

Door-to-door canvassing in the European elections: Evidence from a Swedish field experiment*

Pär Nyman[†]
Department of Government
Uppsala University

December 14, 2016

Abstract

In this paper I report the results from a door-to-door canvassing experiment conducted in Sweden during the 2014 European elections. The canvassing was performed by members of the Social Democratic Party and the experiment closely resembles the partisan nature of most mobilizing campaigns in Europe. The paper is one of the first to provide causal evidence for the mobilizing effectiveness of canvassing outside the United States. Living in a household that was visited by canvassers increases the probability of voting by 3.6 percentage points. This effect is entirely driven by estimates twice as large for occasional and first-time voters. Compared to previous research, the high compliance rate gives precisely estimated effects that are closer to average treatment effects.

Introduction

The rich world has long been plagued by a steady decline in voter turnout. This trend concerns the very core of political science, because lower turnout worsens the existing inequalities in political participation and – ultimately – threatens the legitimacy of our democratic institutions (Lijphart 1997).

In recent research as well as contemporary political campaigns, different kinds of “get out the vote” efforts have been suggested as the cure to this problem. Following the seminal works of Gerber and Green (2000), a fast-growing literature has emerged that uses randomized experiments to test the mobilizing effectiveness of different campaign tactics. Door-to-door

*I am grateful to the two anonymous reviewers, Rafael Ahlskog, Elin Bjarnegård and John Ternovski for their helpful comments and to Peter Gerlach, Carl Kangas, Ingela Nilsson and Fredrik Olovsson at the Social Democratic party for their co-operation. I would also like to direct a special thanks to Nicklas Vinthagen for excellent research assistance and to all the canvassers for their hard work.

[†]Contact details and supplementary information can be found at www.parnyman.com.

canvassing is generally proposed as the most cost-effective of these methods, trumping both phone calls and direct mail.

In addition to optimizing mobilization campaigns, such field experiments are also of scholarly interest because they deepen our knowledge about the reasons why people vote. Previous studies have provided insights into why couples living together share similar voting behaviors (Nickerson 2008), why immigrants often have a lower degree of political participation than the native-born (Pons and Liegey 2013) and whether our decision to vote is most affected by intrinsic satisfaction from norm-compliance or from extrinsic incentives and concern about what other people think of us (Gerber et al. 2008).

While field experiments are blessed with strong internal validity, they are also characterized by limited external validity. It is therefore unfortunate that with few exceptions, almost every experiment on canvassing effects has been conducted in the United States. While door-to-door canvassing is becoming increasingly popular in the rest of the world, many have questioned its effectiveness outside the US.

In this paper I report the results from a field experiment carried out in the Swedish county of Södermanland, where 11 640 citizens were randomly assigned to either the treatment or the control group. The experiment was conducted in cooperation with the Social Democratic Party, as a part of their campaign during the 2014 European elections. Approximately 60 per cent of the households in the treatment group were visited by party members, who asked them to vote in the election and suggested that they should vote for the Social Democrats. This study differs from previous experiments in two important ways.

The first distinguishing feature of this study is the high rate of compliance. A contact rate of 60 per cent is twice as high as in most comparable studies and the number of treated households in the control group is negligible. There are two reasons why compliance matters. First, it is only possible to study how canvassing affects those voters that are actually contacted by the canvassers, but there are many occasions when we might be more interested in how the average voter would respond. The higher our contact rate, the less the effect on the treated differs from what the average treatment effect would have been. Second, the high rate of compliance in this study, in combination with a relatively large sample size, yields a precision in the estimated effects that few other field experiments can match.

The second distinguishing feature is that the European elections offer a setting very different from those of previous studies. With over 400 million eligible voters they are among the largest elections in the world. As such, it is unfortunate that we know so little about the effectiveness of canvassing in Europe (Nielsen 2014). There are several reasons why we should expect the results of door-to-door canvassing to differ from the American experience. First, canvassing in Europe is typically performed by members

of the political parties, but most American experiments have been conducted by non-partisan canvassers. Evidence from mailing campaigns suggest that non-partisan messages are more effective at mobilizing voters than mails that encourage the recipient to vote for a given issue or candidate (Green et al. 2013).

Second, turnout in most of Europe – and especially in Sweden – is much higher than in the United States. However, the elections for the European Parliament can be described as second-order elections, where turnout is lower and less is at stake than in the national elections. How do the relatively strong national voting tendencies, in combination with a low-salience election, affect the mobilizing effect of canvassing? If the voting habits mean that the norms about voting are more pronounced, I would expect the social pressure induced by mobilizing campaigns to be stronger, and thereby the effects of door-to-door canvassing to be larger. Similarly, the lower turnout in the elections to the European parliament means that there are more voters who can be mobilized.

Third, the European political context differs from the American in many other ways. Most electoral systems in Europe are based on proportional representation, but every successful canvassing experiment has been conducted in countries with majoritarian systems. Moreover, most European countries do not require the citizens to register before voting. As a result, it might be more difficult for parties to target likely voters and especially voters affiliated with their own party. On the other hand, there are no unregistered voters that by definition are out of reach for the canvassers. All these arguments are developed in more detail in the next section.

Compared to previous research in similar settings, the results of this experiment strengthen the case for door-to-door canvassing in Europe. The electoral rolls show that being visited by a canvasser increases the probability of voting with 3.6 percentage points. However, the effect differs substantially between subgroups. No effect was found among voters with a very high or very low propensity to vote. For citizens who voted in the last national election, but refrained from voting in the previous European election, the estimated effect was 6.0 percentage points. I also find larger effects for young people, those living in single-family dwellings and for people in large households, compared to old people, people living in multiple-family dwellings and single-person households.

Because the study is conducted in Sweden, it also fills an important gap in the comparative literature on electoral systems. It is widely acknowledged that electoral systems based on proportional representation increase voter turnout. One of the most commonly proposed mechanisms for this regularity is that the voter mobilization achieved by parties and interest organisations is stronger in countries with PR systems (Cox 2015). But to the best of my knowledge, this study is the first well-conducted canvassing experiment that takes place in a country with proportional representation.

The remainder of the paper proceeds as follows. In the following section I discuss the setting of this experiment in relation to previous research. I then proceed by describing the data analyzed in this paper before I eventually present the results. The last section concludes.

The context and previous research

This study takes place in Sweden during the 2014 elections to the European Parliament. The European elections are among the largest in the world, with over 400 million eligible voters. However, the turnout has been falling for almost every election, from 62 per cent in 1979 to 43 per cent in 2014,¹ and the trend is likely to continue (Bhatti and Hansen 2012). Despite its high turnout in national elections, also Sweden has difficulties mobilizing its voters in European elections, and when 51 per cent of its citizens voted in 2014 it was the highest turnout since the country joined the European union in 1995.

The experiment was conducted in 17 electoral districts, located in five municipalities in the county of Södermanland. The districts are characterized by a high share of Social Democratic voters and a low turnout. In the 2014 election for the European Parliament, the support for the Social Democrats amounted to 40.3 per cent of the votes in these districts, compared to 24.2 per cent in the rest of Sweden. The turnout of 41.4 per cent of the eligible voters was almost ten percentage points below the national average of 51.1 per cent.

The experiment was designed with the ambition to resemble ordinary partisan canvassing as closely as possible. The 155 canvassers were members of the Social Democratic Party and they used the same script as the party used in the rest of their national campaign. The age of the canvassers ranged from 16 to 80 with an average of 52, which is very different from the college students often employed in US experiments. Approximately 20 per cent of the canvassers reported that they hold a political position which makes them known to large parts of the electorate.

The canvassing took place during the three weeks that preceded the election and at least a few households were visited every day during this period. On average, there were 11 days between the visit and the election. Most of the canvassing (70 per cent of the successful contact attempts) was done in pairs. The canvassers knew that they participated in a scientific study, but the citizens they visited were not informed about the experiment.

Below I will go through the arguments for why we should expect different effects of this experiment compared to what previous research has found. First I argue why the mobilizing effect can differ between electoral systems and what consequences it may have. Then I discuss the aspect of partisan

¹A large part of this decline is caused by the enlargement of the European Union.

compared to non-partisan canvassing. Third, I provide three reasons for how strong voting habits, combined with the second-order nature of the European elections, can influence the effectiveness of mobilizing campaigns. Lastly I go through the empirical results in previous research and outline my expectations of this experiment.

Canvassing and proportional representation

That voter turnout is higher in countries with proportional representation than in countries with single-member districts is one of the most robust findings in the comparative literature on electoral systems (Lijphart 1999), but little is known about the mechanisms behind this empirical regularity (Blais and Aarts 2006). A commonly proposed explanation is that the incentives for mobilization efforts are stronger in PR systems (Rainey 2015).

This discussion has been focused on how a proportional distribution of seats affects the incentives for mobilizing additional votes (Cox 1999, 2015; Herrera et al. 2014; Karp and Banducci 2008; Karp et al. 2003, 2008). However, the incentives for mobilization efforts are also dependent on how successful mobilization campaigns are in increasing voter turnout. Unfortunately, the literature is surprisingly scarce in this aspect.

There are theoretical arguments for both higher and lower effectiveness of mobilization campaigns in countries with proportional representation. On the one hand, Cox (1999) expects strong ties between parties and social groups to increase the effectiveness of mobilizing efforts. Because partisan links tend to be stronger in PR systems, Cox argues that mobilization efforts should translate into more votes in countries with proportional representation. Karp et al. (2003, 2008) hypothesize that mobilization will be more effective under proportional representation, because the absence of safe seats and "wasted votes" makes it easier for parties to persuade potential voters.

On the other hand, there are also arguments why mobilization should be less effective in PR systems. First, politics in majoritarian systems tend to be more focused on the candidates, with more personal campaigns, more contact between constituents and representatives and weaker party affiliation. It is probably easier to know what to expect from a party than from a specific candidate. Information about a candidate, and trust toward the same, can be obtained through contact with the candidate's campaign. Therefore, mobilization campaigns should be more effective when the need for trust and information is larger. In line with this argument, Górecki and Marsh (2012) find that the effect of canvassing is larger when the geographic distance between the home of the voter and the candidate is large. Second, Cox (1999, 2015) and Powell (1989) argue that mobilization in PR systems has relied on subcontracting to affiliated interest groups. The reason is that the larger number of parties typically found in PR systems give rise to more ideological parties with closer ties to labour unions, churches

and other interest organisations. To explain the lower degree of partisan mobilization campaigns in PR systems, this argument implicitly assumes that the mobilization efforts conducted by affiliated groups have reduced the effectiveness of further partisan campaigns. Third, door-to-door canvassing and other forms of contacting are more common in countries with majoritarian electoral systems than in countries with proportional representation. Even in safe districts, a larger share of citizens report that they have been contacted than in countries with PR systems (Karp et al. 2008). This pattern is recurring in the elections to the European Parliament, where canvassing is most prominent in Ireland and the UK (Gagatek 2010). If these differences reflect rational choices by the campaign generals, we should expect canvassing to be less effective under proportional representation.

Unfortunately, we know very little about the effectiveness of door-to-door canvassing in countries with proportional representation. In an observational study on survey data, Karp et al. (2008) find no clear patterns in the effectiveness of canvassing, but if anything, the results suggest that door-to-door canvassing might more effective in PR systems.² However, such studies are known for biased results because campaigns are more likely to target voters than non-voters and because contact with the campaign increases the propensity of non-voters to falsely report that they voted. Reliable estimation of canvassing effects therefore requires both random assignment and individual-level voter turnout records. To the best of my knowledge, the only canvassing experiments conducted in a PR system are the ones reported in Bhatti et al. (2014b). While they do not find any statistically significant effects, the small sample sizes and the unknown compliance rates makes it difficult to interpret their results. Consequently, this is the first study which convincingly identifies the effect of door-to-door canvassing in a country with proportional representation.

Partisan canvassing

Until very recently, the only solid evidence for the mobilizing effects of door-to-door canvassing came from non-partisan experiments where college students and community organisations asked voters to do their civic duty. While there have been a few partisan experiments conducted during the last few years, non-partisan canvassing still amounts for most of the published studies.

There are many reasons why we would expect the effectiveness to differ between partisan and non-partisan campaigns. First, non-partisan canvassers typically deliver a message that is purely about voting, while partisan can-

²They find statistically significant effects in a larger share of the countries with PR systems. Moreover, they only found a significant effect in New Zealand after the switch to proportional representation. Comparing the size of the effects is difficult because the baseline voting propensity differs between countries.

vassers also try to persuade the subject to vote for a certain candidate or party. Which kind of message that is most effective for mobilizing voters is essentially an open question. On the one hand, partisan messages can provide additional reasons for voting, like affecting policy or expressing support for a certain candidate. On the other hand, voters might respond negatively to a political message of which they disapprove. The latest meta-study of experiments using direct mailing concludes that only mailings with non-partisan messages boost turnout (Green et al. 2013), but it remains to be seen whether the same holds for door-to-door canvassing. Second, it is possible that partisan campaigns target different groups and have a more multifaceted purpose with their campaigns. For example, they might prioritize areas with many swing voters, which might not be the areas best suited for mobilization efforts.

It has proved difficult to conduct experiments with partisan canvassers and most attempts have been crippled by low contact rates or substantial contamination of the control group. One possible explanation could be that it is easier to supervise hired canvassers and university volunteers than it is to control party activists. While the latter are certainly motivated to contact as many voters as possible, they might be less concerned about whether the canvassed voters were assigned to the treatment group. The survey by Green and Gerber (2008) includes three partisan experiments with published results, in which the contact rate ranges from 8 (Nickerson et al. 2006) to 13 per cent (Nickerson 2006) of the target population.³ On the other hand, through cooperation with party organisations, partisan campaigns also have the potential of reaching a much larger number of voters than what can be achieved with non-partisan canvassing. It is therefore not surprising that also some of the most high-powered experiments are conducted by partisan canvassers, like those analysed by Pons (2014) and Pons and Liegey (2013). However, they find very limited mobilizing effects of partisan door-to-door canvassing, which only underscores the need for more experiments in a partisan context.

A second-order election in a country with high turnout

Following the seminal works of Reif and Schmitt (1980), it has become standard to describe elections for the European Parliament as second-order national contests (Hix and Marsh 2007; Marsh 1998). Among other things, the concept implies that voter turnout and intensity of election campaigns are lower than in national elections where executive power is at stake. Most countries in Europe – and Sweden in particular – have a higher turnout than the United States in their national elections. So while the elections for the

³The third study is made by Arceneaux (2007) and includes two subsamples with a contact rate of 7 (canvassing conducted by candidates) and 11 per cent (partisan canvassing by volunteers). In the unpublished studies the contact rates can be even even lower.

European Parliament is the archetype of a second-order election, it takes place in an electorate with a strong habit of voting. Below I discuss how salience and voting habits might condition the effectiveness of door-to-door canvassing and other mobilizing campaigns.

First, in elections with very high turnout, the mobilizing effect is probably smaller. If almost everyone votes also in the absence of a treatment, there are much fewer citizens we have a theoretical chance of mobilizing. More generally, we can assume that the effect is largest for – or even limited to – people who teeter back and forth between voting and not voting. If the distribution of the voting propensity is symmetrical and unimodal, we would expect the share of people with a voting propensity close to 50 per cent to be the largest when also the turnout is close to 50 per cent.⁴ If that assumption is correct, the European elections are suitable for mobilizing campaigns.

Second, the high turnout in most national elections in Europe could indicate that there are strong norms about voting. Many scholars have argued that the reason why "get out the vote" campaigns work is that they induce social pressure and increase the cost for not complying with the norms. The effectiveness of canvassing is therefore dependent on the strength of these social norms. If the norms about voting are stronger in Europe, we would also expect larger mobilizing effects.

Third, some of the largest effects in previous research have been observed in non-competitive elections for school boards and city councils where the turnout has been less than 20 per cent. In such elections, direct contact with the campaign can sometimes be the only way of learning about the candidates, and the canvassing can then be something qualitatively different from the usual mobilizing campaigns. Such low-salience elections are uncommon in Europe and we should think twice before applying the results from those studies in a European context.

The size of the effects

Some of the most cited field experiments have estimated the effect of door-to-door canvassing to an increase in voter turnout by close to 10 percentage points (Gerber and Green 2000; Green et al. 2003; Nickerson 2008). However, they only represent a small share of all canvassing experiments that have been conducted. A recent review by Green et al. (2013) includes 71 canvassing experiments. The authors calculated the average effect in those campaigns, weighted by precision, to be 2.5 percentage points. Because of publication bias and the small samples used in most experiments, this number is probably

⁴Note that it is the voting propensity among the treated, and not in the electorate in general, that matters for the mobilizing effect of the campaign. Citizens with a high probability of being contacted by the canvassers tend to have a higher voting propensity than those who cannot be reached. On the other hand, mobilizing campaigns are often directed to areas with a lower turnout than average.

more accurate than the large effects found in the most cited articles. On the other hand, it is probably an underestimation of the true effects, because most things that can go wrong in an experiment would lead to downward bias in the estimated effect. Besides, many of the experiments study elections with so low turnout that an increase with one or two percentage points would count as quite substantial.

The experience from the few experiments that have been conducted in Europe is less inspiring. No mobilizing effect at all was found for the large-scale canvassing in the Hollande presidential campaign in France in 2012 (Pons 2014). And during the French regional elections in 2010, the effects were limited to certain subgroups (Pons and Liegey 2013). None of the experiments conducted in Denmark have found any statistically significant effects (Bhatti et al. 2014a). The only promising experiment carried out in Europe was conducted in England. In their study, John and Brannan (2008) reported an estimated effect of 6.7 percentage points, which also was statistically significant (although only with one-sided tests and unadjusted standard errors).

There are several reasons to be more optimistic about the expected effects than what previous studies would give reason to be. First, most of the non-existing effects in France and Denmark can be explained by the high turnout in the elections during which the canvassing was conducted. The level of turnout in the Swedish elections for the European Parliament are better suited for a mobilization campaign. Second, the same studies also report uncertainty about compliance. If the contact rate is lower than estimated, or canvassers decide to also do some campaigning in the control group, the true effects will be underestimated. When planning our experiment, we knew that we could not match the sheer volume of contacts reported in Pons (2014), and instead we took every effort to ensure a high rate of compliance.

In the initial planning phase, we wanted to limit the study only to those who voted in the national elections 2010 but stayed at home during the European elections of 2009. This was the group where we expected the largest effect. However, we were afraid that such a narrow sample would be too small, and widened it to all voters where we had information about whether they voted in the previous elections. As a result, we could now test whether our assumption about the heterogenous effect was correct or not.

Data

In Swedish elections, the constituencies are divided into electoral districts with one polling station per district. The size of the district is usually between 1 000 and 2 000 citizens who are entitled to vote. A few weeks after each election, the electoral rolls are accessible for the public as printed books.

We have digitalized the electoral rolls for 17 electoral districts and three

Table 1: Balance of the treatment and control groups

| | Treatment | Control | P-value |
|-------------------|-----------|---------|---------|
| Voted 2009 | 41.4 | 40.8 | 0.554 |
| Voted 2010 | 79.0 | 79.5 | 0.539 |
| Age | 56.7 | 56.7 | 0.948 |
| Household size | 1.9 | 1.9 | 0.493 |
| Women | 52.0 | 53.0 | 0.316 |
| Voting propensity | 45.4 | 45.1 | 0.517 |

Turnout in previous elections, the share of women and the voting propensity are all displayed as percentages. Voting propensity is calculated as predicted values from regressing turnout 2014 on a large set of predictors.

different elections: the 2009 election for the European Parliament, the 2010 national election for the Swedish parliament and – after the field work was finished – the 2014 election for the European Parliament (the outcome variable). The rolls include the person’s name, his or her personal identification number and whether he or she voted in the election. To these records I have added street addresses from the preliminary electoral rolls for the 2014 election. The Swedish register of addresses includes apartment numbers, which makes it easy to identify which individuals live in the same household.

In total, 31 per cent of the citizens eligible to vote were excluded from the sampling frame, with a remaining 11 640 individuals distributed over 7 579 households. Most of them were excluded because they lived somewhere else during the previous elections and I therefore lack information about whether they voted in 2009 and 2010. The households were then randomly allocated to equally sized treatment and control groups, using strata determined by electoral district and household size. That the randomization was successful is supported by the balance between the treatment and the control group that is shown in Table 1. The table presents summary statistics on the treatment and control group, together with the p-values from a t-test of the means. None of the differences are close to being statistically significant.

The personal identification numbers include information about a person’s sex and age. I have also used them to separate the foreign-born from the Swedish-born. From the street addresses in the electoral rolls, I have been able to identify households, calculate the size of each household (this measure only pertains to people eligible to vote) and separate apartments from single-family dwellings. For each contact attempt, the canvassers reported the date and the outcome. If a contact attempt was successful, the canvassers specified who they had been talking to and which canvassers had participated. For each canvasser I have information about sex, age and whether they consider themselves to be ”locally known”. That makes it possible to examine possible identification effects, for which there is some suggestive evidence (Bennion

2005).⁵

I rely on these contact forms to classify subjects as treated or not treated. Because spill-over effects within households can be large (Nickerson 2008), I define treated individuals as everyone living in a household where a door-to-door canvasser has reported a successful contact attempt. It is thus only for single-person households that the treatment effect can be interpreted as the effect of direct contact with a canvasser. The procedures for excluding citizens from the sampling frame, assigning households to the treatment group and creating all the variables mentioned above are described in detail in the supplemental material on the author’s homepage.

Identifying assumptions

All results in this paper rest on the assumption that the allocation to the treatment and the control group is random. Because the allocation process was entirely under my control, randomization was easily fulfilled. However, I also assume *one-sided noncompliance*, *non-interference* and *excludability*. If any of those assumptions are broken, the estimates will be biased. Below I describe those assumptions in more detail.

Throughout the paper I assume one-sided noncompliance, which means that no subjects assigned to the control group have been visited by canvassers. If this assumption does not hold, all estimated effects are biased downwards. In a telephone survey conducted after the election, only 6 percent of the citizens in the control group responded that they had been visited by party workers, compared to 79 per cent in the treatment group.⁶ The survey not only shows that the experiment was conducted according to protocol, but also that control group contamination caused by door-to-door canvassers from other parties is negligible. The reason why the sample of survey respondents reported a higher contact rate (79 per cent) than the canvassers (60 per cent) is probably that those who were reached by the survey company also had a larger probability of being reached by the canvassers.

I also assume non-interference and excludability. Non-interference means that every subject’s probability of being treated, and every subject’s propensity to vote, is unaffected by other subjects’ treatment assignment. Excludability means that the outcome is affected by the actual treatment and not by the treatment assignment. The largest threat to both these assumptions

⁵While Bennion (2005) finds that student canvassers only affected younger voters, we do not know if this was an effect of similar age or if we would have found the same impact if the canvassers were older.

⁶The survey included a total of 1 000 randomly selected respondents, divided equally between the treatment and the control group. Approximately 75 per cent of the respondents claimed that they voted in the election (79 per cent in the treatment group and 70 per cent in the control group). The actual turnout among subjects in this experiment was 46.1 per cent, which shows why we cannot rely on self-reported turnout in this kind of study.

is probably that canvassed citizens might inform their friends about the upcoming election and even tell them to vote. Unless non-treated subjects in the treatment group are more likely than subjects in the control group to interact with treated individuals, all estimated effects are biased downwards if the assumptions about non-interference and excludability do not hold. Several measures have been taken to reduce the extent of excludability violations. Because people are likely to interact with other people in the same household, everyone in a household is classified as treated if at least one of the household members has talked to a canvasser. To leave non-contacted households without treatment, canvassers were instructed to not leave any leaflets or other information after unsuccessful contact attempts.

This study has a much higher contact rate than most similar studies, which affects the estimated effects in several ways that are worth pointing out.⁷ First, the large share of compliers substantially increases statistical power. While the loss in precision that results from low compliance can be compensated for through a large sample size, there are not many canvassing experiments that match the precision of this study.⁸ Second, the high contact rate probably means that the estimated average treatment effect among compliers is closer to what the average treatment effect would be if everyone would have been treated. This is important, because most theories of political behaviour concern all citizens eligible to vote and not only those with a high probability of being canvassed. The data collected in this experiment also confirms that people with a high contact probability are very different from those that are harder to reach.⁹ Third, whether the effect in high contact rate experiments provides a better estimate or not for what to expect in other canvassing campaigns is an open question and depends on the reasons why some people are contacted while others are not. On the one hand, one of the reasons for the high contact rate is that the canvassing has been conducted during different times and different days. That should make the results more generalizable than if the canvassing had only been conducted during either office hours or weekends. On the other hand, if the high contact rate primarily reflects a higher share of those subjects that

⁷Most successful canvassing experiments have contact rates around 30 per cent. Melissa Michelson's study of a local school board election (Michelson 2003) is the only similar study I know with a higher contact rate than mine, but the sample size of her study is too small to yield effects that are significant at conventional levels.

⁸To my knowledge, the only experiments with better precision are those conducted on the SCOPE campaign in the 2008 presidential election (Bedolla and Michelson 2012) and the French Socialist Party's campaigns in the 2010 regional elections (Pons and Liegey 2013) as well as in the 2012 presidential elections (Pons 2014).

⁹Those living in households where the first contact attempt was successful had a turnout of 46 per cent in the 2009 election and 54 per cent 2014. By contrast, among citizens in households where the canvassers made at least three contact attempts without reaching anyone in the household, the average turnout was 30 and 32 per cent in the 2009 and 2014 elections, respectively.

are hard to reach, owing to the dedicated canvassers who visited the same address up to three times, this experiment might give too much weight to groups with a low probability of being contacted.

Results

If the assumptions discussed above hold, comparing the turnout between these groups is enough to identify the causal effect of canvassing. To improve precision and adjust for intra-household correlations, I employ an instrumental variable regression framework. In the next section I describe this approach in more detail and present the results for the full sample. I then proceed by examining how the effect differs depending on voting propensity, household size and other characteristics of the citizens.

Main effects

The most important finding of this paper is shown in Table 2. From left to right, the table displays the joint number of observations in the treatment and control group, the share of households in the treatment group that were reached by the canvassers and the voter turnout in first the control group and then in the treatment group. The last four columns show the estimated causal effects of door-to-door canvassing, first estimated in a bivariate model and then in a model which includes a set of covariates. The two effects can be interpreted as the expected change in the probability of voting that results from living in a household that was visited by a door-to-door canvasser.

Because the assignment to the treatment and control group is random, the difference between the two groups can be interpreted as the causal effect of being assigned to the treatment group, given that the assumptions about one-sided noncompliance and non-interference hold. This is usually called an intent-to-treat (ITT) effect and it is not shown in the table. To estimate the effect of actually being treated – that is, living in a household that was visited by a canvasser – I could divide the ITT with the treatment share in the subgroup. Compared to the ITT estimates, these estimates only add the assumption of excludability.¹⁰

The difference in turnout between the treatment group and the control group was in this experiment 2.32 percentage points (47.21 – 44.89). If I would divide 2.32 with the treatment share of 0.595 I would get the treatment effect of 3.9 percentage points that was reported as the bivariate effect. Because people living in the same household are not independent from each

¹⁰These are sometimes called LATE (local average treatment effect, see Angrist et al. (1996)) or CACE (complier average causal effects, see Gerber and Green (2012)) effects. Note that these estimates only tell how voters that live in contacted household are affected by canvassing. We cannot be sure that those in the treatment group who were not contacted would have been affected in the same way.

Table 2: Full sample results

| Subgroup | Obs | Visited | Turnout | | Bivariate | | Covariates | |
|-------------|-------|---------|---------|-------|-----------|-----|------------|-----|
| | | | Control | Treat | Effect | SE | Effect | SE |
| Full sample | 11475 | 59.5 | 44.9 | 47.2 | 3.9** | 1.9 | 3.6** | 1.6 |

Robust standard errors clustered on households. All tests are two-tailed.

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

other, I have instead used an instrumental variable regression framework, which allows me to calculate robust standard errors that are clustered on the household level. All models are estimated using 2SLS.

As illustrated by equation 1, I regress the voting decision on the treatment outcome (living in a visited household) and instrument the treatment with the allocation to the treatment group. In the bivariate case, the two procedures yield identical estimates of the effect, but not correcting for the strong within-household correlations would deflate the standard errors substantially.

$$\begin{aligned}
 \text{Vote} &= a_1 + b_1 \widehat{\text{Visited}} + \mathbf{b}\chi + \epsilon \\
 \widehat{\text{Visited}} &= a_2 + b_2 \text{Assignment}
 \end{aligned}
 \tag{1}$$

The regression framework makes it easy to add more explanatory variables to the model, symbolized in the equation through the vector χ . If the variables are correlated with the outcome variable, and are unaffected by the treatment, they increase the precision of the estimates without inducing any measurable bias. The covariates I have included are earlier voting behaviour (two dummies for voting in the 2009 and the 2010 elections and a dummy for whether the person is a first time voter), age, age squared and dummy variables for the 17 electoral districts.

The seventh column of Table 2 shows the estimated treatment effect from the multivariate model. When the results differ from the bivariate estimates, I am inclined to believe more in the regressions with covariates, but if the randomization was conducted properly the differences should be small. For the full sample, the estimated effect of living in a visited household is 3.6 percentage points and it is significant at the 95 per cent level. Another way of expressing this result is that for every 28 treated citizens, one more goes to the ballots. Because there are 1.9 persons living in each household on average, this effect amounts to one more vote for every 15 canvassed doors.

Subgroups

In addition to estimating average effects, it is important to analyse how the effects differ between subgroups. Such knowledge is valuable when planning a canvassing campaign, because it allows the canvassers to focus their efforts where they count. Such knowledge is also of scholarly interest because

heterogeneous effects between groups of citizens can tell us something about the mechanisms behind successful canvassing and why some voters are more easily mobilized than others.

Before I proceed to how the effects differ between subgroups, there are three things I want to emphasize about how we should interpret the effects. First, among the people living in a contacted household, the share that actually talked to a canvasser can differ across subgroups. If this share is larger in one subgroup than in another, and the direct effect of contact with a canvasser is larger than the spill-over effect on the other members of the household, this will then appear in the table as a heterogeneous effect of canvassing. Second, belonging to a subgroup might be correlated with also belonging to another subgroup or with unobserved covariates that affect the effectiveness of canvassing. For example, the differences between people living in single-family and multiple-family dwellings might reflect differences in wealth or income rather than housing per se. Third, the risk of falsely rejecting a true null hypothesis increases with the number of subgroups. When drawing conclusions, we must therefore assess the statistical results in comparison with our theoretical priors and with what has been found in previous research.

It is clear from the table that there are large differences between subgroups, not only in the level of turnout but also in how the turnout was affected by the treatment. Starting with the voting history, I can only find an effect among those who voted in the 2010 elections to the national parliament but refrained from voting in the 2009 elections to the European Parliament. The estimated effect in this group is 6.0 percentage points higher turnout for those living in a household that was visited by a canvasser. I had expected the effect in this group to be larger than for the others, because they are obviously prepared to vote when the circumstances are right but nonetheless refrained from voting in the last election. The idea that mobilization efforts are most effective among occasional voters is also supported by previous research (Arceneaux and Nickerson 2009; Niven 2001). Still, the difference between the groups is larger than I had expected. An equally large effect is found among first-time voters, but they are so few that the estimate is imprecise and not statistically significant.

Many have emphasized the social nature of voting and how social pressure is the mechanism behind successful mobilization (Ali and Lin 2013; Gerber et al. 2008; Shachar and Nalebuff 1999; Uhlaner 1989). Besides, both simulations (Fowler 2005) and experimental evidence (Bhatti et al. 2014a; Bond et al. 2012) suggest that effects of secondary mobilization are much larger than primary effects. I was still surprised to see larger effects for people living in houses and large households, compared to the non-existing effects for those living in apartments or single-person households.¹¹ If anything, I

¹¹The small effect in apartment buildings could theoretically reflect contamination of the

Table 3: Results in different subgroups

| Subgroup | Obs | Visited | Turnout | | Bivariate | | Covariates | |
|-----------------------|-------|---------|---------|-------|-----------|-----|------------|-----|
| | | | Control | Treat | Effect | SE | Effect | SE |
| Full sample | 11475 | 59.5 | 44.9 | 47.2 | 3.9** | 1.9 | 3.6** | 1.6 |
| <i>Voting history</i> | | | | | | | | |
| Both elections | 4659 | 63.7 | 75.3 | 75.9 | 1.0 | 2.3 | 0.8 | 2.3 |
| Only 2010 | 4337 | 57.7 | 27.0 | 30.7 | 6.3** | 2.8 | 6.0** | 2.8 |
| None of them | 1408 | 48.6 | 7.5 | 6.8 | -1.5 | 3.0 | -1.4 | 3.0 |
| First-time voters | 760 | 60.6 | 36.2 | 40.2 | 6.5 | 6.2 | 6.4 | 6.0 |
| <i>Dwelling</i> | | | | | | | | |
| Single-family | 4958 | 72.4 | 51.5 | 55.8 | 6.0** | 2.6 | 5.1** | 2.2 |
| Multiple-family | 6517 | 49.6 | 39.9 | 40.6 | 1.5 | 2.8 | 1.7 | 2.4 |
| <i>Household size</i> | | | | | | | | |
| Single person | 4241 | 46.9 | 40.2 | 40.0 | -0.5 | 3.2 | -0.3 | 2.7 |
| Two persons | 5091 | 67.1 | 50.8 | 52.6 | 2.6 | 2.8 | 3.0 | 2.2 |
| More than two | 2143 | 66.0 | 40.1 | 48.7 | 13.0*** | 4.8 | 10.8** | 4.3 |
| <i>Gender</i> | | | | | | | | |
| Male | 5450 | 58.7 | 46.1 | 47.8 | 2.9 | 2.4 | 3.4* | 2.0 |
| Female | 6025 | 60.1 | 43.8 | 46.6 | 4.7** | 2.2 | 3.6* | 1.9 |
| <i>Age</i> | | | | | | | | |
| 18-29 | 1401 | 57.8 | 34.3 | 38.0 | 6.4 | 4.9 | 8.5* | 4.7 |
| 30-59 | 4527 | 57.3 | 44.0 | 47.4 | 6.0* | 3.1 | 4.5* | 2.6 |
| 60-103 | 5349 | 61.5 | 48.2 | 49.8 | 2.5 | 2.6 | 2.6 | 2.1 |
| <i>Land of birth</i> | | | | | | | | |
| Native-born | 10037 | 59.3 | 46.0 | 48.1 | 3.6* | 2.0 | 3.3* | 1.7 |
| Foreign-born | 474 | 59.9 | 44.4 | 39.6 | -7.9 | 8.3 | -5.1 | 6.4 |

The subgroup "none of them" does not include first-time voters. Robust standard errors clustered on households. All tests are two-tailed. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

would have expected the opposite. A large share of people living in large households did not talk to the canvasser themselves, but were only affected by so called spill-over effects.

Previous research offers little guidance about how the size of the household affects the effectiveness of canvassing. While Bhatti et al. (2014a) analyse the mobilizing effects of short text messages and find substantially larger *spill-over* effects in large households, Nickerson and Rogers (2010) found that phone calls were more effective in small households. Gerber et al. (2008) did not find any effect at all of household size. As one could suspect, there is a

control group, if there is a large amount of between-household interaction, but I find that unlikely in a Swedish context. Moreover, within the control group, the turnout was higher in single-family houses than in apartments, controlling for turnout in previous elections.

positive correlation between living in a house and living in a large household, but both categories have a positive effect when controlling for the other. In fact, both the effects are driven by a very large estimated effect for large households (>2 members) living in a house.

One interpretation is that in larger households, other household members might observe or even participate in the discussion, thereby increasing the social pressure of voting. There are indications of such a mechanism in the data. Controlling for household size, the turnout was 7 percentage points higher among citizens that participated in the conversation together with at least one other household member, compared to citizens that participated on their own. However, the evidence can at best be interpreted as suggestive. Because the number of discussants is not randomized the results might reflect that politically interested citizens are more willing to engage in a discussion with the canvassers.¹²

It should also be mentioned that most two-person households are couples, rather than single parents or friends living together. In almost 90 per cent of these households, the two household members are of different sex and have an age difference smaller than 15 years.¹³

In contrast to Pons and Liegey (2013), I do not find larger effects among immigrants. On the contrary, among foreign-born, the turnout is in fact smaller in the treatment group than in the control group. However, the number of observations is small and the difference is not significant. In contrast to studies that find large positive effects among foreign-born and second-generation immigrants, most canvassers in this study are native-born. I also find that the effect declines with age. There is a gender difference in the bivariate regression that disappears when the covariates are included. The reason for this is that there was a (not significantly) larger share of female 2009 voters in the treatment group than in the control group.

In theory it would be possible to differentiate between different kinds of heterogenous effects by including them as interaction variables in a regression framework. However, most differences between subgroups are not significant at the conventional levels (note the distinction between significant differences between subgroups and statistically significant in some subgroups but not others).

It has become popular to analyse heterogenous effects of canvassing in terms of the citizens' underlying propensity to vote. In an attempt to escape the categorical nature of the voting history, I have estimated the probability of voting for individuals in the control group by regressing the decision to

¹²Indeed, running placebo tests on turnout in previous elections as well as controlling for earlier voting behaviour clearly shows that there are systematic differences between households depending on how many people participate in the discussions.

¹³The age difference between the youngest and the oldest person in larger households is above 15 years in 95 per cent of the households.

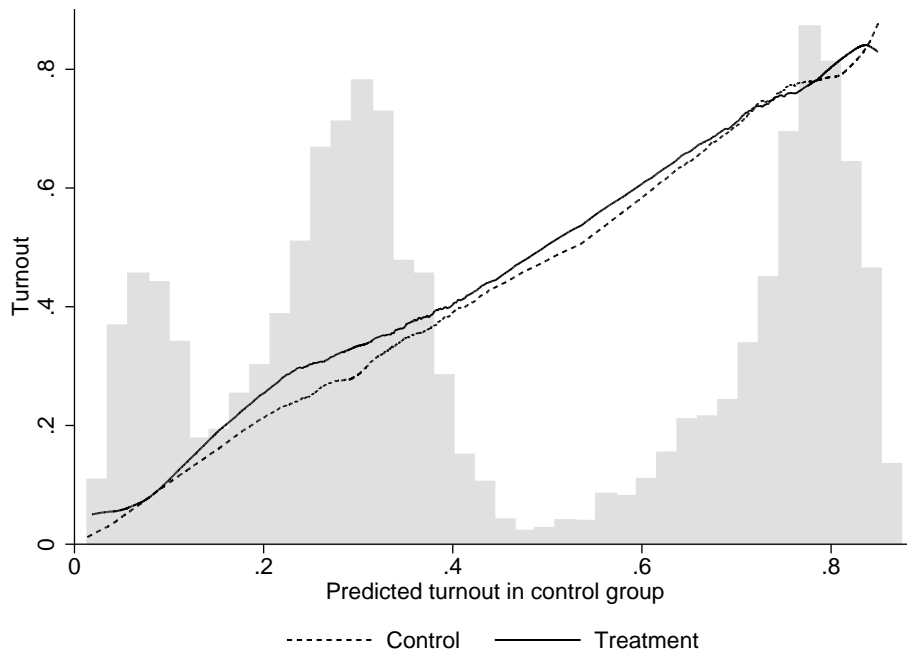


Figure 1: Turnout over predicted turnout in control group

vote on a large number of predictors.¹⁴ These predictions form the horizontal axis of Figure 1 and the histogram shows how they are distributed. In front of the histogram are two less lines, showing the turnout in both the treatment and the control group as locally weighted functions of the predictions. The gap between the two lines makes up the estimated ITT effect. This effect is largest when voters in the control group have a predicted voting propensity between 20 and 35 per cent and remains positive up to approximately 70 per cent. But as the three peaks in the histogram make apparent, the variation in the voting propensity is not really continuous but focused around these three clusters. The clusters correspond to the voting propensity given by previous voting behaviour. The other predictors do not add much information and as a consequence there are few citizens with a predicted probability of voting between 40 and 60 per cent. We should therefore be careful in drawing other conclusions than that there does not seem to be a mobilizing effect among voters with a very low or very high propensity to vote.

¹⁴The model is a logit with the following predictors: voting history, age, age squared, female, form of housing, household size and district dummies. Migrant status was excluded because of missing observations for first-time voters.

Vote choice and other effects

Shortly after the election, a telephone survey was conducted with 1000 respondents. Of the 743 respondents who reported that they had voted, 47 per cent in the treatment group stated that they voted for the Social Democrats, compared to 39 per cent in the control group. That difference corresponds to an intent-to-treat effect of 8 percentage points, or a treatment effect of 10–13 percentage points, depending on what assumptions we make about the contact rate among these respondents. The effects are also statistically significant at the 95 per cent level.

However, I would advise against interpreting these as causal effects on vote choice. To begin with, the effect is both surprisingly large and imprecisely estimated, increasing the plausibility of serious sampling error. Moreover, it is possible that canvassing increases the probability of lying due to social desirability bias. According to the survey, the turnout was 9 percentage points higher in the treatment group than in the control group, which indicates a much larger effect on self-reported than actual turnout.

In the supplementary information on my homepage, I present results from models that were excluded from this paper. Those include probit models, estimations of spill-over effects within the household and models where I analyze whether the mobilizing effect is conditioned by the timing of the contact and the characteristics of the canvassers.

Conclusions

This field experiment has shown that door-to-door canvassing can increase voter turnout also in a European context. Living in a household that was visited by party workers is estimated to increase the probability of voting by 3.6 percentage points.

While the estimated effect is twice as large among occasional voters and first-time voters, I do not find any effect at all among chronic non-voters or voters with a very high vote propensity. This result is in line with previous research, which has shown that canvassing is most effective for citizens who are close to indifferent about voting.

I was surprised to find that the average effect was much greater in large households than in single-person households. If this finding can be replicated in other experiments, it has major implications both for how to conduct canvassing campaigns and for our understanding of voter behaviour. For campaign generals, it means that focusing the efforts on large households will not only increase the number of citizens who are reached by a single contact attempt, but also improve the average effect on every household member. For scholars of political behaviour, this finding underlines the social nature of voting. To learn more about why household size matters, we should design experiments with placebo messages where we can examine how the

mobilizing effect is conditioned on the number of household members that are at home or participate in the conversation.¹⁵

Door-to-door canvassing is widely acknowledged as one of the most effective means for increasing voter turnout. It is therefore puzzling that such campaigns are much more common in countries with majoritarian systems than under proportional representation. The results in this paper confirm that door-to-door canvassing mobilize voters also in countries with PR systems and partisan canvassers. For scholars of comparative electoral research, the reasons why parties in many countries are so reluctant towards this form of campaigning remains to be explored. Are there perhaps institutional barriers that prevent our democracies from working as well as they could?

¹⁵Delivering a placebo message to the control group, which could be any message that does not affect the probability of voting, would make it possible to estimate what the turnout would have been – conditioned on the number of conversation participants – in the absence of treatment.

References

- Ali, Nageeb N. and Charles Lin (2013). ‘Why people vote: Ethical motives and social incentives’. *American Economic Journal: Microeconomics* 5.2, pp. 73–98.
- Angrist, Joshua D., Guido W. Imbens, and Donald B. Rubin (1996). ‘Identification of causal effects using instrumental variables’. *Journal of the American Statistical Association* 91.434, pp. 444–455.
- Arceneaux, Kevin (2007). ‘I’m asking for your support: The effects of personally delivered campaign messages on voting decisions and opinion formation’. *Quarterly Journal of Political Science* Oct 2005, pp. 43–65.
- 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, pp. 1–16.
- Bedolla, Lisa García and Melissa R. Michelson (2012). *Mobilizing Inclusion: Transforming the Electorate through Get-Out-the-Vote Campaigns*. The Institution for Social and Policy St. Yale University Press.
- Bennion, Elizabeth A. (2005). ‘Caught in the Ground Wars: Mobilizing Voters during a Competitive Congressional Campaign’. *The ANNALS of the American Academy of Political and Social Science* 601.1, pp. 123–141.
- Bhatti, Yosef and Kasper M. Hansen (2012). ‘The effect of generation and age on turnout to the European Parliament – How turnout will continue to decline in the future’. *Electoral Studies* 31.2, pp. 262–272.
- Bhatti, Yosef, Jens Olav Dahlgaard, Jonas H. Hansen, and Kasper M. Hansen (2014a). *How voter mobilization spread in households and families - The use of short text messages on cell phones to boost turnout*.
- Bhatti, Yosef, Jens Olav Dahlgaard, and Jonas Hedegaard Hansen (2014b). *Kan man øge valgdeltagelsen?* CVAP Working Paper Series. University of Copenhagen.
- Blais, André and Kees Aarts (2006). ‘Electoral Systems and Turnout’. *Acta Politica* 41.2, pp. 180–196.
- Bond, Robert M., Christopher J. Fariss, Jason J. Jones, Adam D. I. Kramer, Cameron Marlow, Jaime E. Settle, and James H. Fowler (2012). ‘A 61-million-person experiment in social influence and political mobilization’. *Nature* 489.7415, pp. 295–298.
- Cox, Gary W. (1999). ‘Electoral Rules and the Calculus of Mobilization’. *Legislative Studies Quarterly* 24.3, pp. 387–419.
- Cox, Gary W. (2015). ‘Electoral Rules, Mobilization, and Turnout’. *Annual Review of Political Science* 18.1, pp. 49–68.
- Fowler, James H. (2005). ‘Turnout in a Small World’. In: *Social Logic of Politics*. Ed. by Alan Zuckerman. Temple University Press, pp. 269–287.
- Gagatek, Wojciech (2010). ‘Campaigning in the European Parliament elections’. In: *The 2009 Elections to the European Parliament*. Ed. by Wojciech Gagateg. European University Institute. Chap. 2, pp. 13–19.

- Gerber, Alan S. and Donald P. Green (2000). ‘The effects of canvassing, telephone calls, and direct mail on voter turnout: A field experiment’. *American Political Science Review*.
- Gerber, Alan S. and Donald P. Green (2012). *Field experiments : Design, analysis, and interpretation*. New York: W. W. Norton.
- Gerber, Alan S., Donald P. Green, and Christopher W. Larimer (2008). ‘Social Pressure and Voter Turnout: Evidence from a Large-Scale Field Experiment’. *American Political Science Review* 102.1, pp. 33–48.
- Górecki, Maciej a. and Michael Marsh (2012). ‘Not just ‘friends and neighbours’: Canvassing, geographic proximity and voter choice’. *European Journal of Political Research* 51.5, pp. 563–582.
- Green, Donald P. and Alan S. Gerber (2008). *Get out the vote : How to increase voter turnout*. 2nd ed. Washington, D.C.: Brookings Institution Press.
- Green, Donald P., Alan S. Gerber, and David W. Nickerson (2003). ‘Getting Out the Vote in Local Elections: Results from Six Door-to-Door Canvassing Experiments’. *The Journal of Politics* 65.04, pp. 1083–1096.
- 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, pp. 27–48.
- Herrera, Helios, Massimo Morelli, and Thomas Palfrey (2014). ‘Turnout and power sharing’. *Economic Journal* 124.574.
- Hix, Simon and Michael Marsh (2007). ‘Punishment or protest? Understanding European parliament elections’. *Journal of Politics* 69.2, pp. 495–510.
- John, Peter and Tessa Brannan (2008). ‘How Different Are Telephoning and Canvassing? Results from a ‘Get Out the Vote’ Field Experiment in the British 2005 General Election’. *British Journal of Political Science* 38.03, pp. 565–574.
- Karp, Jeffrey A. and Susan A. Banducci (2008). ‘Political efficacy and participation in twenty-seven democracies: how electoral systems shape political behaviour’. *British Journal of Political Science* 38.2, pp. 311–334.
- Karp, Jeffrey A., Shaun Bowler, and Susan A. Banducci (2003). ‘Electoral systems, party mobilization and turnout: Evidence from the European parliamentary elections’. *Journal of Elections, Public Opinion & Parties* 13.1, pp. 210–225.
- Karp, Jeffrey A., Susan A. Banducci, and Shaun Bowler (2008). ‘Getting Out the Vote: Party Mobilization in a Comparative Perspective’. *British Journal of Political Science* 38.01, pp. 91–112.
- Lijphart, Arend (1997). ‘Unequal Participation: Democracy’s Unresolved Dilemma’. *American Political Science Review* 91.1, pp. 1–14.
- Lijphart, Arend (1999). *Patterns of democracy, government forms and performance in thirty six countries*. New Haven: Yale University Press.

- Marsh, Michael (1998). 'Testing the second-order election model after four European elections'. *British Journal of Political Science* 31.1997, pp. 591–607.
- Michelson, Melissa R. (2003). 'Getting out the Latino vote: How door-to-door canvassing influences voter turnout in rural central California'. *Political Behavior* 25.3, pp. 247–263.
- Nickerson, David W. (2006). *Forget me not? The importance of timing in voter mobilization*.
- Nickerson, David W. (2008). 'Is voting contagious? Evidence from two field experiments'. *American Political Science Review* 102.1, pp. 49–57.
- Nickerson, David W. and Todd Rogers (2010). 'Do you have a voting plan? Implementation intentions, voter turnout, and organic plan making'. *Psychological Science*.
- Nickerson, David W., Ryan D. Friedrichs, and David C. King (2006). 'Partisan Mobilization Campaigns in the Field: Results from a Statewide Turnout Experiment in Michigan'. *Political Research Quarterly* 59.1, pp. 85–97.
- Nielsen, Julie Hassing (2014). 'Why use experiments in EU studies?' *Comparative European Politics*, pp. 1–19.
- Niven, David (2001). 'The limits of mobilization: Turnout evidence from state house primaries'. *Political Behavior* 23.4, pp. 335–350.
- Pons, Vincent (2014). 'Does Door-to-door Canvassing Affect Vote Shares? Evidence from a Countrywide Field Experiment in France'. July.
- Pons, Vincent and Guillaume Liegey (2013). *Increasing the Electoral Participation of Immigrants – Experimental Evidence from France*.
- Powell, G. Bingham (1989). 'Voting Turnout in Thirty Democracies: Partisan, Legal and Socio-Economic Influences'. In: *Electoral Participation: a comparative analysis*. Ed. by Richard Rose. Beverly Hills, CA: Sage, pp. 396–412.
- Rainey, Carlisle (2015). 'Strategic mobilization: Why proportional representation decreases voter mobilization'. *Electoral Studies* 37, pp. 86–98.
- Reif, Karlheinz and Hermann Schmitt (1980). 'Nine Second-order National Elections – a Conceptual Framework for the Analysis of European Election Results'. *European journal of political research* 8, pp. 3–44.
- Shachar, Ron and Barry Nalebuff (1999). 'Follow the leader: Theory and evidence on political participation'. *American Economic Review* 89.3, pp. 525–547.
- Uhlaner, Carole J. (1989). 'Rational Turnout: The Neglected Role of Groups'. *American Journal of Political Science* 33.2, pp. 390–422.