

Evaluating the effect of government communication towards young citizens
– a randomized wait-list field experiment

Paper prepared for workshop 26: Causal effects in political science: The promise and pitfalls of experimental methods at the NOPSA Conference, Gothenburg, 12-15 August 2014.

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

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 Parliament (*Folketinget*).

*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.

Abstract

A minor adjustment of an ongoing government information campaign to increase turnout, allowed to evaluate the campaign's efficiency with strong internal and external validity. Since 2005 the Danish government has sent the constitution along with a letter in a closed envelope to all Danes when they turn 18. At the 2013 municipal elections, the campaign was adjusted using a randomized wait-list field experimental design, which exploits delay of the intervention to estimate causal effects. Our target population was 27,550 young voters between 18 and 18½ years old on Election Day. This group was randomly divided into three groups. Two treatment groups received either a traditional or a humorous letter, in both cases along with a copy of the constitution from the Danish Parliament before the election. The control group received a treatment after Election Day. Turnout was accurately measured using voter files of actual voting and linked with government administrative data from the national statistics bureau. The general effect of the traditional and humorous letter was 0.6 and 1.6 percentage point. The effect among voters with low propensity to vote was as high as 5 percentage points while the effect was indistinguishable from zero among high propensity voters, suggesting that the mobilization efforts has decreased the gap between low and high propensity voters.

Keywords: Field experiment, GOTV, randomized wait-list design, turnout, mobilization, young voters, information campaigns, nudging

Introduction

Many public policy initiatives are designed to influence the behavior of the citizens. This is indeed the goal of large government information campaigns, where there are no changes to the current law. Such campaigns can either aim at giving the citizens more information to make a better choice for themselves or they can encourage citizens' behavior in a given direction. The campaigns are used within a range of policy areas (e.g. health, environment, traffic), where new laws often are not perceived as appropriate tools or sufficient to achieve a change in behavioral outcome (see also Stutzer et al. 2011; Costa & Kahn 2013; Fischer et al. 2013; Phillips et al. 2011). Several of the problems that governments are trying to solve with the use of information campaigns are recurring in their nature. For example, new citizens will grow up and become potential smokers, drive too fast in their cars or, as in our case, become non-voters. With this in mind, it is a central concern how we can evaluate existing government programs in a valid and reliable way. Such evaluations can be used to inform the design of coming programs and thereby increase the cost-efficiency of government campaigns.

However, making causal inferences about existing government programs is difficult, both for government itself and for researchers. The campaigns are often designed and implemented as reach-all campaigns. This means that the counterfactual – what would have happened in the absence of a campaign – is impossible to assess. If a government campaign sets out at time 0 and we observe a change in behavior at point 1, we cannot be certain that the behavioral change was caused by the campaign or something else (an unobserved variable). Even more challenging, when we are faced with existing, long-running government programs it is often difficult even to observe the change over time because we have no measurement of the outcome of interest before the program was initiated. A central question is therefore how existing government campaigns can be designed to make it possible to evaluate the effects in a more valid way.

In this article, we present field experimental evidence on the causal effect on turnout from a government program that have been running since 2005. In field experiments, the target group is randomly distributed into one or several treatment groups and a control group that receives no treatment. As the assignment into control and treatment groups is random, the only systematic difference between the two groups in expectation will be whether they were targeted by the campaign. Any difference between the groups regarding their behavior (the dependent variable) can therefore be assigned to the intervention/campaign (the independent variable) with a high degree of certainty. The Danish parliament has since 2005 mailed the Danish constitution to all citizens when the turn 18. Part of the motivation for this have been to increase turnout among young voters (Regeringen & Dansk Folkeparti 2004). In the experiment we present here, we were able to randomly assign if voters that recently had turned 18 received the constitution before or after Election Day. Two groups received the constitution prior to the election along with either a standard letter or a more humorous cartoon. The control group received the constitution after the election.

One challenge with standard field experiments in public administration can be that public agencies are reluctant to exempt a random part of the population from receiving a potentially beneficial treatment (Cotterill & Richardson 2010). In this study, we apply a variant of the field experiment where we consider this concern. More specifically, we apply a randomized wait-list field experiment (e.g. Gerber & Green 2012: 276-81; Brown et al. 2006: 260) where all individuals receives the treatment, but where the different (random) timing of the treatment allowed to make unbiased causal inferences.

The article adds to the growing literature applying field experimental methods to evaluate government policy initiatives in a manner that combine high internal and external validity (Boruch 2005; Cotterill et al. 2009; Cotterill & Richardson 2010; Haynes et al. 2012). Furthermore, because the dependent variable of the study is voter turnout the article also aims to contribute to the

get-out-the-vote literature regarding the effects of direct mail on voter turnout (Green et al. 2013). More specifically, we investigate a new context and consider the difference between messages using humor compared to messages that are more traditional.

The main finding is that, on average, receiving the constitution and an encouragement from the parliament four days ahead of the election had a statistically significant positive effect of 1.1 percentage point on turnout. This is a remarkable effect for two reasons. First, the baseline turnout (in the control group) was already as high as 72.5 percent, thus more than seven out of ten citizens in the control group voted in the absence of receiving the package. Secondly, there was a general increase in the 18 years old turnout by 13.2 percent compared to the previous election. Thus, the 1.1 percentage point is on top of a major general mobilization of the young voters in the election. The letter applying a humorous tone had a significant positive treatment effect of 1.6 percentage point, while the other more traditional letter had an insignificant effect of 0.6 percentage point. This is valuable extra information with regards to how one can design such letters in the future. Across different subgroups, the largest increase is among the low propensity voters where the effects approaches five percentage points while the effect is indistinguishable from zero among high propensity voters.

The next section contains a theoretical introduction to the relevant literature that has guided the design of the letters. Following this, we present the research design and the data used in the analysis that follows in the next section. Finally, we sum up the findings in the conclusion before we provide some relevant perspectives on how this study might inspire public communication campaigns in the future.

Theoretical background

Since one of the purposes of the government program under investigation is to increase turnout, we apply insights from the get-out-the-vote (GOTV) literature (e.g. Gerber & Green 2000; Gerber & Green 2001; Gerber et al. 2008; Nickerson 2008). Several studies have investigated the effect of letters on the propensity to vote. In a meta-analysis across 79 non-advocacy mailings Green et al. (2013) find a positive, but small average effect of 0.194 percentage points (95 percent CI: 0.106; 0.282). Even though the literature is vast and methodologically well founded, our study contributes in two ways to the knowledge on GOTV campaign effects.

First, most of the existing studies are from a US context (e.g. Green et al. 2013: 12-13) and there is a need to investigate whether the results hold in settings different from the US. The context of our study differs in several ways from the US. Most important turnout is much higher. Thus, in the population under investigation in this study, 18-18½ year olds at the Danish 2013 Municipal Elections, the turnout rate was an impressive 72.5 percent in the control group. This may limit the potential for finding effects of the treatment, as most of the target population would vote regardless of treatment status. In that respect, we run a hard test of the established findings. Another difference is that the Danish electoral system is a multiparty, PR system where registration to vote is automatic.

Secondly, we compare a traditional letter from an authority with a more eye catching and humorous cartoon. While humor as a tool has been applied in other GOTV-studies (Nickerson 2007), we are not aware of any previous study designed to test the effect of traditional versus humorous treatments.

The use of humor in communication campaigns is not new, even though it has not been studied much in the GOTV-literature. Advertisers use humor routinely, and public campaigns

have tried this tactic before, too. For instance, a field experiment on melanoma prevention showed that more people read the humor leaflet, but the average knowledge level was not higher in the humor group compared to the group receiving a neutral information leaflet. Both groups had a higher knowledge level than the control group in this study (Richard et al. 1999). Furthermore, several studies of political comedy suggest that there is a positive relationship between exposure to satire and political knowledge and further information acquisition (Cao 2010; Xenos & Becker 2009; Young & Hoffman 2012) as well as a positive impact from watching political comedy on political self-confidence (Brewer et al. 2013). As these outcomes to a considerable extent are correlated with turnout, it is of interest whether the use of humor is a viable way for encouraging young citizens to vote.

Building on the existing research, we put forward five hypothesis of the effect of the campaign. First, we expect that there is a positive overall treatment effect. It should have a positive effect to receive a letter with an encouragement to vote no matter the details in the letter compared to not receiving one (cf. Green et al. 2013). Furthermore, we expect that both letters independently have a positive effect on turnout. Thus, the second hypothesis is that turnout for the group receiving the traditional package will be higher than the control group's turnout and the third hypothesis is that the humor package will have a higher turnout than the control group (the treatments are explained in detail in the research design section).

The fourth hypothesis is that the humor package has a larger effect than the traditional package. As mentioned above, the treatment effect from traditional letters are usually quite small, even in studies where the baseline turnout is lower than is the case here. One potential reason is that the receivers of traditional letters make little notice of the letters. If they are read, they might leave to little impression to be remembered on Election Day. One possible solution is to design more noticeable and surprising letters. Receiving the humor package, which includes a cartoon-like quiz,

from the Danish Parliament is quite unusual, and that might have the effect that the letter will be remembered when Election Day arrives – especially among young voters. This potential effect occur because the cartoon may initiate more talk about politics than the traditional package, which then should be expected to lead to higher participation rates (Klofstad 2007).

Finally, we hypothesize that the treatment effect is larger for voters with a lower propensity to vote compared to high propensity voters. First, there is a larger mobilization potential for this group. When the baseline turnout is so high as in the Danish case (remember that 72.5 percent voted in the absence of the campaign), the campaign will more often than not reach citizens who will vote regardless of the letter. Therefore, the potential to mobilize new voters to the polls is primarily to be found among the low propensity voters (e.g. people living in households where there is a non-voting habit). Secondly, research indicate that humorous political information has a larger impact on citizens with lower levels of political sophistication (Xenos & Becker 2009).

Context and data

We implemented the study prior to the Danish Municipal Elections held on November 19, 2013. The elections take place every four years across all the 98 different Danish municipalities simultaneously. The local government plays a central role in government service provision including public schools, childcare and eldercare. The local governments spend more than 29 percent of the GDP (2010) (Statistics Denmark 2012). The municipalities collect income taxes and decide the tax level under a large degree of autonomy. In 2014, the average income tax in the municipalities was 25 percent. The municipalities are quite diverse with regard to population size ranging from a 1,837 inhabitants in the small island municipality Læsø to 569,000 in the capitol municipality of Copenhagen (as of October 1st 2013).

The municipalities have proportional representation with several parties represented in each municipality. In general, the elections have a high saliency and receive much media attentions. The turnout average across all 98 municipalities in the 2013 elections was 71.9 percent. During recent decades, the turnout has fluctuated around 70 percent in municipal elections, which is somewhat lower than national elections that have fluctuated around 85 percent. In the election of 2013, the turnout increased with 6.1 percentage points compared to the preceding municipality election. This increase is quite exceptional and can have been fueled by among other things mobilization campaign in the main public radio and TV stations of unprecedented intensity. Turnout in the control group for this experiment was 72.5 percent. Considering that the control group is composed of voters aged 18 to 18.5 this is a high very baseline turnout. Compared to the election of 2009 the general turnout among the 18 years olds increased with 13.2 percentage points from 57 to 70.2 percent (Bhatti et al. 2014). This also suggests that any effect from our experiments is on top of a general increase in turnout.

All Danes have a social security ID. The ID number links to a range of background information on all Danes including their date of birth and residential address. This allowed to target only those that turn 18 shortly before the election and randomly assign them to either treatment or control condition. We can treat them precisely since we have accurate information on their residential address. After the election, we were able to merge validated turnout for all individuals with their personal ID number. Thus, we have information on individual level turnout and treatment status. In addition to this, we have a large number of variables from high quality government registers at our availability through *Statistics Denmark*, the official statistic bureau. All analysis were carried out on Statistic Denmark's server were all data are fully anonymous. 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 that are uncorrelated with treatment status. All our data are of high quality with a very low probability of measurement error. Any potential findings, or lack of such, are therefore unlikely to be due to stochastic or systematic measurement error.

Research design

When scholars have traditionally tried to evaluate the impact of various government initiatives they have encountered issues complicating their enterprise. The first issue is how to evaluate the programs. Suppose everyone in the population is exposed to a program. How is one to evaluate the effect of this? One approach could be to do a survey asking respondents if they remember the program and to assess the effect it has had on them. Furthermore, one could rely on the question of whether they remember being exposed to a treatment, and see how this correlates with the outcomes of interest. The problem with this approach is that people's ability to pay notice to and remember the treatment could be correlated with unobserved characteristics that correlate with the outcome of interest, too (Michelson & Nickerson 2011).

Another challenge is that exposure to many types of campaigns is not random. If an organization from a public agency or the civil society for example initiates a door-to-door information campaign not everyone will be home and open the door, when the campaign knocks on the door. If the score on the outcome in question is correlated with the propensity to open the door, an estimate that fails to consider this will be biased.

The result of either problem would be biased estimates of the effect with the risk of making wrong inferences to the population (Gerber & Green 2012: 38). For illustration, one can think of the limits of large-scale anti-smoking campaigns or campaigns to get citizens to sort their

trash. Here, it is an obvious challenge to reach all types of citizens with the campaign and actually get them to understand the message. It is easy to imagine that anti-smoking campaigns have minimal reach among extensive smokers as well as there could be a social bias concerning who actually gets the trash-sorting message. Furthermore, if evaluating by a survey, one should expect to get a biased sample with a lack of input from exactly those citizens, who are an important target group of the campaigns. The consequence would be a biased result with a likely over-estimation of the effect if one base the conclusions on a survey that do not reach those who we expect acts in the least desired way.

Field experiments are in many cases a feasible road to overcome the problems and make valid, causal inference of the effects of initiatives. In a field experiment, like in the more famed laboratory experiments, a subset of individuals is treated and a subset is assigned to a control group, which receive no treatment. The researchers have control over who is assigned to which group and the assignment follows a random procedure where subjects are assigned to either treatment or control with some known probability.

Until recently, this design was rarely used within the public administration literature (Cotterill & Richardson 2010; Margetts 2011; Perry 2012; Boruch 2005), even though field experiments can provide valid causal estimates with a good internal and external validity. However, the use of field experiment seems to be on the rise in both public administration and political science (Perry 2012; (Gerber 2011). Recently, field experiments in the field of public administration have been applied to study a wide array of problem. Researchers have evaluated public policy programs on garbage disposal (Dur & Vollaard 2014), recycling campaigns (Cotterill et al. 2009), an anti-drug project (Stoker & John 2009), and on public management programs' effect on employees work motivation (Jakobsen & Andersen 2013). There are more examples out there, and the use of field experiments seems to be on its way up.

Since the researchers assign treatment randomly, it is per definition expected to be uncorrelated with all observable and unobservable characteristics of the subjects. This implies that the estimated effect will not be subject to omitted variable bias. Furthermore, since assignment is done prior to the outcome of interest and independent of the expected outcome, there is no simultaneity. Thus with random assignment the only systematic difference between treatment and control group is the treatment, and any difference in the outcome of interest between the two groups can be ascribed to the treatment¹.

In this study, we use a variant the field experiment, namely the randomized wait-list field experiment. Oftentimes policy makers and researchers face situations where for ethical or political reasons it is impossible or at least difficult to exclude a group from a certain initiative or public program (Cotterill & Richardson 2010). In such a case a wait-list design can sometimes be applied. In this design all individuals receives the treatment but the timing of the treatment is varied between randomly selected groups. Measuring the difference on the outcome between the treatment and control group after the treatment group received the treatment and before the control group received it yields the short-term causal effect of the treatment.

Since 2005, the parliament has sent the Danish constitution to all citizens on their 18th birthday along with a letter welcoming them as adult citizens. One of the arguments for sending this material to the young citizens was that it could have a positive impact on turnout (Regeringen & Dansk Folkeparti 2004). However, until now it has remained untested whether it actually has such an effect. To make up for this knowledge deficit, while ensuring that all young Danes received the constitution, we designed a randomized wait-list field experiment with the parliament that allowed

¹ Two other assumptions are required (Gerber & Green 2012, 45). 1) The excludability assumption that says that the outcome of interest responds only to treatment when assigned to treatment and not to treatment assignment or indirect by-products of treatment assignment. 2) Non-interference (also known as SUTVA). The outcome of interest for the individual is independent of treatment status for other subjects, and the treatment status for the individual has no impact on the outcome of interest for other individuals.

for testing the effect of receiving the constitution, while only altering the parliament's practice of sending the constitution to all newly enfranchised citizens marginally.

Thus, we test the impact on turnout for 18 to 18½ year olds of simultaneously receiving the Danish constitution and a letter prior to the municipality election in Denmark. Usually, Danes receive the constitution when they become entitled to vote in the general elections. For native-born Danes this happens on their 18th birthday². In cooperation with the Danish Parliament, who is responsible for the distribution of the constitution, we postponed the mailing for all natives who turned 18 on May 9 until November 19 and did not share household with others who turned 18 in that period. This gave a sample of 33,520 individuals (cf. table 1). We removed a random sample of 3,000 individuals for use in an experiment presented elsewhere (see Bhatti et al. 2014a), leaving us with 30,520 individuals. From this, we sampled two groups of 10,000 individuals each. One of these groups were sent the constitution along with a traditional letter from the speaker of the parliament (traditional package). The other group were sent the constitution along with the humorous and informative cartoon (humor package). The last group contained 10,520 who functioned as the control group. As a part of the design, we originally wanted to test whether an SMS would have an independent impact. However, it was only possible to enrich less than 10 percent of the target group with cell phone numbers (cf. table 1). We chose to put aside this voters to analyses condoned elsewhere. The analysis here are restricted to the groups without phone numbers. There are no systematic bias with regard to the drop-out in each treatment group and control group. Thus, the causal inference in the analysis remain valid (For more info, see Bhatti et al. 2014).

² Persons with permanent residency in Denmark who obtain citizenship receive a constitution, too.

Table 1: Individuals at different stages of the randomization process

Originally population of young voter aged 18 to 18½ on election day		33,520
Random drawn to alternative experiment.	3,000	
Sub population		30,520
Obtain a cell phone match among entire sub population (exclude to use in SMS experiment)	2,657	
Sub population		27,863
No voting record found	313	
Final population		27,550
Random control group	9,475	
Random group receiving cartoon and constitution (humor package)	9,046	
Random group receiving letter and constitution (traditional package)	9,029	

On Thursday, November 14, the respective packages were sent out to both of the treatment groups, in order for them to receive it the next day. The control group received the constitution after the election. This means that they were unaffected by the treatment prior to the election. An attractive feature of this design is as indicated above that we excluded no one from receiving the constitution. Oftentimes policy makers and researchers face situations where for ethical or political reasons it is impossible or at least difficult to exclude a group from a certain initiative or public program. Our design demonstrates how one is sometimes able to overcome such a hurdle. This is obviously no magic wand useful in every situation, but it is a relevant opportunity to keep in mind.

After the election, we matched turnout obtained from the official voting files to the personal ID number. Of the subjects in the sample prior to the election a small proportion of 313 did not figure on the voting files. The reasons for this were that a few municipalities made administrative errors with some of their voting files and that some subjects were removed from the voting files either because they left the country or died. Since we removed a random subset from our sample and experienced minor attrition uncorrelated with the treatment assignment we end up with groups that are slightly smaller than described above.

For our experiment to yield valid inference, our new sample must be a random subset of the original sample and balanced on covariates. To ensure this, we made a logistic regression of being in the new sample on the original assignment as well as a multinomial logistic regression of treatment status in the new sample on all covariates prior to the treatment assignment. In both cases, the model was unable to make significant prediction of the outcomes. This assures us that our sample is still suitable for valid inference of the treatment effect. This is not surprising as the two sources that lead subjects to be removed from the experiment, phone-number enrichment and not figuring on the voter file, were both uncorrelated with the treatment.

Traditionally, the parliament has sent the constitution to everybody in the target group together with a short, formal letter signed by The Speaker. An important aspect of this field experiment was to evaluate the effect on turnout of the letter that the parliament normally uses. Therefore, the traditional letter was used with some slight adjustments to focus it on the municipal election. This letter together with an easy-readable version of The Constitution made up the traditional package (see appendix A for the letter).

However, the parliament was not only interested in evaluating their traditional letter. They were also interested in testing whether another type of communication would have the same, a smaller or larger effect on turnout. Therefore, they designed a humorous type of letter. As described above, such communications can be more eye-catching and start more discussion in general, which again can have an effect on turnout. The final product was a sort of quiz that the receiver could do by herself, where she should answer questions and follow the steps in accordance with her answer. The comic can be found in appendix B.

To sum up, we have valid and reliable measure of treatment and turnout for control and treatment group with treatment randomly assigned. This allows us to estimate the causal effect of receiving the constitution with either a standard or a humorous letter. We now turn to this matter.

Analysis

Our first hypothesis stated that a postal encouragement increase turnout compared to no postal encouragement. In table 2, we show the actual turnout in the treatment groups along with a test of whether or not the increase in turnout caused by the treatments were statistically significant. As our hypotheses are one-sided - increase over no effect - we apply one-sided tests of significance, which also is the dominant approach in related experimental work (Gerber & Green 2012:64).

Table 2: Turnout across experimental groups

	Turnout percent	Increase in turnout compared to control	N
Control group	72.5		9,475
Combined	73.6	1.1* (0.56)	18,075
Traditional package	73.2	0.6 (0.65)	9,029
Humor package	74.1	1.6* (0.65)	9,046

The differences to the control group are tested by applying a logit regression with turnout as dependent variable and treatment as independent with control group as reference. * $p < 0.05$ (one-sided test). The difference between the two treatments is not statistically significant $p = 0.072$ (one-sided test). Standard errors from average marginal effects are in parentheses.

The turnout in our combined group, composed of individuals who either received the traditional letter or the cartoon, is 73.6 percent. This is a statically, significant increase of 1.1 percentage point compared to the control group who received delayed treatment. Thus, we find support for our first

hypothesis; a postal encouragement together with the Danish constitution and sent by the parliament increase turnout significantly compared to no encouragement.

Looking into the next two hypotheses, we find that the traditional package does not have a significant effect on turnout though our point estimate is an increase of 0.6 percentage point (hypothesis 2). On the other hand, the humor package increased turnout significant with 1.6 percentage point compared to the control group (hypothesis 3). The 1.0 percentage point difference in turnout between the traditional and the humor package is non-significant (hypothesis 4) though the point estimate of the difference is noticeable³.

Calculated as numbers of persons our best estimate is that the humor package made 145 extra young voters go to the polls. If we set the price of the material and postage to €2, this mean that, each extra vote costs €125. This is of course quite expensive. However, we have to remember that this is the cost based on the effect measured only on the single citizen receiving the package and in a single election. Studies have documented substantial spill-over-effects from GOTV-campaigns to other co-resident voters (Nickerson 2008) as well a habit forming element as the treatment effect carries on the future elections (Davenport et al. 2010). The consequence is that the price pr. vote calculated here probably is over-estimated.

Heterogeneous treatment effects

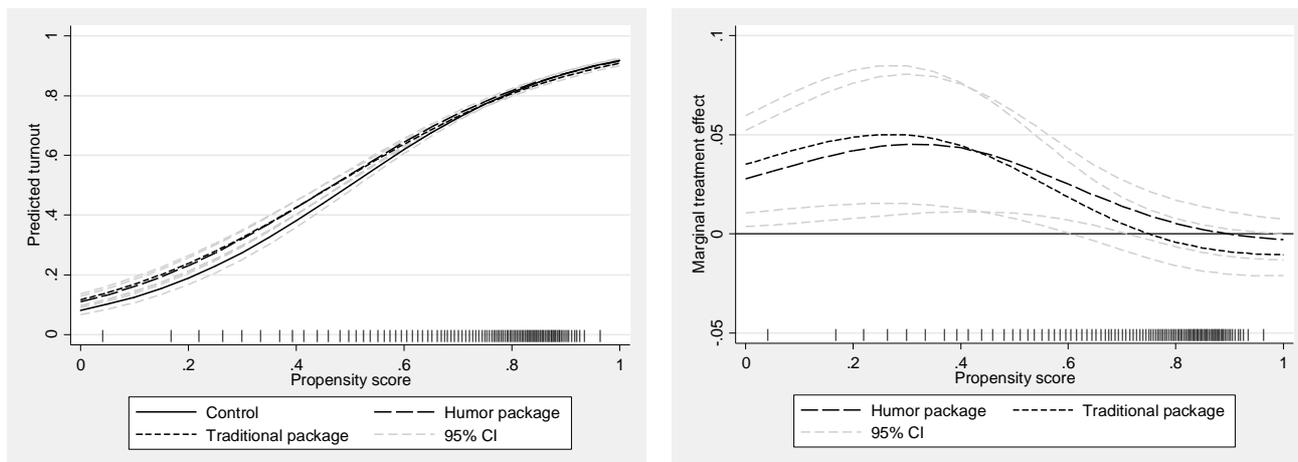
Our fifth hypothesis stipulated that we would expect that persons in the treatment groups that initially had a low propensity to vote would be more affected by the treatment. These persons have

³ Of the 27,550 persons in our experimental groups, 14,608 were assigned to vote at polling stations applying digital voting lists. For this subset we know if the vote was cast before the voter could have received the constitution. Of the 14,608, 778 voters casted their vote on November 14th or earlier (i.e. before they could have received the constitution). Excluding these 778 from the analysis does not alter the results.

the largest opportunity to improve whereas other might already vote regardless of being treated and as such, the high propensity to vote group might experience a ceiling effect. In order to investigate this we regress whether or not the person voted in our control group on a range of variables of variables that are commonly known to predict turnout. We know from extensive studies of turnout in Danish municipality elections that first and second generation immigrants, young voters that have moved away from their parents, young voters where their parents abstain from voting in the preceding election, men and voter living far away from the polling station vote substantial less than others (Bhatti & Hansen 2010; Bhatti et al. 2014b; Bhatti & Hansen 2013; Bhatti 2012).⁴ We use the obtained regressions coefficients to predict the turnout for all individuals in our experimental groups. The resulting prediction can be termed “the propensity to vote” (see Enos et al. 2014 for a similar approach). The next step is to take the actual turnout of the treatment groups interacted with the propensity to vote in order to see whether the effect of the treatments vary by turnout propensity. The result is visible in figure 1 showing the marginal, predicted probabilities from the logistic regression of turnout on propensity to vote interacted with treatments in the left panel. The right panel shows the marginal treatment effect. Each plot includes a rug plot where we plot each percentile from the distributions of propensity to vote including the maximum and minimum value.

⁴ We also include municipality fixed effect to control for differences between the municipalities. We apply a logit regression with robust standard errors.

Figure 1: Turnout across experimental groups and propensity to vote



The left panel of figure 1 shows that there is no effect of the treatments among voters with high propensity to vote as the high turnout comparing control and treatment group is no different. Among voters with low propensity to vote, we see difference among the control group and the treatment groups. In the panel to the right, the differences are clearer. The right panel show that the marginal treatment effect (MTE) peaks among voters with about a 0.3 propensity to vote. At this level, the MTE is about 5-percentage point. We can also see that the MTE is significant for voters with a propensity to vote up until 0.6. When the propensity to vote exceeds 0.6 any differences become non-significant. This supports the fifth hypothesis that stated that we should expect to find stronger effects among the least likely to vote, whereas our treatment has no effect among the most likely to vote in the first place. This finding has the important implication that the gap between the turnouts across subgroups become smaller due to the treatments as those least likely to vote catch somewhat up to those most likely to vote. This contradicts Enos et al. (2014) who find an increased gap due to mobilization efforts. One explanation for the difference between our study and Enos et al. (2014) could be the context and especially the saliency of the election (see also Arceneaux & Nickerson 2009). Compared to any US election, the Danish municipal elections have a very high turnout. As has previously been noted (Arceneaux & Nickerson 2009; Enos et al. 2014) it seems to be easier to mobilize low propensity voters in high saliency elections.

A last point we draw from figure 1 is that the effects of the two treatments are similar across most propensities. There is a tendency that the humor package is stronger than the traditional package among the persons with high propensity to vote less strong among the low propensity voters. As there is a majority of persons in our experiment with a high propensity to vote persons, these individuals tend to drive the results found in table 1 where the humor package has a higher effect than the traditional package compared to the control group. This means that if the focus is exclusively on mobilizing low propensity voters we find no difference between the two treatments.

From a practical policy perspective, it is perhaps hard to get a grasp of who a low propensity voter could be. In additional analyses, we calculated effect sizes for various low-propensity groups. The highest effects we found for second-generation immigrant. This group has a turnout of 51.5 percent in the treatment condition with the humor package compared to a 44.6 percent in the control group, i.e. a treatment effect of 6.9 percentage points. If we apply the cost estimate of €2 pr. humor package sent to this subgroup an extra vote among second generation immigrants costs €29 pr. extra vote. This is still a considerable expenditure, but substantial less than the cost of €125 calculated across all treatment groups. While second generation immigrants is of course a specially selected group, this does remind us that we can design interventions focusing on selected target groups and that average treatment effects might cover considerable heterogeneous effects when we run experiments.

Conclusion

In the public administration literature, we are often faced with the challenge of evaluating existing government programs without the opportunity to exclude individuals from receiving a potential beneficial treatment. In this study, we evaluated the Danish parliament's practice of sending the constitution and a letter to young adults with only a small adjustment to the on-going government information campaign. Using a randomized wait-list field experiment where a control group got a delayed treatment after the point of measurement, we were able to test the program's effect on voter turnout with strong internal and external validity and in a way that was feasible considering the administration's continuing practice of sending material to all young adult. The study furthermore contributes to our existing knowledge of how to increase voter turnout.

The study shows that voter turnout among voters between 18 to 18½ years old at the municipality election of 2013 was mobilized by 1.1 percentage point through a postal encouragement including the constitution and a traditional letter or a humorous letter. Two type of postal encouragement were tested. The traditional letter increased the turnout by 0.6 percentage points and the humorous letter with 1.6 percentage point. These treatment effects are somewhat larger than traditional found in US GOTV-studies. Furthermore, we found that among the young voters with low propensity to vote the effect can be up to five percentage points whereas young voters with a high propensity to vote seem not to be mobilized by the treatment. This suggests that our mobilization effort has decreased the gap in turnout between the low and high propensity to vote persons.

We have two concerns regarding the generalizability of our findings. First, we should note that our turnout in the control group is high and much higher than in other GOTV-studies and that we experienced a strong, general increase of 13.2 percentage point among the 18 years old in

the election under investigation compared to the previous election. It is quite likely that these two factors have worked against stronger effects of our treatments. In other words, we would expect our findings to be stronger if there for instance had been fewer media campaigns to increase turnout. Furthermore, as this election is a high salience election compared to any U.S. election, where most previous studies have been conducted, our finding, in a setting with an overall turnout rate of 71.9 percent, should be more likely to apply to average European elections than to the low salience elections of the U.S.

Second, we have estimated that the cost of an extra vote with the humor package was €125 for the cost of an extra vote in the general sample, but substantially less for low propensity voters. Furthermore, these cost measures are quite conservative in the sense that it does not include any spill-over effects to other individuals or habit effects spill-over effects to subsequent elections (habit effects) – two spill-over effects which have been established in the literature.

As of present, the Danish parliament has decided to continue sending out the constitution on the birthday of all 18 years old with the traditional letter. It has yet to be decided if they want to adjust how the constitution is sent out due to the evidence from this field experiment.

Appendix A: Parliament traditional letter send out together with the constitution

FOLKETINGET



Folketingets formand

Christiansborg
DK-1240 København K

Tlf. +45 33 37 55 00
Fax +45 33 32 85 36

www.ft.dk
ft@ft.dk

Tillykke

Du er blevet 18 år og har fået valget. Det giver dig ret til at stemme til kommunalvalg, og hvis du har dansk statsborgerskab også til bl.a. folketingsvalg og folkeafstemninger. På den måde kan du få indflydelse på, hvem der bestemmer i Danmark, og hvad det er for en politik, der skal føres.

For at værne om det danske folkestyre er det vigtigt, at vi som borgere gør brug af muligheden for at stemme og af muligheden for selv at blive aktive i politik. Det var i 1849 – for mindre end 200 år siden – at Danmark afskaffede enevælden, hvor kongen bestemte, og fik den første grundlov. Grundloven er fundamentet for vores demokrati – den er grundlaget for det danske folkestyre.

Grundloven fastsætter først og fremmest de grundlæggende rammer for, hvordan landet skal styres. Det gælder f.eks. princippet om magtens tredeling i den lovgivende, den udøvende og den dømmende magt. Grundloven indeholder også det værdisæt, som er bærende for det danske folkestyre, nemlig frihedsrettighederne.

Grundloven tilhører os alle sammen, og den er vigtig for os. Derfor har Folketinget besluttet, at alle, der fylder 18 år eller bliver dansk statsborger, skal modtage grundloven med forklaringer. Jeg håber, at du vil læse grundloven og bruge den som fundament for din deltagelse i den demokratiske debat.

Med venlig hilsen

A handwritten signature in black ink, appearing to read 'Mogens Lykketoft'.

Mogens Lykketoft
Folketingets formand

Appendix B: Parliament humorous letter send out together with the constitution

TILLYKKE! DER ER VALG OG DU MÅ STEMME!

... HVORFOR? DERFOR:

- ER DU ET MENNESKE? (Are you a human?)
 - JA: ER DU TILFREDS MED DIT LIV? (Are you satisfied with your life?)
 - JA: SYNES DU VALG ER KEDELIGT? (Do you think elections are boring?)
 - JA: ER DU LIDT DOVEN? (Are you a bit slow?)
 - JA: KAN DU LIDE AT FESTE? (Do you like to party?)
 - JA: HAR DU VENNER? (Do you have friends?)
 - JA: KAN DU LI' SPÆNDENDE BØGER? (Do you like exciting books?)
 - JA: KAN DU LI' KAGE? (Do you like cake?)
 - STEM
- NEJ: ... I TVIVL? TAG TESTEN: (In doubt? Take the test!)
 - ER DU ET MENNESKE? (Are you a human?)
 - NEJ: IKKE SÅ TOSSET! (Not so stupid!)
 - GAVE TIL DIG: (Gift to you)

- STEM
- STEM PÅ DEM (Vote for them)
- STEM PÅ MIG (Vote for me)
- STEM PÅ DIG SELV (Vote for yourself)
- STEM PÅ EN DOKTOR (Vote for a doctor)
- STEM PÅ DEN HVIDE TIGER (Vote for the white tiger)
- STEM PÅ DEN RØDE TIGER (Vote for the red tiger)
- STEM PÅ DEN GRØNNE TIGER (Vote for the green tiger)
- STEM PÅ DEN BLÅ TIGER (Vote for the blue tiger)
- STEM PÅ DEN GULDE TIGER (Vote for the golden tiger)
- STEM PÅ DEN SØLV TIGER (Vote for the silver tiger)
- STEM PÅ DEN BRUNE TIGER (Vote for the brown tiger)
- STEM PÅ DEN HVIDE TIGER (Vote for the white tiger)
- STEM PÅ DEN RØDE TIGER (Vote for the red tiger)
- STEM PÅ DEN GRØNNE TIGER (Vote for the green tiger)
- STEM PÅ DEN BLÅ TIGER (Vote for the blue tiger)
- STEM PÅ DEN GULDE TIGER (Vote for the golden tiger)
- STEM PÅ DEN SØLV TIGER (Vote for the silver tiger)
- STEM PÅ DEN BRUNE TIGER (Vote for the brown tiger)

FACT CHECK: DIN STEMME UDGØR:

- MANGE SÅDAN ER DET! MEN TÆNK HVIS INGEN INGEN STEMTE, SÅ VILLE DIN VERDEN HVORLÆNG KUNNE TIL... AT SE SÅDAN UD.
- ALDRIG INTERNET
- ALDRIG KLADDE MUSIK
- ALDRIG KÆNDE

HUSK DET HER:

Kærlig hilsen Folketinget

...NÅR DU SKAL STEMME - HVORDAN? SÅDAN:

... ER DU NU KLAR TIL VALG! ... KOM, SOM DU ER!

VALG

QR code

References

- Arceneaux, K. & Nickerson, D. W. 2009. Who Is Mobilized to Vote ? A Re-Analysis of 11 Field Experiments. *American Journal of Political Science*, 53(1): 1-16.
- Bhatti, Y. 2012. Distance and Voting: Evidence from Danish Municipalities. *Scandinavian Political Studies*, 35(2): 141-158.
- Bhatti, Y., Dahlgaard, J. O., Hansen, J. H. & Hansen, K. M. 2014a. Kan man øge valgdeltagelsen? Analyse af mobiliseringstiltag ved kommunalvalget den 19. november 2013 *CVAP Working Paper Series, CVAP WP 3/2014*, 2014(3): 1-91.
- Bhatti, Y., Dahlgaard, J. O., Hansen, J. H. & Hansen, K. M. 2014b. Hvem stemte og hvem blev hjemme? Valgdeltagelsen ved kommunalvalget 19. november 2013. Beskrivende analyser af valgdeltagelsen baseret på registerdata. *CVAP Working Paper Series, CVAP WP 2/2014* 2014(2): 1-406.
- Bhatti, Y. & Hansen, K. M. 2010. Valgdeltagelsen ved kommunalvalget 17. november 2009. Beskrivende analyser af valgdeltagelsen baseret på registerdata. *Arbejdsrapport Københavns Universitet, Institut for Statskundskab*, 2010(3).
- Bhatti, Y. & Hansen, K. M. 2013. The effect of co-residence on turnout. *MPSA Annual Conference 2013*. Chicago.
- Boruch, R. 2005. Preface: Better evaluation for evidence-based policy: Place randomized trials in education, criminology, welfare, and health. *The Annals of the American Academy of Political and Social Science*: 6-18.
- Brewer, P. R., Young, D. G. & Morreale, M. 2013. The Impact of Real News about "Fake News": Intertextual Processes and Political Satire. *International Journal of Public Opinion Research*, 25(3): 323-343.
- Brown, C. H., Wyman, P. A., Guo, J. & Peña, J. 2006. Dynamic wait-listed designs for randomized trials: New designs for prevention of youth suicide. *Clinical Trials*, 3(3): 259-271.
- Cao, X. 2010. Hearing it from Jon Stewart: The impact of The Daily Show on public attentiveness to politics. *International Journal of Public Opinion Research*, 22(1): 22-46.
- Costa, D. L. & Kahn, M. E. 2013. Energy Conservation "Nudges" And Environmentalist Ideology: Evidence From A Randomized Residential Electricity Field Experiment. *Journal of the European Economic Association*, 11(3): 680-702.
- Cotterill, S., John, P., Liu, H. & Nomura, H. 2009. Mobilizing citizen effort to enhance environmental outcomes: A randomized controlled trial of a door-to-door recycling campaign. *Journal of environmental management*, 91(2): 403-410.
- Cotterill, S. & Richardson, L. 2010. Expanding the Use of Experiments on Civic Behavior: Experiments with Local Government as a Research Partner. *The ANNALS of the American Academy of Political and Social Science*, 628(1): 148-164.
- Davenport, T. C., Gerber, A. S., Green, D. P., Larimer, C. W., Mann, C. B. & Panagopoulos, C. 2010. The enduring effects of social pressure: Tracking campaign experiments over a series of elections. *Political Behavior*, 32(3): 423-430.
- Dur, R. & Vollaard, B. 2014. The Power of a Bad Example-A Field Experiment in Household Garbage Disposal. *Tinbergen Institute Discussion Paper*, 61(1): 1-43.
- Enos, R. D., Fowler, A. & Vavreck, L. 2014. Increasing Inequality: The Effect of GOTV Mobilization on the Composition of the Electorate. *The Journal of Politics*, 76(1): 273-288.
- Fischer, R., Akin, A., White, B., Arant, D., Chamberlain, T. & Bolton, S. 2013. The smallest user campaign to decrease energy consumption in two Mid-South neighborhoods: A field experiment. *Public Relations Review*, 39(4): 391-393.

- Gerber, A. 2011. Field experiments in political science. pp. 115-140 In: James N. Druckman, D. P. G., James H. Kuklinske, Arthur Lupia (ed.) *Handbook of Experimental Political Science*. Cambridge: Cambridge University Press.
- Gerber, A. S. & Green, D. P. 2000. The Effects of Canvassing, Telephone Calls, and Direct Mail on Voter Turnout: A Field Experiment. *The American Political Science Review*, 94(3): 653-663.
- Gerber, A. S. & Green, D. P. 2001. Do phone calls increase voter turnout?: A field experiment. *Public Opinion Quarterly*, 65(1): 75-85.
- Gerber, A. S. & Green, D. P. 2012. *Field Experiments: Design, Analysis, and Interpretation*, New York: W. W. Norton Limited.
- Gerber, A. S., Green, D. P. & Larimer, C. W. 2008. Social Pressure and Voter Turnout: Evidence from a Large-Scale Field Experiment. *American Political Science Review*, 102(1): 33-48.
- Green, D. P., Mcgrath, M. C. & Aronow, P. M. 2013. Field experiments and the study of voter turnout. *Journal of Elections, Public Opinion & Parties*, 23(1): 27-48.
- Haynes, L., Owain, S., Goldacre, B. & Torgerson, D. 2012. Test, learn, adapt: developing public policy with randomised controlled trials. *Cabinet Office-Behavioural Insights Team*.
- Jakobsen, M. & Andersen, S. C. 2013. Intensifying social exchange relationships in public organizations: Evidence from a randomized field experiment. *Journal of Policy Analysis and Management*, 32(1): 60-82.
- Klofstad, C. A. 2007. Talk Leads to Recruitment: How Discussions about Politics and Current Events Increase Civic Participation. *Political Research Quarterly*, 60(2): 180-191.
- Margetts, H. Z. 2011. Experiments for public management research. *Public Management Review*, 13(2): 189-208.
- Michelson, M. R. & Nickerson, D. W. 2011. Voter mobilization. pp. 228-242 in: James N. Druckman, D. P. G., James H. Kuklinske, Arthur Lupia (ed.) *Cambridge Handbook of Experimental Political Science*. Cambridge: Cambridge University Press.
- Nickerson, D. W. 2007. Does email boost turnout. *Quarterly Journal of Political Science*, 2(4): 369-379.
- Nickerson, D. W. 2008. Is Voting Contagious? Evidence from Two Field Experiments. *American Political Science Review*, 102(1): 49-57.
- Perry, J. L. 2012. How Can We Improve Our Science to Generate More Usable Knowledge for Public Professionals? *Public Administration Review*, 72(4): 479-482.
- Phillips, R. O., Ulleberg, P. & Vaa, T. 2011. Meta-analysis of the effect of road safety campaigns on accidents. *Accident Analysis and Prevention*, 43(3): 1204-1218.
- Regeringen & Dansk Folkeparti 2004. Aftaler om Finansloven for 2005 in: Finansministeriet (ed.). Copenhagen: Schultz Information.
- Richard, M. A., Martin, S., Gouvernet, J., Folchetti, G., Bonerandi, J. J. & Grob, J. J. 1999. Humour and alarmism in melanoma prevention: a randomized controlled study of three types of information leaflet. *British Journal of Dermatology*, 140(5): 909-914.
- Stoker, G. & John, P. 2009. Design experiments: Engaging policy makers in the search for evidence about what works. *Political Studies*, 57(2): 356-373.
- Stutzer, A., Goette, L. & Zehnder, M. 2011. Active Decisions and Prosocial Behaviour: a Field Experiment on Blood Donation. *Economic Journal*, 121(556): 476-493.
- Xenos, M. A. & Becker, A. B. 2009. Moments of Zen: Effects of The Daily Show on Information Seeking and Political Learning. *Political Communication*, 26(3): 317-332.
- Young, D. G. & Hoffman, L. 2012. Acquisition of Current-Events Knowledge From Political Satire Programming: An Experimental Approach. *Atlantic Journal of Communication*, 20(5): 290-304.