within the group of first time voters, we next turn to examining if our treatment had a different impact conditional on propensity to vote.

What we want to do is to search for heterogeneous treatment effects. However, to search for heterogeneous effects of treatment is a dangerous venture as you do multiple comparisons with a large amount of discretion for researchers (Green and Kern, 2012). To minimize this problem we follow Enos et. al. (2014) and estimate a propensity score for turnout in our experiment. 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 have shown to predict turnout well among first-time voters at Danish elections (Bhatti and Hansen 2012). Furthermore, the control is identical to the list of covariates applied in table 2. If our treatment had heterogeneous effects conditioned on propensity to vote we should expect an interaction between our treatment and the estimated propensity score.

Enos et.al. (2014) note that it seems most sensible to only test for heterogeneous effects where the overall treatment effect is positive, since an interaction between an average null effect and propensity to vote would imply that we would demobilize some voters with the treatment. Following this line of thought, we choose to narrow our analysis to the only treatment that seemed to work; the letter where we delivered all arguments to turnout simultaneously. In table 4, we present the results from this analysis.

Table 4: Interaction effects

	Sample 1	Sample 2
	One first-time-voter in household	**Two or three first-time-voters in household**
Control group	-	-
Benefit, gain	0.0802	0.308
	(0.103)	(0.272)
Benefit, loss	0.117	-0.0858
	(0.101)	(0.267)
Probability, gain	-0.0262	0.338
	(0.104)	(0.240)
Probability, loss	-0.0178	0.233
	(0.104)	(0.272)
Duty, gain	-0.0145	0.349
	(0.106)	(0.239)
Duty, loss	0.270**	0.221
	(0.101)	(0.246)
Cost	0.202	0.0528
	(0.104)	(0.266)
All arguments	0.192	0.231
	(0.104)	(0.258)
Propensity score	4.634***	4.711***
	(0.0566)	(0.104)
All arguments * Propensity score	-0.243	-0.286
	(0.166)	(0.434)
Constant	-2.338***	-2.397***
	(0.0355)	(0.0636)
n	104,912	20,597

NOTE: Entries are logistic regression coefficients. Standard errors in parenthesis. *p<0.05, **p<0.01, ***p<0.001, two-sided test.

As is evident from table 4, both the interaction terms are insignificant but of similar size and direction. To get a better grip of the patterns of the interaction we look at figure 1 and 2. In figure 1 we display for sample 1 the marginal, predicted probabilities from a logistic regression of turnout on propensity to vote interacted with treatment in the left panel and in the right panel we display the marginal treatment effect in the two samples. Each plot includes a rug plot where we have plotted each percentile from the distributions of propensity to vote including the maximum and minimum value. Figure 2 displays the same for sample 2.

Figure 1: Interaction effect: Propensity to vote and the all arguments treatment for sample 1

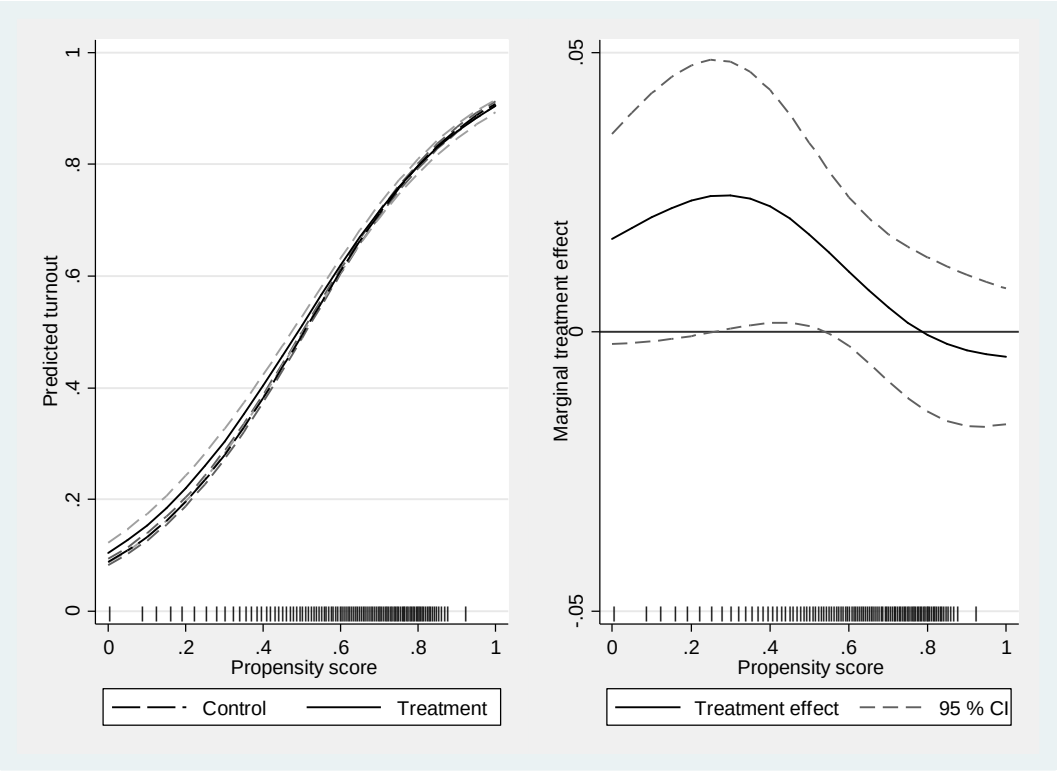

Figure 2: Interaction effect: Propensity to vote and the all arguments treatment for sample 2

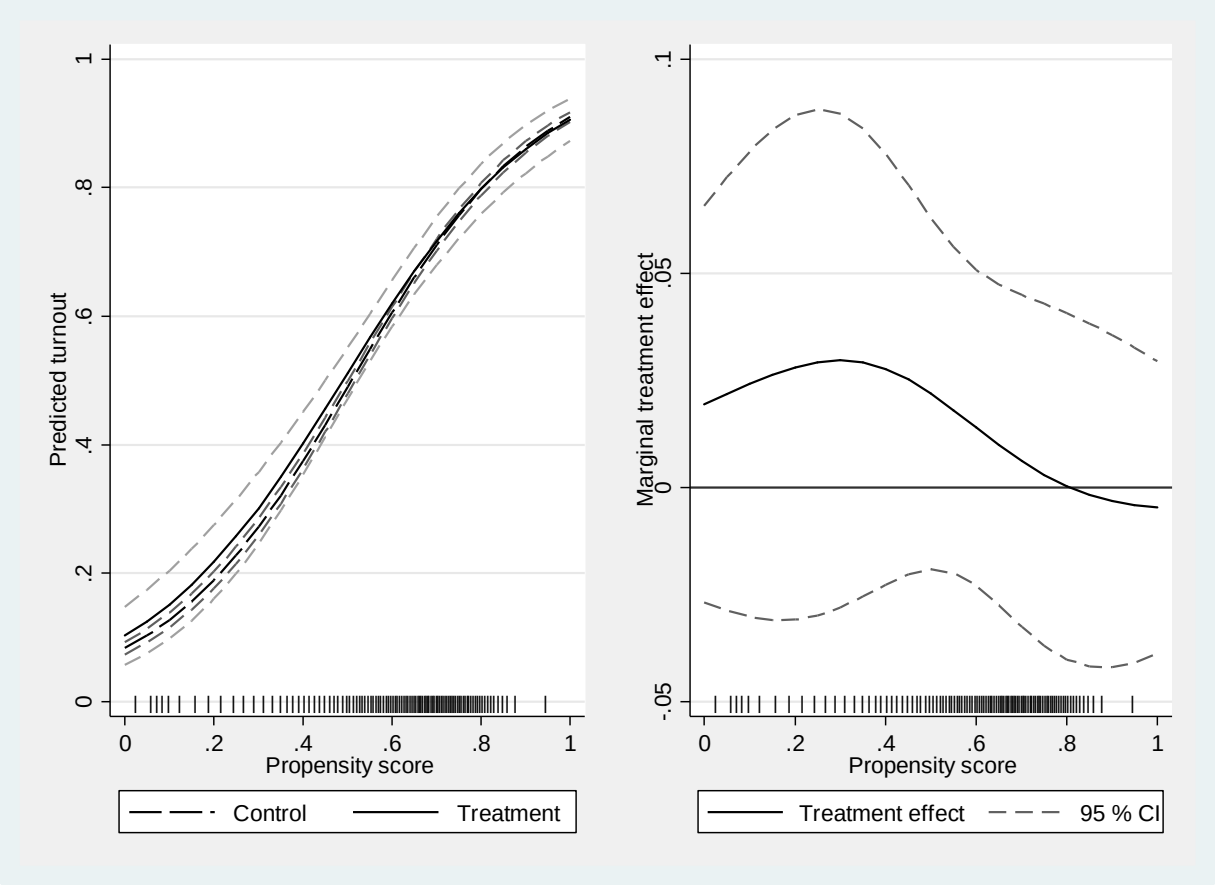

Figure 1 shows that the marginal effect from the treatment follows a curvilinear pattern where the treatment effect tops at around a propensity score of 0.4 and in general is higher for the least likely voters. The effects are still rather small and estimated with a high degree of uncertainty compared to their size, so we should not be to firm in our conclusion. In figure 2, we see a similar pattern. Due to the smaller power of this study, these effects are however far from precise enough for us to say anything with certainty.

To sum up on the interaction between the all treatment and the propensity to vote, the results do suggest that the effort was more efficient in mobilizing voters with a low propensity to vote than high propensity voters. Especially when considering the fact that the lower quartile for propensity to vote was 0.48 in sample 1 and 0.51 in sample 2.

Conclusion

In this article, we have investigated the effect of direct postal letters in a new context. The Danish 2013 municipal elections were conducted in a multiparty system with proportional representation, automatic registration and a high turnout. The overall effects of the treatments were statistically insignificant with point estimates 0.22 percentage points and 0.86 percentage points for all treatment letters which is close to the findings from a U.S. context.

Eight different versions of the letter were assigned randomly to first time voters between 18 and 21 years old. We found no difference in the treatment effects depending on which of the arguments in the calculus of voting was applied. Likewise, we found no difference between the treatments that, applying insights from prospect theory, used a positive or a negative framing. However, one treatment did stand out. Individuals receiving letters with all arguments had a 1.2-1.3 percentage points higher turnout than the control group and the highest point estimate of any of the treatments. This lends support to the idea that more arguments can be more convincing than few, at least with the amount of information present in the treatment letters used in this study. This suggests an additive effect of arguments and the individual might be able find the argument among many arguments that influence them the most. Furthermore, it also suggests that we should not worried about providing multiple arguments in order to expect the largest mobilization effect in postal letters to the 18-21 years old. The result calls for further studies of the importance of the number of arguments on the treatment effects, including when information overload kicks in.

Finally, we do find some evidence for a curvilinear relationship between the individuals' propensity of voting and the treatment effect. Our power is too low to say anything conclusive, but the results do suggest that the treatment effect would probably have been higher in a lower salience election.

A central question that we raised in the beginning of this paper is to which degree the findings from the primarily US-based studies can be transferred to a quite different context. As such the existing non-U.S. GOTV-studies are few, which in some ways is ironic and problematic. Ironic because that many policy makers and organizations outside U.S. uses the same techniques and refer to the American experiences with these. Problematic, as the same organizations do not have any near solid evidence that the U.S.-findings can be transferred to a context like the Danish, with high turnout, high saliency, multiple parties and proportional representation. Our study shows that the findings with regards to the use of direct, non-partisan mailings seem to work in approximately the same degree as the average American experience with this type of treatment. In this way our study suggests that the findings do seem to be transferable. However we must stress two important points. Firstly, this is only a single study. There is surely a need for more field experimental evidence in non-US contexts across countries and elections. Secondly, this is a test of one specific type of treatment. To which extent the effects of other types of treatments, e.g. text messages (SMS) or door-to-door canvassing, can be transferred surely is a question for future research.

References

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

Arceneaux, Kevin and David W. Nickerson (2010): “Comparing Negative and Positive Campaign Messages: Evidence From Two Field Experiments, *American Politics Research,* 38(1): 54-83.

Bedolla, Lisa Garcia, and Melissa 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.

Bhatti, Yosef, Kasper M. Hansen, and Hanna Wass (2012): "The relationship between age and turnout: a roller-coaster ride." *Electoral Studies* 31(3): 588-593.

Dale, Allison & Strauss, Aaron (2009) Don’t forget to vote: text message reminders as a mobilization tool. American Journal of Political Science, 53: 787–804.

Downs, A. 1957. *An Economic Theory of Democracy*, Harper & Row

Chong, D. & Druckman, J.N (2007). Framing theory. *Annual Review of Political Science,* 10**,** 103-126.

Enos, Ryan D. and Anthony Fowler (2014): Pivotality and Turnout: Evidence from a Field Experiment in the Aftermath of a Tied Election. Political Science Research and Methods.

Enos, Ryan D., Anthony Fowler and Lynn Vavreck (2014): “Increasing Inequality: The Effect of GOTV Mobilization on the Composition of the Electorate”. The Journal of Politics, 76: 273-288.

Fowler, J.H. (2005): Turnout in a Small World. In A.S. Zuckerman (ed.), *The Social Logic of Politics: Personal Networks as Contexts for Political Behavior*. Philadelphia: Temple University Press.

Gerber, Alan S. and Donald P. Green (2000): The Effects of Canvassing, Telephone Calls, and Direct Mail on Voter Turnout: A Field Experiment, 94(3): 653-663.

Gerber, Alan S. og Donald P. Green (2012). *Field Experiments: Design, Analysis, and Interpretation,* W. W. Norton Limited.

Gerber, Alan S.; Donald P. Green og Christopher 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, Donald P., Mary C. McGrath, and Peter M. Aronow (2013): "Field experiments and the study of voter turnout." *Journal of Elections, Public Opinion & Parties* 23(1): 27-48.

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, Arend (1997): Unequal participation: democracy's unresolved dilemma. *American Political Science Review*, 91: 1–14.

McNulty, John E. (2005): "Phone-Based GOTV - What's on the Line? Field Experiments with Varied Partisan Components, 2002-2003.", *The Annals of the American Academy of Political and Social Science*. 601: 41-65.

Nelson, T.E., Oxley, Z. M. & Clawson, R.A. (1997). Toward a Psychology of Framing Effects. *Political Behavior*, 19**,** 221-246.

Nickerson, David 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, David W. (2007): Does email boost turnout? Quarterly Journal of Political Science, 2: 369–379.

Nickerson, David W. (2008): "Is voting contagious? Evidence from two field experiments", *American Political Science Review,* 102(1): 49–57

Panagopoulos, Costas (2009): "Partisan and Nonpartisan Message Content and Voter Mobilization: Field Experimental Evidence", *Political Research Quarterly* 62(1): 70-77.

Panagopoulos, Costas (2011): "Timing is Everything? Primacy and Recency Effects in Voter Mobilization Campaigns," *Political Behavior*, 33: 79-93.

Panagopoulos, Costas (2013): Extrinsic Rewards, Intrinsic Motivation and Voting, *Journal of Politics*, 75(1): 266-280.

Petty, Richard E. and John T. Cacioppo (1984): "The Effects of Involvement on Responses to Argument Quantity and Quality: Central and Peripheral Routes to Persuasion", *Journal of Personality and Social Psychology*,46, 69-81.

Riker, W.H. and Ordeshook, P.C. (1968): A Theory of the Calculus of Voting. *American Political Science Review,* 62: 25-42.

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.

Winke, Michaela, Herbert Bless and Barbara Biller (1996): Subjective Experience Versus Content of Information in the Construction of Attitude Judgments, *Personality and Social Psychology Bulletin*, 22: 1105-1113.