BRINGING THE PARTY BACK IN:
Mobilization and Persuasion in Constituency Election Campaigns

D.Phil Thesis

Florian Foos

Department of Sociology and Nuffield College

University of Oxford

492321

Introduction, Copyright © 2015 by Florian Foos

Chapter 1 ‘The Heuristic Function of Party Affiliation in Voter Mobilization Campaigns’, Copyright © 2015 by Florian Foos & Eline de Rooij

Chapter 2 ‘Household Partisan Composition and Campaign Mobilization’, Copyright © 2015 by Florian Foos & Eline de Rooij

Chapter 3 ‘Reaching Across the Partisan Divide’, Copyright © 2015 by Florian Foos

Chapter 4 ‘Impersonal but Noticeable?’, Copyright © 2015 by Florian Foos

Chapter 5 ‘The National Effects of Subnational Elections’, Copyright © 2015 by Elias Dinas & Florian Foos

Conclusion, Copyright © 2015 by Florian Foos

All rights remain with the authors.

Foos, Florian. 2015. Bringing the Party Back in: Mobilization and Persuasion in Constituency Election Campaigns. Oxford: D.Phil Thesis.

ABSTRACT

In this thesis, I report the results from the first randomized field experiments conducted in collaboration with party-affiliated candidates and campaigns in the United Kingdom. The papers presented as part of this thesis test both the limits and possibilities of campaign influence, in a partisan political environment. During election campaigns parties provide signals to voters, voluntarily or involuntarily imposing a structure, and thereby constraints, on individuals' electoral decisions. By integrating insights about heuristic and social decision-making into the experimental campaign literature, I formulate testable hypotheses about the direct and indirect effects of party cues on campaign mobilization and persuasion.

The first paper, *The Heuristic Function of Party Affiliation in Voter Mobilization Campaigns,* addresses how the provision of party cues, used during campaign phone calls, affects turnout among party supporters, opponents and unattached voters. The second paper on *Household Partisan Composition and Voter Mobilization,* explores the spillover effects from the previous experiment, testing whether campaign-induced mobilization between household members is conditioned by the partisan composition of a household, and the partisan intensity of a campaign message. Paper three investigates if candidates who are *Reaching Across The Partisan Divide* can win over supporters of rival parties. In the fourth paper, I test if *Impersonal, But Noticeable* methods of voter contact, such as door hangers and text messages, affect the turnout decisions of partisans and unattached voters. The final paper, *The National Effects of Subnational Representation*, highlights the importance of local party organization for the outcomes of national elections. The results of this thesis show the electoral consequences of direct and indirect interactions between campaigns and voters of different partisanship, and point to strategies that allow constituency campaigns to successfully navigate challenging partisan environments.

ACKNOWLEDGEMENTS

This thesis would not have been written, and these experiments would never have been conducted, had it not been for two of the most impressive campaigners in British politics. I owe Caroline Badley and Rowenna Davis a tremendous amount, for their trust, hours invested, and willingness to support experimental research. Over three years I have learnt a lot about campaigns, and most of it in the field. I met people who genuinely care about their constituents, who get excited if they are able to change something locally, and who work very hard, much harder than I do. I have been very lucky that Caroline and Rowenna agreed to collaborate on this research.

I am also deeply indebted to all the volunteers who have worked on my field experiments without remuneration over the past three years. Without their support, scripts would not have been printed, calls would not have been made, interviews would not have been conducted, and data would not have been entered. Some are my friends, but others are good people I had never met before. I am particularly grateful to Charles di Castiglione for his excellent work on the Birmingham experiment, to Lisa Mitchell and Elizabeth Jackson for their invaluable help in Southampton, and to Ben Sayah, Richard Johnson, and Clare Ford for their support with a telephone survey. I am also thankful to my good friends Meghan Laws, Paul Stoddart, Silas Walton, Toni Rodon, and Zsuzsi Szőllősi for proof reading, and for putting in that extra hour more than once.

Of course none of this would have happened without Marcus Roberts' enthusiasm. I first met Marcus in February 2012 after a short email exchange. Marcus was the first person I talked to within the Labour Party who was excited about the idea of doing GOTV experiments. Having worked on countless US campaigns, and as the field director for Ed Miliband's leadership campaign, he was familiar with the

Gerber and Green experiments. Better even, Marcus offered to open his address book for me. I was in the room for less than ten minutes, when he started calling every party organizer he knew who could be interested in collaborating with me. It was him who put me in touch with Caroline and Rowenna. Marcus has been supportive throughout, and willing to stand up for me whenever needed. Within the Labour Party I am also thankful to Andy Hull, who was a great help getting started, to the Fabians' Richard Speight, and to Kevin Cunningham and Tim Waters of the Labour Analytics Team.

Working with practitioners has been a great privilege, and so has been my work with some of the most brilliant minds in academia. First, I want to thank my supervisors at Oxford, Steve Fisher and Geoffrey Evans. I would not have been able to research for my D.Phil at Oxford without Geoff's support. He single-handedly arranged funding for me. I also thank him for the good advice, and the discussions we had over the years about my papers and about academic life more generally. I am sorry that I still have to write a single paper on class non-voting. Steve Fisher has been my academic supervisor since I arrived at Oxford. He was hands-on in the beginning when it counted, and then gave me the freedom to follow my own path and research ideas. From the outset, and despite the risks to my thesis, he was very supportive of my plans to conduct field experiments. I was lucky to have two supervisors who allowed me to follow my ideas wherever they let me. I also gratefully acknowledge the generous funding provided by the Nuffield Sociology Group and the Nuffield Politics Group that allowed me to concentrate on my research while undertaking graduate study. Within the Sociology Department I am especially grateful to Anne Millard who has been a constant source of help with the administrative perils of graduate study and a great source of support. Moreover, I am

thankful to an inspiring generation of Political Scientists who coincided with me at Oxford, especially to Catherine de Vries, Sergi Pardos-Prado, and Anja Neundorf.

Three of the papers included in this thesis are the results of long-standing collaboration with Eline de Rooij and Elias Dinas. I first started talking to Eline about campaign experiments when I was quasi-randomly allocated the seat next to her at Nuffield High Table. Eline had worked with Don Green at Yale, and after an engaging dinner conversation, we decided to start experimenting in Britain. I regret that Eline now lives and works in Vancouver, Canada, but our long-distance collaboration is going strong and the results are visible in this thesis. Without Eline, there probably would not have been any experiments. But most importantly, she has been great to work with over the past three years, and I have learned a lot from her.

Elias Dinas has been my mentor and friend since I arrived at Nuffield as a Master student. He got me interested in causal inference when I barely knew how to run an OLS regression. He was the first person at Oxford to believe in me, and to take me seriously as a researcher. It is no surprise then that Elias had by some margin the biggest influence on my academic development over the past four years. I will miss the countless discussions we have had over coffee, or over a drink until late at night. Without Elias, this thesis would have been mediocre at best.

The second person that had a great influence on my academic development is Don Green. Eline introduced me to Don in summer 2012, and he invited me to come to Columbia to learn about experiments and to work on my thesis. To join Don at Columbia was the best academic decision I ever took. Most what I know about experimental methods, I learned in the course he taught at Columbia in Spring 2013. I had never seen a Professor who was more excited about teaching a class than Don was about teaching his course on field experimentation. Since then, he has read every

single draft I have ever sent him, and his suggestions and advice have greatly improved the papers included in this thesis. I know only one person who is equally passionate about field experiments, and that is Peter John, who has been a great source of advice on my work in the past year. I also thank all of the exceptional experimentalists and behavioral scientists who coincided with me at Columbia, especially Alex Coppock, Yegor Lazarev, Adam Zelizer, and Andrea Lozano. Alex has helped me with code on several occasions, and I am thankful to him for his advice.

I also thank all my friends and colleagues at Nuffield College, particularly Dingeman Wiertz who started as a graduate student with me, and who I owe both academically and personally. Over the years I learned more about statistics from Dingeman than in any statistics course, and I only realized that I needed a schedule once our schedules no longer overlapped. I also thank my Nuffield cohort, especially Ksenia Northmore-Ball, Stefanie Reher, Matilde Gawronski and Silvana Tarlea, and my dear friends Hector Solaz, Christian Arnold, Kristen Lau, Nemanja Dzuverovic, Mathieu Ichou, and Giulia Bistagnino for many deep and not so deep conversations, and lasting friendships.

A study based on interviews with Oxford students recently found that while many acknowledged the educational and social disadvantages that lead some students to be underrepresented at the university, they "did not make connections to the advantages that they themselves enjoyed".[1] During the interviews students attributed their educational success mostly to intelligence and hard work. The truth is that my life has been full of advantages, and I would have never ended up at Oxford without

---

Warikoo, Natasha Kumar and Christina Fuhr. 2014. "Legitimating status: perceptions of meritocracy and inequality among undergraduates at an elite British university." *British Educational Research Journal* 40(4): 699–717.

the support I received earlier in life.

I had the advantage of having a brilliant social science teacher in high school, Gunther Rohe. I signed up to write a paper with him in my last year, and I remember being shocked when I found my entire 20 pages draft crossed out with red ink. Identifying Helmut Kohl as the cause of declining turnout in Germany occurred before my engagement with causal inference. After my initial shock, he suggested that I try the scientific method, and encouraged me to conduct a survey in my high school. The resulting project kept me busy for the rest of the school year. I sampled students by class, calculated odds by hand, and found that political participation was stratified by education, gender and family background. The results intrigued me. I probably would not have decided to become a social scientist were it not for Gunther Rohe.

For both my parents, Ulrike and Stefan, education played a very important role in how they brought my brother Alex and me up. Although they have not enjoyed the privilege of a university education, my mother spent years making sure that we would be able to go to university. Although she had to argue so many times with me about studying, she rarely gave in, and I cannot thank her enough for what she did. Once I had the education I needed, my parents gave me the freedom to do whatever I wanted with it. Of course they were surprised when I told them at the age of 18 that I applied to study abroad, in England. But they always encouraged me to follow through. Throughout the last eight years, they have lifted me up when I was down, when things got difficult or when I was ill. To know that whatever happens, I can always count on their support, gave me peace of mind. It is probably no coincidence that my brother Alex and I both decided to study politics. My late grandfather Theo was a Social Democrat and the mayor of my little hometown. Moreover, my

grandfather Hans is a Christian Democrat and a union man. It is an interesting fact of life that my brother ended up voting for one party, and me for the other. Growing up in a family, in which people were passionate supporters of opposite parties, resulted in engagement with politics, not in withdrawal. One paper of my thesis shows that this pattern holds more generally.

Finally, I want to thank the person who has helped me the most over the past seven years. After I failed on my own, it was Eliza who mobilized all her networks to put me in touch with her colleague Andy, who then introduced me to Marcus. It was her who encouraged me to go to the United States, insisting for hours before I finally came around and sent an email to Don. It was her who worked both on the Birmingham and the Southampton experiments, and her and her friends who proofread some of this thesis. I even got support from Eliza's family, most notably from her sister Diana, who has helped me on many occasions, and who will make a much better Sociologist than I ever could. I just wish I had an equally important influence on Eliza's work as she has on mine. I do not know how many times she had to deal with my notorious lack of organization, or my "newest ideas". But she did all of that, and I guess that is a sign of love. This is why I dedicate my thesis to my fiancée Eliza, and to my parents Ulrike and Stefan.

Table of Contents

INTRODUCTION 1-30

REFERENCES 24-30

1. THE HEURISTIC FUNCTION OF PARTY AFFILIATION IN VOTER MOBILIZATION CAMPAIGNS 31-81

ABSTRACT 31
REFERENCES 67-74
APPENDIX 75-81

2. HOUSEHOLD PARTISAN COMPOSITION AND CAMPAIGN MOBILIZATION 82-131

ABSTRACT 82
REFERENCES 117-121
APPENDIX 122-131

3. REACHING ACROSS THE PARTISAN DIVIDE 132-169

ABSTRACT 132
REFERENCES 161-166
APPENDIX 167-169

4. IMPERSONAL BUT NOTICEABLE? 170-196

ABSTRACT 170
REFERENCES 193-195
APPENDIX 196

5. THE NATIONAL EFFECTS OF SUBNATIONAL REPRESENTATION 197-244

ABSTRACT 197
REFERENCES 227-233
APPENDIX 234-244

CONCLUSION 245-266

REFERENCES 261-266

INTRODUCTION: MOBILIZATION AND PERSUASION IN A PARTISAN CAMPAIGN ENVIRONMENT

Oxford, March 18th 2014

"Just finished! 300 done." I receive a text from Rowenna who is the Labour parliamentary candidate for Southampton Itchen, one of two constituencies located in the Southern English port city. As one of the most marginal Labour seats in the country, with the incumbent Labour MP retiring and a demographic profile that suits UKIP, Itchen is likely to become one of the key Labour – Conservative battlegrounds in the upcoming UK General Election. Rowenna, 29, with a background in journalism and grassroots politics, hand-wrote 300 letters to constituents who are unlikely to support her candidacy in the General Election. Their entry in the Labour campaign database indicates that each of them told canvassers that they intend to support the Conservative Party, the Liberal Democrats, UKIP, or the Green Party instead. "I think I can persuade Tories like my Grandma", Rowenna said when we discussed the idea a couple of weeks earlier. What is special about Rowenna's letters? It is not only that she made the effort to reach out to people who are unlikely to vote for her, inviting each person individually to meet her for tea and talk about their concerns; it is also that the letters are part of a randomized field experiment.

"Randomization", I explained in a meeting with Rowenna's campaign staff and volunteers a couple of months earlier, "works like a coin flip. Have a look at this two pounds coin. I'll flip it repeatedly. If the first throw shows heads, you knock on door one, if it shows tails, you leave door one untouched. Then we'll do the same with door two, with door three and so on. If we go about allocating canvassing by coin flip, there really can't be any other reason than the canvass for why people are more likely to vote. Normally you decide which voters to target, and because of that decision you won't be

able to tell if canvassing really caused voters to come out on Election Day. It might well be the case that people voted for all kind of other reasons, probably the same that made you contact them in the first place."

It is a lively session, and there are a lot of questions following my short presentation. Not everyone is convinced. Some people fear that not attempting contact with all potential supporters might affect candidates' chances of winning in the Council elections. The solution we find is to conduct campaigns on top of the party's usual routine: pilot projects supported by a hard core of volunteers who are ready to put in extra time, and most importantly, by the candidate. Rowenna invests around 30 hours of her time hand-writing every letter, and posting them to the addresses listed on the electoral register. Although a politician comes to experimentation from a different perspective than a researcher, in this case Rowenna and I shared the desire to test new ways of engaging with voters who are not natural supporters of the Labour party. Indeed, much of my research has shown that campaigns have trouble influencing the electoral behavior of individuals who support other parties. Most voters ignore GOTV calls from rival parties. Could Rowenna's letter make the difference? Initially, we were rather surprised and disappointed when only four constituents reply to the letter. Nevertheless, we field a survey to see if the letters affected vote intentions, even if our original plan to organize meetings failed. The survey shows that thanks to the letter, vote intentions for Rowenna increased by around 8 percentage-points compared to the control group that was not contacted. This persuasion experiment marks the end of lengthy journey that began one and a half years earlier, a few hundred miles up North in the city of Birmingham.

Birmingham, November 12th 2012

I just missed the last train back to Oxford because I got lost talking to activists after a long day instructing and supervising campaign volunteers working for the late Bob Jones, who sadly passed away only two years into his term, in the 2012 Police and Crime Commissioner Elections. I end up sleeping on the couch a local campaign organizer offers me who I owe tremendously for agreeing to participate in the first randomized experiment I conducted. As most experimentalists will note, conducting experiments in a field setting is often exhausting, but highly rewarding. For someone who studies politics, witnessing a real election campaign from the inside for the first time is an eye-opening experience.

After one week of campaigning, I have the impression that I got to know some of the volunteers well. There are people who return every day to volunteer their time, often after a long day of work. Some volunteers provide assistance for the benefit of the candidate, while others offer their services simply because the campaign manager called. We talk a lot about their experience campaigning, and the hundreds of conversations they have with voters before each election. After a few days, they are relaxed about my presence, follow my instructions without hesitation, and are happy to endure the briefings at the beginning and the end of each campaign day. What I ask of volunteers is not easy: In one treatment condition, voters are given a telephone script that is written in support of the candidate, but callers are instructed not to volunteer the candidate's party affiliation unless asked. For a seasoned campaign worker, omitting the word "Labour" is difficult. But in the end, I successfully position myself in the 'war room' with my laptop and keep a running number of contacts that volunteers make each day. Volunteers are eager to know the outputs of their day's work. And, indeed, the turnout data from the PCC election shows that volunteers are highly effective at mobilizing voters. Turnout was 15% higher[1] for

[1] 19,96% compared to 17.30% in the control group.

Labour supporters who were randomly assigned to receive a phone call from campaign workers.

Introduction

This thesis is about campaign influence and its limitations. I show that the two local campaigns that I studied mobilized and persuaded voters, but the results also suggest that campaigns are unable to influence voters at will because they operate under the severe constraints that the partisan environment imposes on them. Local campaigns often begin more than one year before an election. However, by the time a party selects her candidate to stand in a seat, large shares of the electorate will already be predisposed toward or against her simply based on which party she affiliates with (Green et al. 2002; Hillygus and Shields 2008). Although some campaigns can be transformative and shape the partisan environment in the long run, partisan identities are simply a given for most local candidates. Previous field experiments have raised questions about the consequences of party affiliation for campaign influence (Nickerson et al. 2006; Panagopoulos 2009). But for a sub-discipline that often relies on non-partisan interventions, finding answers proves difficult.

In this thesis, I report the first randomized experiments that vary partisan campaign contact in the United Kingdom. I conducted the field experiments on which this thesis is based from the inside of two separate election campaigns that took place in England in 2012 and 2014. Both campaigns were associated with Labour Party candidates, and were conducted in the context of nation-wide elections. The campaigns did not impose any conditions on the publication of the results from these experiments.

The salience of both elections differed markedly. The Labour campaign for the newly established office of Police and Crime Commissioner in the West Midlands

Constituency was the first campaign with which I worked. The election took place on 15 November 2012, and was the first of its kind in England and Wales. The fieldwork happened in the week prior to the election, during which I spent one week in the Labour campaign headquarters in Birmingham, where I supervised a volunteer phone call and text messaging experiment. The second set of experiments was part of a parliamentary election campaign in the Southern English Labour-Conservative marginal Southampton Itchen. In collaboration with the Labour Parliamentary Candidate and the campaign organizer, I randomly assigned a voter persuasion effort, one year before the General Election, and a voter mobilization campaign using door-hangers in the 2014 European and council elections.

Both campaigns shared fully anonymized versions of their local targeting database, which enabled me to select experimental samples based on partisanship, turnout history, and demographic variables. Voter turnout was measured based on public records included in the voter register, which are updated after every election. In contrast, vote intentions are, of course, not publically recorded, and are more difficult to register as a result of this. The data on vote intentions from Southampton is based on a telephone survey that volunteers conducted after the treatment was administered.

I gained access to both campaigns through snowball sampling. I initially planned to conduct the first experiment in a local council election in a West Yorkshire Town in May 2014. However, the campaign cancelled the experiment one week before the election because the election was judged 'too close' to keep the assigned control group. The second attempt proved successful thanks to the dedication of the organizer and the support of a campaign volunteer who liaised with me and who was of tremendous importance in devising the calling sheets and entering the data.

The resulting thesis contains five stand-alone papers. The first four papers are linked in that they assess how election campaigns shape and are themselves shaped by the partisan environment in which they operate. The last paper stands on its own as it takes an aggregate view of how sub-national campaigns affect party competition at the national level. I include the last paper in this collection not only because it shares the conclusion that decentralized, local campaigns matter, but because it provides the background to my intellectual journey over the last four years of graduate study. I became interested in questions of causal inference while working on this paper as a Master's student. All papers were written to be publishable on their own, and were edited only marginally for publication in this thesis. Three papers are co-authored, which is reflected in the inconsistency of the personal pronouns employed. I decided against changing the pronouns because they reflect my collaborations. Papers one and two are co-authored with Eline de Rooij (Assistant Professor of Political Science, Simon Fraser University, Vancouver), and the fifth paper is co-authored with Elias Dinas (Associate Professor of Comparative Politics, Brasenose College, University of Oxford). I am the lead author on all papers included in this thesis, with the exception of the last paper, to which Elias Dinas and I contributed equally. My contributions are clearly identified in the declaration of authorship submitted alongside this thesis, and I hope that the reader will excuse any inconsistencies, as well as the inevitable repetition of theory.

Concepts and Definitions

Before I outline the theory and methods that will be further elaborated, tested, and discussed in the papers that follow, I will first define the most important concepts.

The constituency campaign

Four of the five papers assembled in this thesis deal with one particular aspect of election campaigns, the constituency campaign. Constituencies are relevant electoral districts that send a defined number of representatives to an assembly. In the first-past-the-post system, one Member of Parliament represents each constituency. At times, I will also refer to the constituency campaign as a local election campaign. 'Local election campaign' should not be confused with a campaign for local office. Local campaigns can be fought for any type of office, ranging from ward councilors to Members of Parliament. The papers included in this thesis highlight one aspect of the constituency election campaign: Field operations or what Issenberg (2011) refers to as the 'ground game'. I follow Kleis-Nielsen (2012) in avoiding the term 'grassroots-campaign' because field campaigns contain important top-down elements such as co-ordination by candidates and professional campaign staff. Nevertheless, field campaigns contain grassroots elements: A field campaign would be impossible without party activists who volunteer their time during election campaigns. 'The field' is often contrasted with 'the air', which involves the traditional media, especially TV, and radio advertisement (Issenberg 2011). Conceptually, the local or constituency campaign stands in contrast to the centralized, national election campaign that is run by party head quarters, and involves party leaders and the national media. Of course the popularity of party leaders, the health of the economy, incumbency, the national media and TV debates are likely to have an influence on the outcome of an election. However, while some of these factors relate to events that coincide with constituency election campaigns, they are not the objects of this thesis.

Aims and Outcomes of Campaign Interventions: Mobilization and Persuasion

Mobilization and persuasion are the two forms of political influence that campaigns have at their disposal to affect the outcome of an electoral contest (Bartels 1998; Bailey et al. 2013). By winning over voters who were previously undecided or inclined to support a rival candidate, persuasion is about changing whom an individual intends to vote for. Mobilization, in contrast, refers to turnout, not vote intention. Parties and candidates mobilize voters if they want to change the make-up of the electorate. Therefore, persuasion and mobilization refer, at the same time, to the different strategies that campaigns employ, and to two distinct observable outcomes of campaign interventions. A campaign strategy can be aimed at persuasion, but result in mobilization, or the other way around. Interventions that aim to mobilize the electorate are often targeted at core supporters, or individuals who are already inclined to vote for the party. In contrast, interventions that aim to persuade are often targeted at swing voters, or supporters of other parties. However, in reality this distinction is not always clear cut, and campaign interventions often seek to persuade and mobilize (Nickerson et al. 2006, Bailey et al. 2013). Partisan campaign messages, therefore, often combine content targeted at mobilization (for instance, information about where and how to vote), and content targeted at persuasion, such as information about candidates and issues. The source of campaign mobilization or persuasion is usually the political party, the candidate, or volunteers who act on behalf of these political actors.

Different types of campaign interventions

Campaign interventions can be divided into different categories. Campaign interventions can be direct or indirect, personal or impersonal, partisan or non-partisan. Mobilization or persuasion is direct if it involves personal contact between a voter and someone who

represents the campaign (Gerber and Green 2008); it is indirect if it is mediated by another individual (Arceneaux 2007; Rolfe 2012), or by impersonal campaign materials such as leaflets or letters. It is important to recognize that even if campaigns only manage to talk to a fraction of the electorate directly, they reach a larger universe of voters through indirect types of influence. Indirect mobilization, as a consequence of campaign contact, occurs between spouses and family members within the household (Nickerson 2008, Sinclair et al. 2012), and within larger social networks (Fowler et al. 2012). Some go as far as to say that "all turnout is mobilized" (Rolfe 2012).

Campaigns use different methods to communicate with constituents. Some campaign interventions, such as door-to-door canvassing, volunteer phone calls, and leaflets belong to the standard repertoire of political parties in the UK; others, such as door hangers, and text messages have not been widely used to mobilize voters in the British electoral context. However, scholars have documented their large-scale use in the United States (Nickerson et al. 2006; Dale and Strauss 2009; Malhotra et al. 2011). Campaign interventions are often grouped into personal and impersonal campaign tactics, and based on non-partisan GOTV experiments, it is widely believed that personal contact with individuals is the most effective method of mobilizing voters (Green and Gerber 2008). However, campaign interventions differ on many other variables, such as quality (Nickerson 2007) and noticeability (Dale and Strauss 2009). Attributes of campaign interventions are often correlated. Personal methods are, for instance, more noticeable than impersonal methods of voter contact. Hence, the apparent effectiveness of personal contact might be confounded by other attributes.

Partisan and non-partisan campaign interventions

From a theoretical point of view, the most important distinction is the conceptual difference between partisan and non-partisan campaign interventions. Non-partisan interventions are often defined by their goals, the mobilization of registered voters, without privileging a particular partisan group. In reality, however, non-partisan campaigns might target specific ethnic or socio-economic subgroups, which are more likely to vote for one party than for another. Imagine, for instance, a "non-partisan" campaign that intends to increase turnout among Hispanic voters in the United States (Bedolla and Michelson 2012). Would this intervention be defined as partisan, or non-partisan? Defining interventions by their goals is therefore difficult. A more promising distinction that I rely on in this thesis is the classification of campaign interventions by their content. I define the partisan category narrowly as campaign interventions that are explicitly affiliated with a political party. This definition excludes advocacy campaigns that are affiliated with outside groups, and campaigns for independent candidates.

More than two hundred randomized campaign experiments have been published to date (for reviews, see Gerber and Green 2008; Green et al. 2013), but only a small fraction deals with campaigns that were explicitly affiliated with political parties. All but five of them have been conducted in the United States. Pons and Liegey (2013) and Pons (2013) report the results of door-to-door canvassing campaigns affiliated with the French Socialist Party in the 2010 local and the 2012 Presidential Election. Moreover, Brown et al. (2010) randomize a leafleting campaign of a Green Party candidate in the 2007 Ontario elections in Canada, and Kendall et al. (2014) work with the Italian Democratic Party in a mayoral election. None of these studies report statistically significant effects on turnout. There has been one partisan experiment conducted in collaboration with the Labour Party in the UK (Bochel and Denver 1971), but this study was not fully randomized.

Meta-analyses of non-partisan interventions, limited to the United States, have consistently shown positive average treatment effects of door-to-door and volunteer phone calls on turnout (Gerber and Green 2008; Green et al. 2013). Moreover, they show smaller, but statistically significant average treatment effects of non-partisan direct mail on electoral mobilization.

There is only one available meta-analysis of partisan direct mail, and it shows that partisan mailers are ineffective at increasing turnout (Green et al. 2013). There are fewer studies of partisan door-to-door canvassing and telephone calls, but the picture appears equally mixed. As we would expect from the non-partisan GOTV literature, some studies report a positive and statistically significant effect of door-to-door canvassing or volunteer phone calls on turnout (Nickerson et al. 2006; Arceneaux 2007; Barton et al 2014). However, the partisan canvasses, randomized by Green (2004), McNulty (2005), Michelson (2005), Panagopoulos (2009), and Bailey et al. (2013), report null effects on mobilization. Should we therefore conclude that partisan campaigns are less effective than non-partisan campaigns? This conclusion would be premature for many reasons, not limited to the small number of partisan experiments. First, partisan campaigns do not intend to mobilize the entire electorate, but only those individuals who support them. Hence, studies that randomize contact at the aggregate level (Pons 2014; Pons and Liegey 2013; Kendall et al. 2014) are unable to identify heterogeneous mobilization effects. It is possible that contacting party supporters increases turnout, while contacting supporters of rival parties decreases turnout. In the end, the average treatment effect will mask the fact that the intervention was effective with party supporters. Moreover, all partisan experiments run the risk of targeting individuals who are thought to be supportive of a particular party, but due to data limitations, run into the dangers of contacting the "wrong people" (Nickerson 2005; Nickerson et al. 2006). High quality pre-treatment data on individual partisanship or

predictive modeling is, therefore, a prerequisite for identifying individuals who should be responsive to partisan interventions.

Second, to generalize from the limited data we possess to the universe of partisan campaigns would be misleading because we only have data about campaigns that have agreed to work with researchers (and that are able to follow through with an experiment). It is no secret that researchers usually contact campaigns based on the snowball principle, and that they receive few replies. What is more, only some campaigns are able to follow through with experiments that are initially agreed on for various reasons that are often related to their organization. We need to be transparent about these limitations.

Table 1 presents a list of campaigns that initially agreed to work with me during the last three years. Are the campaigns that followed through with an experiment more or less effective than those who were unable to execute the experiment? In reality we have no idea about the size or the direction of the bias introduced by missing data. The results reported in this dissertation should thus be interpreted as context-dependent. They give us an idea about the effectiveness of specific campaign interventions in specific electoral contexts. The results are useful for the local campaigns that I worked with, and if replicated in different electoral settings, the effects of the specific campaign intervention such as telephone calls or door-hangers, might prove generalizable. What this study cannot provide is an evaluation of the effectiveness of Labour GOTV efforts in general.

But this is not my intention; neither is it my aim to engage in the controversial debate about the origins and the exact nature of partisanship. Other scholars have started to explore these important causal questions using experimental methods (Gerber at al. 2010; Green 2013). Instead, the papers that follow are concerned with how partisan environments shape individuals' responses to campaign interventions.

Table 1 Contacts with Election Campaign 2012-2013

Year	Region	Election Type	Contact	Outcome	Reasons	# Experiments
2012	West Yorkshire	Local Council	Candidate	Failure	Campaign cancelled	0
2012	West Midlands	Police & Crime Commissioner	Organizer, volunteer	Success		2
2014	South Yorkshire	General	Candidate	Failure	Failure to follow through	0
2014	West Yorkshire	General	Candidate	Failure	Failure to follow through	0
2014	Hampshire	General, European & Local Council	Candidate, organizer	Success		2

The promises and limits of campaign influence in partisan environments

Most of what we know about mobilization and persuasion in constituency campaigns in the United Kingdom is based on correlations found in observational data. The non-partisan field experiments conducted by Edward Fieldhouse, Peter John, David Cutts and their colleagues at Manchester (John and Brennan 2008; Cutts et al. 2009; Fieldhouse et al. 2012) have contributed much to our understanding of non-partisan mobilization methods. However, the non-partisan context of the randomized campaigns constitutes a loss of realism and makes it difficult to infer how campaigns influence voters and how voters influence each other when they are constrained by the realities of the partisan environment.

Constituency election campaigns do not operate in a political vacuum. They operate within a national political environment and within a party system that has been shaped over decades, and sometimes over centuries. Paraphrasing Herbert Simon (1955), the environment is not understood as the objective world in its totality, but as those characteristics of the totality that are relevant to the "life space" of the individual. The

political environment consists of the political institutions and organizations relevant to political life (Mondak 1993; Kuklinski et al. 2011).

In line with Simon's work, political scientists have found that most individuals are badly informed about politics (e.g. Page and Shapiro 1992; Popkin 1994; Delli Carpini and Keeter 1996). However, they can utilize the cues and symbols provided by the political environment to overcome their cognitive limitations (Sniderman et al. 1991; Mondak 1993; Popkin 1994; Druckman 2001). When individuals apply heuristic processing (i.e. mental shortcuts) they pay less attention to a message's content, instead focusing more on "noncontent cues", such as characteristics of the source in judging its validity (Chaiken 1980; Chaiken and Maheswaran 1994).

Party affiliation is often described as the most useful heuristic that the political environment provides to the voter during election campaigns (Sniderman et al. 1991; Rahn 1993; Snyder and Ting 2002; Arceneaux and Kolodny 2009). Serious candidates for political office in the United Kingdom affiliate with political parties. Political parties not only provide candidates with campaign volunteers and funds, they also offer a party label that relates to longstanding partisan attachments in the electorate. According to Aldrich (2006), party affiliation provides candidates with their basis of political support, their starting point in an election campaign. As Green et al. emphasize: "A large fraction of the electorate start each campaign with a strong proclivity to support one candidate or another." (2002: 222). The existence of parties and party attachments therefore generates "a latent expectation about which candidates a voter will support in a generic election" (Hillygus and Shields 2008).

At the same time, heuristic reasoning does not always result in a vote that approximates the fully informed preferences of the individual (Lau and Redlawsk 2001). Individuals are subject to biases (Tversky and Kahneman 1974), and they sometimes

dismiss information out of hand that would have been useful if processed systematically. For instance, party affiliation conditions the receptiveness of voters to campaign communication (Zaller 1992; Iyengar and Kinder 2000). Partisans are less likely to both perceive and accept campaign communication that is associated with a party that they do not support (Zaller 1992).

An important aspect of the partisan environment relevant to the individual's decision-making is, therefore, the provision of cues. In order to explore how the partisan environment shapes the response of individuals to campaign interventions, I will rely on the insight that "the state of information may as well be regarded as a characteristic of the decision-maker as a characteristic of his environment." (Simon 1955: 100). I manipulate the partisan environment by randomly assigning whether individuals are exposed to partisan signals and to partisan content. The consequences of the partisan environment for reception and acceptance of direct campaign communication are explored in papers one, two, three and four.

The first paper addresses how the provision of party cues by the electoral environment changes the behavior of constituents in response to campaign telephone calls. In collaboration with Eline de Rooij, I designed a randomized field experiment that varied whether a party's campaign phone call referred to the candidate's party affiliation during a low-information election in the UK. The findings from this study support the idea that voters use party affiliation as a cue to decide whether to accept or reject the standard GOTV content of the mobilization message, rather than as a mental shortcut that lowers the cost of deciding which candidate to vote for, or as a prompt to turn out in support of their partisan team.

The existence of a partisan environment might also condition social interaction between ordinary voters within a household, particularly if housemates disagree on which

party to support. Individuals are thought to be conflict averse, and are likely to withdraw from politics in order to avoid cognitive dissonance (Festinger 1962), and partisan cross-pressure (Hirschman 1970). Coleman (2004) interprets this type of behavior as a sign of 'social conformity'. Mutz (2002; 2006) has therefore concluded that partisan disagreement within social networks should result in lower electoral participation. This should especially be the case if the political environment provides partisan stimuli. The second paper exploits the spillover effects from the previously conducted Birmingham field experiment to test whether the messages' spillover effects on the turnout of the non-experimental household member varies conditional on the partisan composition of the household, and the partisan content of the message. The downstream experiment suggests that regardless of partisan content, spillover of campaign messages is stronger in heterogeneous partisan households than in homogeneous households.

In paper three, I test if candidates can overcome the constraints imposed upon them by the partisan environment by exploiting voters' use of cues competing with party affiliation. This randomized persuasion experiment tests whether signaling costly effort on the part of a UK parliamentary candidate persuades supporters of other parties to change their vote intentions. In co-operation with a first-time Labour Parliamentary candidate, I randomly assigned whether constituents who support other parties receive a hand-written, personalized letter from the candidate in which she invites them to a one-to-one meeting. However, only two experimental subjects take up the invitation to meet the candidate, and in a post-treatment survey, I find that Conservative supporters are entirely unresponsive to the letter. Nevertheless, there is some support for the hypothesis that, on average, the letters had a positive effect on subjects' candidate evaluations and vote intentions even if they refused to meet the candidate. But this effect is limited to supporters of smaller parties, such as the Liberal Democrats or UKIP. Moreover, the experiment suggests that

non-verbal cues might be most effective at influencing opinion in a positive direction when consistent with verbal communication.

Paper four reports the results of two mobilization experiments that use impersonal campaign methods: Text messages, and door hangers. I find that Labour door-hangers had large substantive effects on the turnout of Labour supporters, but do not affect turnout among constituents who are not preconditioned to support the party. Partisan text messages were ineffective at increasing turnout. Taken together, the results from papers one, three and four, therefore support the conclusion that individuals are relatively unresponsive to direct campaign contact from the main rival party.

The final paper leaves the context of individual constituency campaigns in the United Kingdom, and emphasizes the importance of sub-national election campaigns for national party competition. The paper asks the question: Do the outcomes of subnational elections in Germany produces electoral and organizational benefits that materialize at higher levels of government? According to scholarly wisdom, party competition at the subnational level plays a negligible role in national elections. This paper provides theory and evidence that qualifies this view. Subnational elections determine entrance into subnational assemblies, which is accompanied by organizational resources. Since in most cases the same political actors compete at all levels of government, they can make use of region-specific resources to improve their status in national party competition. Together with Elias Dinas, I test this argument by employing the discontinuities that the five percent electoral threshold in German state elections generates. We find that parties that marginally cross the threshold for state parliamentary representation perform better in subsequent national elections in that state. We also present evidence regarding the mechanism: Consistent with organizational explanations, bottom-up effects take some time to materialize, are more pronounced where parliamentary parties receive more resources, and

state party organizations recruit more members as a result of entering subnational parliament

Testing Causal Theories

Many theories in political science make causal claims. However, testing these causal hypotheses is much more difficult than most Political Scientists claim. Almost every student knows that correlation does not equal causation. However, until the turn of the millennium causal effects of campaign interventions were inferred from correlations in observational data, despite the fact that the data generation process was unknown. Observational studies on constituency campaigning in the UK usually find that voters who report that they have been contacted by a political party are more likely to vote for the party than voters who report no contact (Whiteley and Seyd 1994; Pattie and Johnston 2003; 2009; Johnston et al. 2012). In reality, it is difficult to attribute any causal meaning to this correlation between contact and vote choice. Even if we were able to 'control' for observable covariates, such as partisanship, demographics, or political engagement, and even if we had perfect knowledge about the variables that determine the strategic targeting of individuals by political campaigns (information that we usually do not have), it is likely that individuals who are contacted differ on many unobserved characteristics from those who are not contacted. Moreover, it is likely that unobserved heterogeneity explains why a campaign succeeds in contacting some individuals who were initially targeted, but not others. Individuals who go untreated might, for instance, be ill, on holiday, at work during the evenings, or have recently moved. Illness or vacations are of course also reasons for why individuals may not vote. A regression of turnout on reported (or validated) contact is, therefore, likely to be biased because a group of people who were successfully contacted is compared to a "control group" that is composed of both individuals who were not initially

targeted, and individuals who were targeted but who the campaign failed to reach. Without the use of randomization, estimates of campaign effects are always in danger of unobserved variable bias. In fact, in the field of voter mobilization, it has been shown on multiple occasions that matching estimators do not meet the benchmarks set by fully randomized experiments (Arceneaux et al. 2006; 2010).

Over the past fifteen years, the scientific study of constituency election campaigns has undergone a methodological revolution. Starting at Yale in the late 1990s, Gerber and Green (1999; 2001) established randomized field experiments as the gold standard for testing the effectiveness of various campaign methods. Gerber and Green (1999) understood that if they randomly assigned whether individuals were contacted by a campaign, they could expect that who was targeted had no systematic relationship with any observed or unobserved background characteristics, or with the potential responses of the individual to being targeted by the campaign.

As Arceneaux explains: "had the researcher merely randomly assigned individuals to different treatment groups but failed to carry out the experiment, there should be no differences between the treatment groups on the dependent variable" (2010: 206). Random assignment ensures that whether a campaign attempts to contact an individual or not is unrelated to the individual's potential response to the event that contact does or does not occur. These potential responses are usually called potential outcomes, and may be written $Y_i(d)$, where *d* refers to the treatment value, which in this thesis would indicate whether an individual is contacted by a campaign. Following Gerber and Green (2012: 24), potential outcomes are conceptualized as fixed attributes of each individual (at a given point in time) and represent the outcomes that would be observed hypothetically in the event that an individual was contacted or in the counterfactual event that the individual was not contacted by the campaign. In the case of campaign experiments that assign individuals or

larger units such as households to one intervention, and one control group, $Y_i(1)$ denotes the potential outcome if unit i was contacted, and $Y_i(0)$ refers to the potential outcome if unit i was not contacted. It is impossible to observe the individual treatment effect τ_i, defined as $\tau_i = Y_i(1) - Y_i(0)$, since we can only observe the treated (contacted) *or* the untreated potential outcome for a single unit (at the same point in time). The inability to observe the individual treatment effect is often referred to as "the fundamental problem of causal inference" (Holland 1986).

As a function of random assignment, the units that reveal their treated potential outcomes and the units that reveal their untreated potential outcomes, differ, in expectation, only on one variable: the treatment. In order to get an estimate of the average treatment effect (ATE) we can thus subtract the average outcome in the control group from the average outcome in the treatment group. The ATE equals $\mu_{Y(1)} - \mu_{Y(0)}$, where $\mu_{Y(1)}$ is the sum of the observed outcomes for the treated units, divided by the number of treated units, and $\mu_{Y(0)}$ is the sum of the observed outcomes for the untreated units, divided by the number of untreated units (Gerber and Green 2012: 25).[2]

For the difference-in-proportions or the difference-in-means estimator to be unbiased, we need to fulfill two further assumptions. First, individuals should only respond to their own treatment, and not to the treatment of other individuals. This assumption is called the *non-interference* or Stable Unit Treatment Value Assumption (SUTVA). The non-interference assumption is often violated if experimental units interact with each other and if interactions are not taken into account in the experimental design. Moreover, it is important that treatment assignment z_i does not set in motion causes of the observed outcome Y_i other than the treatment d_i. This assumption is called the *excludability*

[2] This holds under the assumption of full compliance, meaning all units assigned to treatment are treated, and all units assigned to control go untreated. If there is only partial compliance with the treatment, the Intent-to-Treat (ITT) estimator and the Complier Average Causal Effect (CACE) estimator will be unbiased.

assumption or the exclusion restriction, and is for instance violated if treatment affects a systematic response of a third actor.

Ethical Considerations

The field experiments on which this thesis is based were conducted in collaboration with local Labour Party campaigns.[3] This research is, therefore, an example of 'embedded experimentation' (Humphreys 2011). It falls in between a 'natural experiment' and a 'framed experiment' (Levitt and List 2009) in that the intervention was established both for non-research and for research reasons, and that it was varied for research reasons. Some people object to experimentation that is not based on informed consent on principle (Peyton 2012). However, the only feasible solution of informing subjects would have been to ask them at the beginning of each campaign phone call or on the top of each campaign letter if they agreed to participate in an experiment. Doing so would have defeated the point of the field experiment, namely the study of an intervention as it occurs in a natural setting. It would have endangered both the internal and the external validity of this research. As Humphreys (2011) points out, knowing that a subject is part of an experiment is likely to affect the subjects' response to the treatment. The modification of responses as a result of 'being studied' is known as the 'Hawthorne Effect' (Peyton 2012). Moreover, the requirement of fully informed consent would introduce randomization bias arising from individuals refusing to take part in a study.

Second, subjects were not exposed to treatments that they do not encounter in their everyday life in the course of a normal election campaign. This aspect is one of the most

[3] The Birmingham experiment was approved under CUREC1A/12-090 by the Social Sciences & Humanities Inter-Divisional Research Ethics Committee at the University of Oxford and the Southampton experiment was approved by the Research Ethics Committee of the Department of Sociology (DREC) at the University of Oxford under Reference Number 2014_03_03_R1 C1A.

important criteria Morton and Williams (2010) list. Canvassing belongs to the standard campaign repertoire of all political parties in the UK. Moreover, because the experiments took place in marginal parliamentary constituencies, constituents are used to receiving large amounts of campaign communication. Most definitely, there was no danger of harming anyone physically in these experiments, and minimal danger of causing psychological harm.[4] None of the treatments were designed to affect subjects negatively. Of course, not all voters enjoy being called by campaign volunteers. But canvassing is an integral part of British democracy, and if subjects do not wish to talk to campaigns, they are free to reject a call.

Third, these experiments did not affect the privacy rights of individuals. Personal data was fully anonymized before being shared with researchers, and campaigns were only provided with experimental results at the aggregate, not at the individual level.

Fourth, although the experiments were conducted in marginal parliamentary constituencies, they were not conducted in marginal elections, and did not affect the outcome of any electoral contest. While the experiments were conducted in collaboration with local Labour Party campaigns, I am neither a member of the Labour Party, nor have I ever been employed or remunerated by the Labour Party. Both campaigns agreed in advance that all results of the experiments could be published independently of findings. Moreover, I followed best practice (Loewen et al. 2010) and drafted agreements with both campaigns that outlined what was asked of everyone involved in the research project and informed collaborators about their right to withdraw at any time if they so wished. These

[4] The Research Ethics Committee of the Department of Sociology (DREC) at the University of Oxford shares this assessment in their review of the Southampton persuasion experiment: "The DREC has determined that intervention you are proposing into a standard practice for political candidates poses minimal risk to human subjects" (Ethics Approval for 'A Randomised Voter Persuasion Experiment' – Ref. # 2014_03_03_R1 C1A 14-005).

agreements were signed both by the person who was politically responsible for the campaign, and by all researchers involved in the experiment. Throughout my research partnerships with political practitioners I made every effort to follow Loewen et al.'s advice that field experiments "should not be viewed as one-shot affairs, but instead as the beginning of a relationship" (2010: 172).

Finally, I think that these campaign experiments are justifiable because the benefits of experimentation in this case outweigh the costs. There are, of course, the scientific benefits of gaining unbiased evidence on questions of voter mobilization and persuasion that have intrigued Political Scientists for decades. The use of experiments in this thesis significantly improves our ability to test theories of voter mobilization, heuristic reasoning, and social influence in politics. Moreover, there are the practical political and civic benefits of these experiments. Parties are important actors in Western democracies, and they play a crucial role in mobilizing citizens to participate in democratic elections. This paper tests relatively unobtrusive and cost-efficient ways of getting out the vote. Moreover, the interaction between politicians and voters of different parties is important to a tolerant and open society. Voters would probably be interested in receiving campaign communication that captures their interest and at least has the potential to influence their vote rather than calls or leaflets that are meaningless and ineffective. Overall, if employed responsibly, the use of experimentation can positively affect democratic engagement and the quality of democratic discourse. At the same time, Political Science becomes richer if it engages with real organizations and real voters that make up the fabric of our democracy.

References

Aldrich, John H. 1995. *Why Parties? The Origin and Transformation of Party Politics in America.* Chicago: University of Chicago Press.

Arceneaux, Kevin. 2010. "The benefits of experimental methods for the study of campaign effects." *Political Communication* 27(2) 199-215.

Arceneaux, Kevin, Alan S. Gerber, and Donald P. Green. 2006. "Comparing experimental and matching methods using a large-scale voter mobilization experiment." *Political Analysis* 14(1): 37-62.

Arceneaux, Kevin, Alan S. Gerber, and Donald P. Green. 2010. "A cautionary note on the use of matching to estimate causal effects: an empirical example comparing matching estimates to an experimental benchmark." *Sociological Methods & Research* 39(2): 256-282.

Bartels, Larry M. (1998). Where The Ducks Are: Voting Power in a Party System. In Geer, John G. (ed). *Politicians and Party Politics.* Baltimore: John Hopkins University Press: pp. 43-79.

Bailey, Michael A., Daniel J. Hopkins and Todd Rogers. 2013. "Unresponsive, Unpersuaded: The Unintended Consequences of Voter Persuasion Efforts", http://scholar.harvard.edu/files/todd_rogers/files/unresponsive.pdf

Barton, Jared, Marco Castillo and Ragan Petrie. 2013. "What Persuades Voters? A Field Experiment on Political Campaigning." *The Economic Journal*. 124: 293–326.

Bochel, John M., and David T. Denver. 1971. "Canvassing, turnout and party support: an experiment." *British Journal of Political Science* 1(3) 257-269.

Brown, Steven D., Andrea ML Perrella, and Barry J. Kay. 2010. "Revisiting Local Campaign Effects: An Experiment Involving Literature Mail Drops in the 2007 Ontario Election." *Canadian Journal of Political Science* 43(1): 49-67.

Cardy, Emily Arthur. 2005. "An Experimental Field Study of the GOTV and Persuasion Effects of Partisan Direct Mail and Phone Calls." In Donald P. Green and Alan S. Gerber (eds). The Science of Voter Mobilization. *The Annals of the American Academy of Political and Social Science*. 601: 28-40.

Chaiken, Shelly. 1980. "Heuristic Versus Systematic Information Processing and the Use of Source Versus Message Cues in Persuasion." *Journal of Personality and Social Psychology* 39(5): 752-766.

Chaiken, Shelly, and Durairaj Maheswaran. 1994. "Heuristic Processing Can Bias Systematic Processing: Effects of Source Credibility, Argument Ambiguity, and Task Importance on Attitude Judgement." *Journal of Personality and Social Psychology* 66(3): 460-473.

Conover, Pamela J. 1981. "Political Cues and the Perception of Candidates." *American Politics Research* 9(4): 427–448.

Coleman, Stephen. 2004. "The effect of social conformity on collective voting behavior." *Political Analysis* 12(1): 76-96.

Conover, Pamela J. 1981. "Political Cues and the Perception of Candidates." *American Politics Research* 9(4): 427–448.

Delli Carpini, Michael X., and Scott Keeter. 1996. *What Americans Know about Politics and Why It Matters.* New Haven: Yale University Press.

Festinger, Leon. *A Theory of Cognitive Dissonance*. Vol. 2. 1962. Stanford: Stanford University Press.

Fieldhouse, Edward, David Cutts, Paul Widdop, and Peter John. 2012. "Do impersonal mobilisation methods work?" Evidence from a nationwide Get-Out-the-Vote experiment in England. *Electoral Studies* 32(1): 113-123.

Fieldhouse, Edward, David Cutts, Peter John, and Paul Widdop. 2013. “When Context Matters: Assessing Geographical Heterogeneity of Get-Out-The-Vote Treatment Effects Using a Population Based Field Experiment”. *Political Behavior:* published online.

Gerber, Alan S., and Donald P. Green. 1999. Does Canvassing Increase Voter Turnout: A Field Experiment. *Proceedings of the National Academy of Sciences* 96(19): 10939-10942.

Gerber, Alan and Donald P. Green. 2012. *Field Experiments: Design, Analysis and Interpretation.* New York: Norton and Company.

Gerber, Alan S., Gregory A. Huber, and Ebonya Washington. 2010. "Party affiliation, partisanship, and political beliefs: A field experiment." *American Political Science Review* 104(4): 720-744.

Green, Donald P. 2004. "Results from a Partisan Phone and Canvassing Mobilization Campaign in Pennsylvania Primary Election." Unpublished Manuscript. Institution for Social and Policy Studies, Yale University.

Green, Donald P. 2003. "Breaking empirical deadlocks in the study of partisanship: An overview of experimental research strategies." *Politics and Governance* 1(1): 6-15.

Green, Donald P., and Alan S. Gerber. 2008. *Get out the vote: How to increase voter turnout*. Brookings Institution Press.

Green, Donald P., Peter M. Aronow, and Mary C. McGrath. 2013. “Field Experiments and the Study of Voter Turnout.” *Journal of Elections, Public Opinion & Parties*: published online.

Hirschman, Albert O. *Exit, voice, and loyalty: 1970. Responses to decline in firms, organizations, and states*. Cambridge: Harvard University Press, 1970.

Hillygus, Sunshine D. and Todd G. Shields. 2008. *The Persuadable Voter: Wedge Issues in Presidential Campaigns*. Princeton: Princeton University Press.

Humphreys, Macartan. 2011. Ethical Issues for Embedded Experimentation. Comparative Democratization Newsletter, October 9(3): 10-29.

Holland, Paul W. 1986. "Statistics and causal inference." *Journal of the American Statistical Association* 81: 945-960.

Issenberg, Sasha. 2012. *The Victory Lab: The Secret Science of Winning Campaigns.* New York: Crown Publishers.

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(3): 565-574.

Johnston Ron J., and Charles J. Pattie. 2003. "Do Canvassing and Campaigning Work? Evidence from the 2001 General Election in England." *British Elections & Parties Review* 13(1): 248-273.

Johnston, Ron J. and Pattie, Charles J. 2009. MP's expenditure and General Election campaigns: do incumbents benefit from contacting their constituents? *Political Studies* 57(3): 580-591.

Johnston, Ron, David Cutts, Charles Pattie, and Justin Fisher. 2012. "We've got them on the list: contacting, canvassing and voting in a British general election campaign." *Electoral Studies* 31(2): 317-329.

Kendall, Chad, Tomasso Nannicini and Francesco Trebi. 2014. How Do Voters Respond to Information? Evidence from a Randomized Campaign. Working paper.

Kleis-Nielsen, Rasmus. 2012. Ground Wars. *Personalized Communication in Political Campaigns*. Oxford: Oxford University Press.

Kuklinski, James H., Paul J. Quirk, Jennifer Jerit and Robert F. Rich. 2001: "The political environment and citizen competence." *American Journal of Political Science* 45(2): 410-424.

Lau, Richard R., and David P. Redlawsk. 2011. “Advantages and Disadvantages of Cognitive Heuristics in Political Decision Making." *American Journal of Political Science* 45(4): 951-971.

Levitt, Steven D., and John A. List. 2009. "Field experiments in economics: the past, the present, and the future." *European Economic Review* 53(1): 1-18.

Loewen, Peter John, Daniel Rubenson and Leonard Wantchekon. 2010. “Help Me Help You: Conducting Field Experiments with Political Elites.” *The Annals of the American Academy of Political and Social Science* 628(1): 165-175.

McNulty, John E. 2005. "Phone-Based GOTV - What's on the Line? Field Experiments with Varied Partisan Components, 2002-2003." In Donald P. Green and Alan S. Gerber (eds). The Science of Voter Mobilization. Special Editors: *The Annals of the American Academy of Political and Social Science* 601: 41-65.

Michelson, Melissa R. 2005. “Meeting the Challenge of Latino Voter Mobilization” In Donald P. Green and Alan S. Gerber (eds). The Science of Voter Mobilization. Special Editors: *The Annals of the American Academy of Political and Social Science* 601: 85-101.

Mondak, Jeffery J. 1993. "Source cues and policy approval: The cognitive dynamics of public support for the Reagan agenda." *American Journal of Political Science*: 186-212.

Mullainathan, Sendhil, Joshua Schwarzstein and Andrei Shleifer. 2008. “Coarse thinking and persuasion.” *The Quarterly Journal of Economics 123(2): 577-619.*

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): 85-97.

Page, Benjamin I., and Robert Y. Shapiro. 1992. *The rational public: Fifty years of trends in Americans’ policy preferences*. Chicago: University of Chicago Press.

Panagopoulos, Costas. 2009. "Partisan and Nonpartisan Message Content and Voter Mobilization: Field Experimental Evidence." *Political Research Quarterly* 62(1): 70-76.

Pattie, Charles J., and Ronald J. Johnston. 2003. "Hanging on the Telephone? Doorstep and Telephone Canvassing at the British General Election of 1997." *British Journal of Political Science* 33(2): 303-322.

Peyton, Kyle. 2012. Ethics and Politics in Field Experiments. The Experimental Political Scientist. Newsletter of the APSA Experimental Section 3(1): 20-37.

Pons, Vincent. 2014. Does Door to door canvassing affect vote shares: Evidence from a Countrywide Field Experiment in France. Working paper: http://economics.mit.edu/files/9000.

Pons, Vincent and Guillaume Liegey. 2013. Increasing the Electoral Participation of Immigrants. Experimental Evidence from France. Working paper: http://economics.mit.edu/files/8885.

Popkin, Samuel. 1994. *The reasoning voter: Communication and persuasion in presidential campaigns*. Chicago: University of Chicago.

Rahn, Wendy M. 1993. "The Role of Partisan Stereotypes in Information Processing about Political Candidates". *American Journal of Political Science* 37(2): 472-496.

Schaffner, Brian F., and Matthew J. Streb. 2002. "The Partisan Heuristic in Low-Information Elections." *The Public Opinion Quarterly* 66(4): 559–581.

Sinclair, Betsy, Margaret McConnell, and Donald P. Green. 2012. "Detecting Spillover Effects: Design and Analysis of Multi-level Experiments." *American Journal of Political Science* 56: 1055-1069.

Simon, Herbert A. 1955. "A behavioral model of rational choice." *The Quarterly Journal of Economics* 69(1) 99-118.

Simon, Herbert A. 1956. "Rational choice and the structure of the environment." *Psychological review* 63(2): 129.

Snyder, James M. Jr., and Michael M. Ting. 2002. "An Informational Rationale for Political Parties." *American Journal of Political Science* 46(1): 90-110.

Tversky, Amos, and Daniel Kahneman. 1974. "Judgment under uncertainty: Heuristics and biases." *Science* 185(4157): 1124-1131.

Whiteley, Paul and Patrick Seyd. 1994. Local Party Campaigning and Electoral Mobilization in Britain. *The Journal of Politics* 56(1): 242-252.

THE HEURISTIC FUNCTION OF PARTY AFFILIATION IN VOTER MOBILIZATION CAMPAIGNS:

Evidence from a Randomized Field Experiment

Florian Foos & Eline de Rooij[1]

Abstract

Party cues are widely believed to be of central importance to voter decision-making during election campaigns. Yet, it is unclear whether they have an independent effect on electoral mobilization. In a randomized campaign experiment we vary whether a campaign's volunteer phone calls refer to the candidate's and the opponent's party affiliation during a low-information election in the United Kingdom. Our findings support the idea that voters use party affiliation as a mental shortcut to accept or reject the standard Get-Out-The-Vote content of a campaign message, rather than as an information or identity cue that increases their likelihood of voting in support of the closest candidate, or their partisan team.

[1] Simon Fraser University, Vancouver.

Introduction

The majority of real-world election campaigns are run by candidates affiliated with political parties. During these campaigns the party affiliation of candidates is widely communicated and generally easily ascertained by voters. Nevertheless, we know little about how voters' knowledge of a candidate's party affiliation affects the way they respond to typical campaign appeals such as Get-Out-The-Vote (GOTV) messages (Nickerson et al. 2006, Panagopoulos 2009). There are good reasons to assume that the party brand influences voter decision-making in election campaigns (Sniderman et al. 1991; Rahn 1993; Lau and Redlawsk 2001; Snyder and Ting 2002; Arceneaux and Kolodny 2009). Since candidates do not operate in a political vacuum, and many individuals have pre-defined party preferences, affiliation with a political party might win the candidate a "block of potential supporters" (Aldrich 2006). At the same time party affiliation is not costless for candidates. It is likely to constrain the effectiveness of campaign appeals with subgroups other than party supporters: Individuals might discount a message's content if it is associated with a source they dislike or distrust based on previous experience or socialization (McGuire 1969; Zaller 1992; Rahn 1993; Kuklinksi and Hurley 1994; Iyengar and Simon 2000). In the worst case, a GOTV message that reveals a candidate's party affiliation might succeed at mobilizing supporters of rival parties, but without persuading them to switch their vote.

This ambiguity around the workings of party cues shows that understanding their usage in campaign appeals is important not only to gain insight into the particulars of voter decision-making, but because it can have serious consequences for candidates' campaigns. Almost every campaign mistakenly contacts supporters of rival parties as information on partisan preferences within the electorate is at best incomplete (Huckfeldt and Sprague 1992; for an illustrative case see Nickerson 2005). This problem is particularly pronounced in countries

like the United Kingdom, where citizens do not publically register their partisanship in order to participate in primaries. Moreover, it is not uncommon for candidates to target individuals who have previously voted for other parties in an attempt to persuade and mobilize them to vote (Johnston and Pattie 2003; Nickerson et al. 2006; Arceneaux and Kolodny 2009; Roberts 2013).[2] Hence, a candidate would want to know whether a campaign message might mobilize individuals who are unlikely to be persuaded by it (Bailey et al. 2013).

Studying the role of party affiliation in electoral decision-making has long been the privilege of lab and survey experiments (Conover 1981; Rahn 1993; Druckman 2001). The large literature on the use of heuristics in decision-making has demonstrated that party cues strongly affect how individuals evaluate candidates and political issues. However, these studies are limited by the fact that we do not know to what extent the impact of party cues is short-lived (Mutz 2006), and therefore whether their impact might fail to translate into a higher propensity to vote in real-world elections. Evidence based on observational studies that compare turnout levels in partisan and non-partisan local elections is mixed, but most find that turnout levels are higher in partisan contests (Alford and Lee 1968; Karnig and Walter 1983; Schaffner et al. 2001; Holbrook and Kaufmann 2012). This evidence is often used to support the claim that the party brand mobilizes voters. However, observational work faces great problems of unobserved heterogeneity: Partisan and non-partisan elections differ on many observable and unobservable variables other than party affiliation. And even if the party is driving the observed effect, the mechanism by which it is affecting turnout remains unclear. As Aldrich (2006) explains, parties do not only provide an initial reputation, but also volunteers who participate in GOTV efforts. Hence, it might be that partisan campaigns are

[2] Persuasion attempts are particularly common in constituencies where parties do not hold a core voter advantage (Johnston and Pattie 2003).

characterized by a greater number of GOTV appeals made by experienced volunteers, and that this is what drives up turnout, not the use of the party label.

There have also been some field experimental studies that have randomly assigned subjects to receive party campaign messages (Gerber 2004; McNulty 2005; Nickerson et al. 2006; Pons 2013; Bailey et al. 2013, Barton et al. 2013). Given the relatively small number of these studies it is not surprising that results have so far been inconclusive. Importantly, when the results of party campaign experiments are compared to the well-established results of non-partisan GOTV experiments (for a review see Green et al. 2013), the heterogeneity across the different studies –in campaign interventions, campaign goals, messages and electoral settings– greatly complicates the effort to disentangle mechanisms that might lead to differential effects of partisan and non-partisan campaigns on turnout. Panagopoulos (2009) has first highlighted this problem, and addressed it by randomly assigning individuals to receive similar partisan and non-partisan GOTV messages within the same field experiment. Although his study design thus improves on previous studies, both the partisan and non-partisan treatments proved ineffective compared to not receiving treatment, and compared to each other. This null finding therefore only demonstrates that commercial phone banks using partisan messages are as ineffective at raising turnout as commercial phone banks using non-partisan messages.

We build on Panagopoulos's (2009) approach of comparing the effects of interventions within the same experiment. In order to test how knowledge about a candidate's party affiliation affects voting decisions, we exploit the unique low-information environment created during a nationwide election for a newly established elected office in England: The Police and Crime Commissioner (PCC) Election, which was first held in November 2012. We conducted the randomized field experiment in the city of Birmingham in cooperation with a local Labour Party. The PCC Election provided us with the opportunity to conduct one of the

closest real-world experiments imaginable that corresponds to the ideal scenario in which voters are either assigned to know the party affiliation of a candidate or to be deprived of this information during a GOTV campaign. Unlike in elections that are either clearly partisan and where candidates' party affiliations are widely known, or clearly non-partisan and therefore hardly comparable to partisan elections, this election provides a unique opportunity: Before the elections, parties hesitated whether to field candidates under their party brand (Travis 2011), none of the candidates were incumbents, constituencies were very large and no free electoral materials were available (Garland and Terry 2012). The 2012 Birmingham campaign experiment on which this paper is based is also the first field experiment to be run in co-operation with a political party in the United Kingdom since Bochel and Denver's (1971) Dundee canvassing study.

In what follows we first develop theoretical expectations about how knowledge of party affiliation might affect voter decision-making in election campaigns. We then describe the research design and present the results. Our experiment shows that a telephone campaign that reveals the candidate's Labour Party affiliation was very effective at mobilizing voters in the PCC Election. However, the positive effect on turnout is not due to revealing the party affiliation of the candidate since Labour partisans and unattached voters who received a similar message delivered by the same volunteers in the same election, but without party cues, were no less likely to be mobilized. We do find some evidence for the idea that the use of party cues conditions the effectiveness of the GOTV element of the message, as supporters of rival parties were unresponsive to the message revealing the candidate's Labour Party affiliation. We discuss the implications of our findings in the concluding section.

The Role of the Party Heuristic in Mobilizing Turnout

It has been shown that most individuals are badly informed about politics (e.g. Page and Shapiro 1992; Popkin 1994; Delli Carpini and Keeter 1996). Moreover, individuals' cognitive abilities when processing political information and forming reliable judgments are said to be limited (Mondak 1993a). Although such constraints are likely to deter individuals from voting, social scientists have argued that individuals are able to reduce the economic and cognitive costs associated with voting through heuristic processing, utilizing easily accessible elite cues and political symbols to form an opinion on who or what to support (Sniderman et al. 1991; Mondak 1993a; Popkin 1994; Druckman 2001).

Heuristic processing implies that individuals use heuristics –mental shortcuts– to simplify opinion formation (Chaiken 1980; Chaiken and Maheswaran 1994; see also Petty and Cacioppo 1986 on "peripheral route" processing). When individuals apply heuristic –as opposed to systematic– processing they pay less attention to a message's content, instead focusing more on "noncontent cues", such as characteristics of the source, the majority's opinion, or the message's length, in judging its validity (idem). Individuals are more likely to engage in heuristic processing when either their motivation or capacity for systematic processing is low (Chaiken and Maheswaran 1994), for instance when the topic has little personal relevance (Chaiken 1980; Petty and Cacioppo 1986) –like most political issues– or when limited information on a complex issue is available (Lupia 1994). Endorsements by elites (Popkin 1994) or interest groups (Arceneaux and Kolodny 2009) are examples of source cues, which have been shown to affect voter decision-making (Mondak 1993a). Arceneaux and Kolodny suggest that to voters, information on "[w]hether a candidate is affiliated with their political party, shares their cultural background, or has the backing of trusted political groups may communicate just as much about the candidate's issue positions and job qualifications as

an in-depth study of all available information" (2009: 757).

Of the wide variety of source cues available to voters, party affiliation is widely seen as the most decisive signal that the political environment can provide (Conover 1981; Aldrich 1995, 2006; Rahn 1993; Druckman 2001; Snyder and Ting 2002). The "initial reputation" (Aldrich 2006) that affiliation with a particular party confers on a candidate, allows voters to associate candidates with readily available stereotypes about the typical representative of that party (Conover 1981; Rahn 1993; Green et al. 2002). Messages that stress the party affiliation of a candidate should therefore enable individuals to decide more easily between candidates, and consequently positively affect turnout. This impact should be particularly pronounced in a low salience and/or low information election. Accordingly, we expect that, in the context of the PCC Election, GOTV phone messages advocating individuals to vote for the Labour candidate should be more effective at mobilizing Labour partisans or unattached voters who hold favorable attitudes towards the Labour Party than GOTV messages that advocate in favor of the same candidate without mentioning his Labour Party affiliation.

A Labour GOTV phone call could also be beneficial to supporters of rival parties. Arceneaux and Kolodny (2009) show how Republicans are more likely to declare support for a Republican candidate when called by a well-known liberal interest group that endorses a Democratic candidate. The rationale is that voters will use party cues in their political decision-making regardless of whether such cues conflict with their own party affiliation. Depending on whether voters support or oppose the party that initiates contact, the cue will either persuade or dissuade them from supporting a particular candidate. When a source is seen as having different interests, message recipients are less likely to be persuaded (Lupia and McCubbins 1998) and have been shown to take on opposing opinions (Carmines and

Kuklinski 1990; Mondak 1993b; Lupia 1994). Thus, the party cue might mobilize individuals to either vote for or against the party who initiates contact.[3]

So far our discussion has focussed mostly on decision-making. We have suggested that when economic and cognitive costs of opinion formation are reduced, voters are more likely to turn out. There are two reasons though, why one could question whether simplifying the choice between candidates indeed leads to electoral mobilization. First, the use of heuristics is said to be particularly useful for those with the greatest need for cognitive efficiency (Mondak 1993a), i.e. low-information voters (Arceneaux and Kolodny 2009), and those who are casually involved in politics (Zaller 1992); however, such individuals are the same individuals whose propensity to vote is so low that they are unlikely to be mobilized by campaign messages in the first place (Arceneaux and Nickerson 2008). We try to account for this problem by working in a low information environment that deprives all voters of crucial candidate information. Moreover, source cues work best for individuals for whom the source cue conveys the most meaning (Mondak 1993a), for instance, those with a strong partisan preference. To take this into account, we distinguish those without any party preference (unattached voters) from those who indicate a preference for the Labour Party or a rival party. Second, and more importantly, the hypothesized mechanism from simplifying candidate choice to turnout has been called into question (Green and Shapiro 1994): Even if individuals are better able to decide between candidates, the costs of voting will still outweigh the expected benefits.

It might therefore not be the role of party affiliation in facilitating decision-making that makes a campaign appeal effective at mobilizing individuals. According to some, party identification functions as a social identity (Green et al. 2002; Greene 2004). As such,

[3] Of course if rival party supporters are not relying on source cues in their political decision-making, but are engaging in systematic processing of the content of the message, they might be persuaded to switch their vote.

partisanship is, at least in part, an attachment to a group, rather than simply an attitude (Greene 2004). Partisans see themselves as part of a team –their party– which is in constant competition with other political parties to win public support and be elected to government (Green et al. 2002: 219). For partisans then, the act of voting is to a large extent expressive; it is a show of support for the team (idem: 49; 220): "Party identification predicts the vote because partisans pull for their team and the social groups that it symbolizes while at the same time rooting against the other party and its allied social groups […]. Partisans derive what in rational choice parlance might be described as expressive benefits from supporting their team" (idem: 220). In addition to the enjoyment derived from supporting their team, partisans might also feel more compelled to respond to a request for support from fellow team members. Pressure to conform to the social norm on voting has been shown to be an important motivator in turnout decisions (e.g. Gerber et al. 2008) and is more effective when applied by members of the own group (Terry and Hogg 1996). Thus, rather than functioning as a mental shortcut that reduces the costs associated with voting, campaign messages that can clearly be associated with a political party could equally be viewed as rallying cries, to which the party's supporters in particular are likely to respond by turning out to vote (rival party supporters might also be mobilized to show support for their own team as a result of the rallying cries of a competitor). Either way, both the perspective that turnout is mobilized by helping individuals choose between candidates or through the activation of partisan identities would lead us to expect a higher turnout rate among Labour Party supporters, and potentially also among rival party supporters, as a result of the GOTV phone messages that mention the candidate's Labour affiliation than as a result of those that do not. The two mechanisms can be distinguished based on the behavior of unattached voters. Since unattached voters do not

identify with a political party, providing the party affiliation of a candidate should cut their information costs, but not activate their partisan identities.

A third perspective on the role of party source cues in campaign messages however, suggests that they allow the receiver to make a judgment about the source of the message and whether to accept or reject the message's content. If this judgment is not unfavorable, party cues should have little impact on turnout over and above that of a non-partisan GOTV message. GOTV messages typically provide useful information on where, when and how to vote, and if delivered in a personable way these messages can be very effective at increasing turnout.[4] It might therefore simply be the act of a campaign volunteer personally asking individuals to vote that might affect turnout (Arceneaux 2007) –as long as the volunteer is not associated with a negatively evaluated source. This perspective is reflected in Zaller's (1992) argument that individuals who receive a message from a party they oppose are unlikely to even consider the subsequent content of the message. Source cues might thus simply serve as acceptance-rejection cues that allow individuals to judge the remainder of a campaign message (Zaller 1992, Iyengar and Simon 2000). That is to say, in the presence of a party cue, potential voters might choose to ignore the mobilizing content of the message if they dislike the party associated with the messenger. Since many individuals are attached to political parties (Campbell et al. 1960; Green et al. 2002; Bartle and Bellucci 2009), individuals might base their choice of responding to the turnout appeal on whether they like the party affiliated with the volunteer who delivers the message. If this acceptance-rejection mechanism holds, we would expect personalized mobilization messages to be equally effective at increasing turnout when they are not associated with a party as when they are associated with a party. Only for

[4] A number of studies using GOTV field experiments have shown that more personal GOTV messages –such as door-to-door canvassing or calls from volunteer phone banks– are more effective than less personal messages –e.g. direct mailers of standard calls from professional phone banks– (Green and Gerber 2008).

voters who receive a message that is clearly associated with a party they oppose, should we find similar turnout rates compared to individuals who do not receive a campaign message, as they are expected to reject the message.

In sum, the provision of Labour Party cues is likely to either lower decision-making costs of candidate choice or to activate partisan identities. We can disentangle the two potential mechanisms indirectly by looking at the observational consequences of adding partisan cues to the message. In both cases, campaign messages including such cues should lead to higher turnout than messages without such cues, especially among Labour Party supporters, but potentially also among rival party supporters. However, while the party affiliation of a candidate should provide information about his policy positions to all subjects regardless of partisanship, the social identity mechanism should leave subjects who do not identify with any party unaffected. In contrast to increasing turnout, a third perspective suggests that party cues only enable individuals to decide whether to accept or reject the subsequent standard GOTV elements of the campaign message. We now turn to our empirical investigation of these mechanisms.

Research Design

The problem of identifying the effects of a candidate's party affiliation on turnout constitutes a classical problem of causal inference because candidates who belong to political parties differ on many observable and unobservable factors from unaffiliated candidates running campaigns without party support. Although field experiments are a promising means to address this issue of confounding factors in the context of real-world political campaigns, knowledge of party affiliation is normally difficult to manipulate in a field experimental context given extensive media coverage, normal campaign activity, and the distribution of free electoral materials. We

take advantage of the context of an election that had never been held before in England and Wales to conduct our randomized field experiment.[5] The experiment was fielded in the city of Birmingham.

Elections for the newly created office of Police and Crime Commissioner were first held on November 15th 2012, and can be compared to US-Sheriff elections. The PCC elections took place across 41 newly formed constituencies. After some initial hesitation of whether political parties should compete in these elections, both the Labour Party and the Conservative Party decided to field candidates under their respective party labels in all constituencies. The Conservative Party won 16 constituencies, the Labour Party won 13 constituencies, and Independent candidates were successful in 12 contests. Unlike parliamentary elections in the UK, the supplementary vote system was used in the PCC elections unless there were only two declared candidates (Electoral Commission 2013). Importantly, very little information was available to voters about the elections, the candidates and their party affiliation. As the Electoral Reform Society concludes in its report on the election: "Voters were then left in the dark about who they could vote for with a lack of centrally provided candidate information" (Garland and Terry 2012: 7). In fact, the government failed to provide funds to the Electoral Commission to distribute leaflets with information about candidates to eligible households (idem: 10).

Since the position of PCC was newly established, none of the candidates were incumbents with high name recognition. Moreover, none of the candidates in the West Midlands PCC Election –the setting of our experiment– had served in national office before, and, finally, the sheer size of the West Midlands Constituency, which spans 902 km^2 and

[5] The elections were not held in London and in the two remaining parts of the UK (Scotland and Northern Ireland), because the London Metropolitan Police is accountable to the Home Secretary and the two regions have their policing powers devolved to regional assemblies.

includes 2.7 million people, contributed to the fact that candidates were unknown to the general public. Seven candidates stood in the West Midlands PCC election, one candidate each for the Labour Party, the Conservative Party, the Liberal Democrats and the UK Independence Party. Moreover, there were three independent candidates competing for the position. In order to receive information about candidates or the election, potential voters had to actively search the Internet or request a leaflet from the Electoral Commission. The low-information context of the election therefore both substantially increased the probability that candidates were unknown to the electorate and serves as a strong test of the role of party affiliation in campaigns. If voters are unaware of candidates' positions on policy issues, party affiliation should serve as the only yardstick that individuals have to decide between candidates. Arguably, if in the context of such a low-information election awareness of a candidate's party affiliation does not reduce voting costs and increase turnout, it is unlikely to do so in an election in which information about candidates' policy positions is abundant.

In conducting the experiment we worked together with a local Labour Party. Due to the relatively small number of individuals included in the experiment compared to the number of eligible voters in the constituency, there was no concern of affecting the election outcome.[6] In fielding the experiment we only marginally deviated from a standard Labour GOTV effort.[7] The study relies on a three-armed experimental design, which guarantees that, in expectation, there are no observable or unobservable differences between individuals assigned either of the

[6] Our experimental dataset covered 1.3% of eligible voters in the West Midlands constituency (26,827 out of 2 million). In total, there were 238,384 votes cast in the West Midlands PCC Election (12%). Given that we mobilized 145 votes in our experiment, and the Labour candidate's winning margin amounted to 61,703 votes, our experiment did in no way affect the result of the election. We neither expected, nor found, *demobilization* of opposition supporters, as turnout was widely predicted to be very low. For more information on the election see: <http://www.westmidlands-pcc.gov.uk/transparency/bob-jones-archive/> and <http://www.bbc.co.uk/news/uk-england-birmingham-20343913>.

[7] The experiment was approved by the research ethics committees at the University of Oxford, and at Simon Fraser University, Vancouver.

two treatment groups or to the control group. We will now discuss the study population, experimental assignment and treatment in more detail.

Study Population

Our study population comprises those individuals living in a select number of electoral wards in Birmingham, and whose names were included in the Labour Campaign's electoral database, a total of 26,827 individuals. This database had most recently been updated after the May 3, 2012 local elections.[8] Only individuals registered to vote are included in the database. We used a fully anonymized version of the database to randomly assign individuals into treatment and control groups.[9] To fulfill the so-called assumption of non-interference (SUTVA), meaning individuals should only respond to their own treatment and not to the treatment of others surrounding them, we randomly selected one person per household to be included in the sample. After excluding additional household members our sample consisted of 15,461 individuals. For our experiment, we then selected those individuals for which the contact variable in the database indicated that they had provided a landline phone number only. This left us with a total of 13,065 experimental units.[10]

In addition to information on whether individuals share the same household, the database include information on gender ('Mr' or 'Ms'), year of birth, whether individuals were registered as postal voters, as well as individuals' latest recorded party preference and their turnout history (as available from public records or recorded during previous Labour campaigns).

[8] The database is updated after every election. Information on individuals included in the database combines publically available sources, and canvassing records.
[9] All names and contact details were removed and replaced by a unique personal identifier code by the Labour campaign.
[10] We used the remaining 2,395 units, for which the records provided a mobile number, to conduct a small text message experiment, the results of which will be reported elsewhere.

Experimental Assignment

In order to test whether our treatments had a different effect for Labour supporters than for unattached voters and supporters of rival parties, we stratified the sample by pre-recorded partisanship. We distinguish those whose latest recorded partisanship in the database was 'Labour', those whose latest recorded partisanship was a (specified) 'rival party' (i.e. Conservative, Liberal Democrat, Green, BNP, Respect, UK Independence Party, or who otherwise indicated to oppose Labour), and those 'unattached' whose latest recorded partisanship was unknown or who were identified as non-partisans or non-voters (i.e. 'won't say', 'don't know', 'non-voter' or missing).

To get a sense of whether the Labour Party measure of partisanship is a good predictor of vote choice, we instructed volunteers delivering the message with references to the candidate's party to ask at the end of the script which party an individual would vote for in a General Election. Records show that among those who answered this question (the response rate was 62%), 77% of subjects in the Labour group said that they would vote Labour in a General Election, while only 27% in the grouping of other partisans and 33% of unknown partisans made such a statement. Callers were entirely blind to a subject's previously recorded partisanship. These figures confirm that although the Labour Party's information about whether individuals support the party is incomplete, it is valid since voters who are labelled as Labour supporters are 3-4 times more likely to declare that they would vote for the Labour Party than those who are labelled as unattached or as supporting rival parties.

In the final step, individuals in each of the three partisan groups were randomly allocated into one of two treatment groups and one control group. The size of both treatment groups was equal, but limited compared to the size of the control group by the expected

capacity of party volunteers to contact individuals.[11] Individuals in the first treatment group received a baseline campaign message that did not include a party label, while individual in the second treatment group received a similar message that affiliated the candidate with a particular party, namely the Labour Party. The allocation of individuals into treatment and control groups is shown in Figure 1.

Figure 1: Blocked random assignment to treatment and control groups.



[11] Within each partisan group the ratio of allocation is approximately 1:1:3.22.

Treatment

Individuals in the two treatment groups were called by telephone by Labour Party volunteers[12] in the week leading up to the election (November 10th – 15th), and were encouraged to vote in the West Midlands PCC Election on November 15th 2012. In formulating the GOTV scripts we worked closely with the local Labour Party in order to ensure the messages paralleled normal GOTV efforts. Both messages provided practical information to subjects on the election date and their local polling station and encouraged them to vote for the candidate. However, the content of the scripts did vary across the two treatment groups. The 'party affiliation treatment' message explicitly referred multiple times to the candidate's party affiliation and gave more partisan information, while the 'baseline GOTV treatment' message did not mention the candidate's party affiliation or any other political party. Both message scripts are displayed below.[13]

Box 1: Baseline GOTV Message.

"Hello, my name is …. I am phoning to remind you to go out and vote for [candidate name] in the Police and Crime Commissioner Election on Thursday. Your local polling station is located at … during the usual opening hours from 7am to 10pm.

Have you heard of [candidate name]? [Candidate name] is a candidate for Police and Crime Commissioner and he is determined to fight the cuts in frontline policing. As [former role] [candidate's first name] has a strong record in reducing crime and protecting our Police Force. [Candidate name] has been fighting for the victims of crime for over 30 years.

Thanks a lot for taking the time to talk to me.

[12] The great majority of volunteers were local party activists. In order to reach our target contact rates, some additional volunteers from Oxford and London participated in the phone bank. We took care to ensure that local and out-of-town volunteers administered both the party and the candidate treatments.

[13] It is important to note that the policy positions mentioned in the scripts, such as the candidate's opposition to cuts in police numbers, did not necessarily provide a cue to his partisan affiliation as similar pledges have also been made by other candidates running in the West Midlands PCC Election (Secretary of State for the Home Department 2012).

Box 2: Party Affiliation GOTV Message.

"Hello, my name is …. I am phoning from your local Labour Party. I just wanted to remind you to go out and vote for Labour candidate [candidate name] in the Police and Crime Commissioner Election on Thursday. Your local polling station is located at … during the usual opening hours from 7am to 10pm.

Have you heard of the Labour candidate [candidate name]? Labour's [candidate name] is determined to fight the Tory cuts to frontline policing that will hit Birmingham hard if a Conservative is elected. The Conservatives have sacked Police Officers and closed down Police Stations. In contrast, the Labour Party put more Police Officers on the ground and will protect police numbers.

Thanks a lot for taking the time to talk to me.

Campaign volunteers were asked to fill out a form on whether contact with the targeted individual was made, any reasons for why contact had failed (i.e. answering machine, no answer, hang-up, etc.), and the number of call-backs made (see Online Appendix). After the election the local Labour Party provided us with an updated version of the anonymized database, which included turnout data for the PCC Election from the marked electoral register.

Balance Test

We test for balance on a number of important covariates of turnout using randomization inference to assess whether any existing imbalances between treatment and control groups are greater than those we would expect to occur from chance alone. We first extract the log-likelihood statistic regressing treatment assignment on the following pre-treatment covariates (using multinomial logistic regression): validated turnout in all available previous elections, electoral ward, gender and age. We then compare the resulting log-likelihood to the mean of the test-statistics we obtain after simulating the random assignment procedure 10,000 times. According to Gerber and Green, "when all possible randomizations have been simulated, the collection of test statistics represents the exact sampling distribution under the null hypothesis

that no covariates have any effect on the assigned treatment" (2012: 107). The resulting p-value of .86 indicates that any imbalances between treatment and control groups are likely due to chance.[14]

Contact Rates and One-sided Noncompliance

The contact rates displayed in Table 1 include the percentage of those individuals who have been successfully contacted by volunteers, meaning that the caller spoke to the named person. The table shows that we need to deal with one-sided noncompliance or failure-to-treat, a common issue in GOTV studies. We failed to administer the treatment to around 60% of those who were assigned to be treated because phone lines were dead, subjects had moved away, or could not be reached on a first or second attempt. Unfortunately, our manpower was not sufficient to call back more than once. This rate of noncompliance is similar to the rate recorded in other GOTV studies in the UK (John and Brannan 2008). We will return to the issue of non-compliance in the Analysis and Results section.

Analysis and Results

Table 2 shows the turnout rates for the candidate advocacy only ('baseline GOTV') and the 'party affiliation' treatment groups, and the control group for each of the three partisan blocks

[14] Before proceeding to the balance test and the analysis, we removed all individuals from the dataset who could be identified as postal voters based on a pre-treatment measure, as most postal voters would have voted before our GOTV effort and we therefore did not expect the Labour campaign to have any impact on these individuals' turnout in the PCC Election.

There are a small number of individuals in our dataset for which we could not verify whether they voted, as their names were removed from the updated electoral register. These individuals had either passed away or moved to another city. This issue of attrition is common in GOTV studies and is not usually considered a problem because the resulting 'missingness' is unlikely to be a function of treatment assignment. To test this assumption, we regressed whether turnout was reported or not on treatment assignment and compared the F-statistic of this regression to the mean of the F-statistic over 10,000 simulated randomizations. The resulting p-value of .59 confirms that attrition is likely to have occurred independently of treatment assignment.

Overall, 21% of the sample was removed. The final number of individuals included in the analysis is shown in parentheses in Figure 1.

separately, as well as averaged over all partisan groups. Turnout in the PCC Election among those in the control group who identify either with Labour or a rival party was 17%. For those whose party attachment is unknown or who the Labour Party listed as non-voters turnout was, not surprisingly, much lower at only 8%.

Table 1: Contact rates by partisanship, in percentages.

Partisanship	Treatment	N assigned	N treated	Contact rate
Labour partisanship	Party affiliation treatment	531	226	42.56
	Baseline GOTV treatment	501	227	45.31
Rival partisanship	Party affiliation treatment	601	279	46.42
	Baseline GOTV treatment	611	262	42.88
Unattached	Party affiliation treatment	930	310	33.33
	Baseline GOTV treatment	906	315	34.77
Total		**4080**	**1619**	**39.68**

Note: Numbers exclude postal voters and individuals with missing turnout data.

Analysis and Results

Turnout in both treatment groups was on average 2-3 percentage-points higher than in the control group, suggesting that both treatments were effective at increasing turnout. The only exception is turnout among rival party supporters who received the party affiliation treatment. For this group turnout was similar to that of the control group at 17%. These raw numbers do not seem to support the hypothesis that mentioning a candidate's party affiliation increases turnout, either by lowering decision-making costs or by activating partisan identities. Instead, comparing turnout rates between the two treatment groups, we do not detect much difference for Labour and unattached voters, while turnout in the party affiliation treatment group seems lower for rival party supporters than in the baseline GOTV treatment group.

Table 2: Percentage turnout for treatment and control groups, by partisanship.

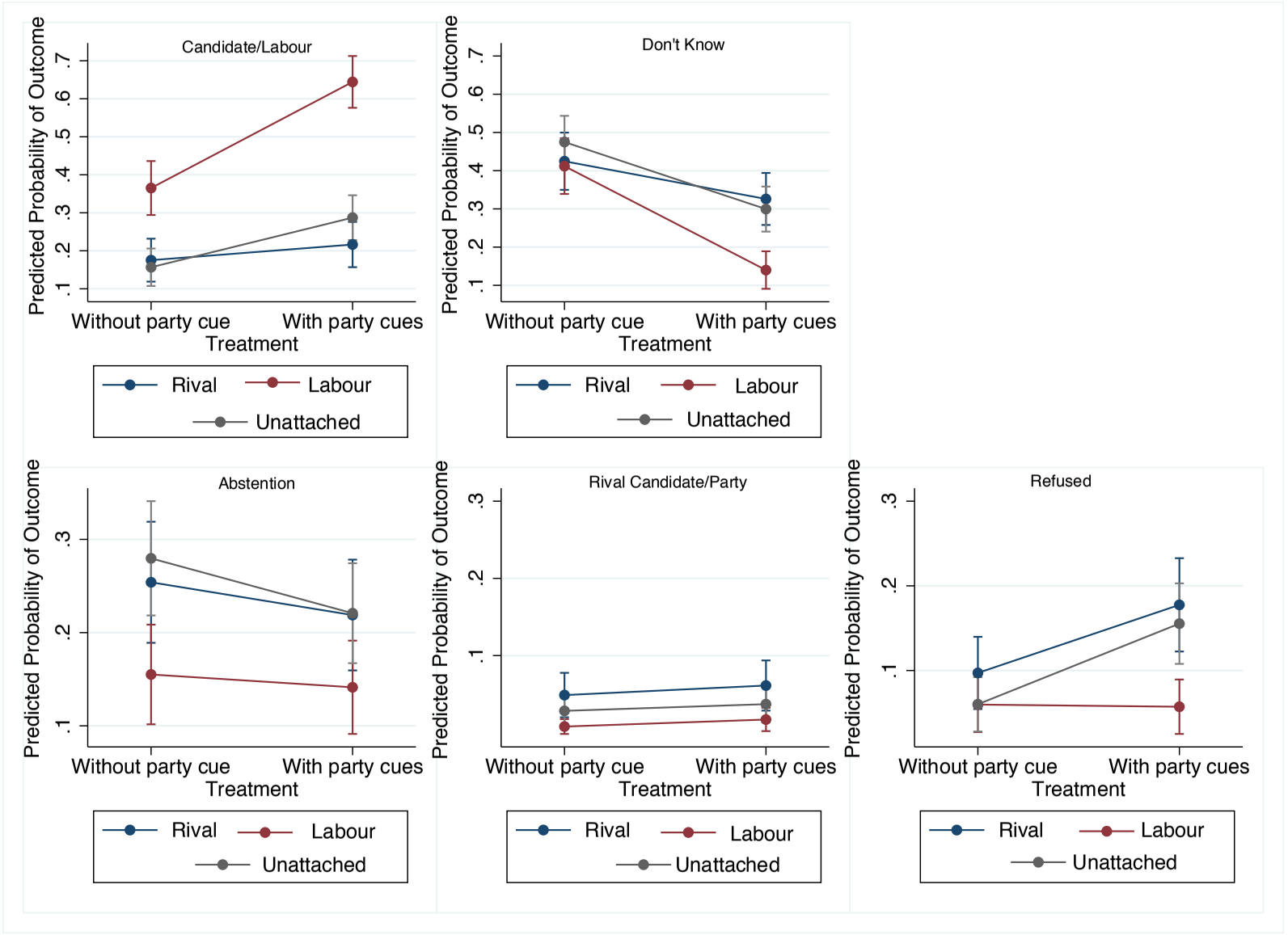
	Labour	Rival Party	Unattached	Combined
Party affiliation treatment	19.96	17.30	10.32	14.84
Baseline GOTV treatment	19.56	19.64	11.37	15.91
Control group	17.30	16.74	8.27	13.01
N Party affiliation treatment	531	601	930	2062
N Baseline GOTV treatment	501	611	906	2018
N Control	1561	1858	2876	6295

Note: Numbers exclude postal voters and individuals with missing turnout data.

Before we turn to evaluating our competing mechanisms in more detail, it is important to first establish that subjects perceived the party cues provided by the party affiliation treatment. At the end of the call subjects in both treatment groups were asked who they would support in the PCC Election. The question was open-ended, and allowed respondents to either give a candidate's name, a party name or to respond that they were undecided or planned to abstain. In Figure 2 we display the predicted probabilities from a multinomial logit regression of five responses to this question (a. voting for the Labour candidate; b. don't know; c. abstention; d. voting for a rival candidate; e. refused to answer) on the party affiliation treatment, partisan group, their interactions, and the pre-treatment covariates.

Figure 2 shows that under the baseline GOTV treatment ("without party cue") all partisan groups have the highest probability, ranging between .4 and .5, to answer "don't know" when asked about the party they would support in the PCC Election. Under the party affiliation treatment ("with party cues") the probability to answer "don't know" is substantially lower; especially for unattached voters and Labour supporters whose predicted probabilities are now only about .3 and .15, respectively.

Figure 2: Complier Average Causal Effect (CACE) of party cues on PCC vote choice intention.



Note: Bars denote 95% confidence intervals.

Instead Labour supporters, and to some extent unattached voters, have a much higher probability to give either the Labour candidate's name or the Labour Party as an answer under the party affiliation than under the baseline GOTV treatment (approximately .65 versus .35 for Labour supporters and .3 versus .15 for unattached voters).[15] This latter difference is striking given that the baseline GOTV message does provide subjects with the name of the candidate, just not his party affiliation. Most notable among rival party supporters is the slight uptick in the probability to refuse to name the candidate or party they would support when comparing the party affiliation to the baseline GOTV treatment (approximately .17 versus .1). Overall, these results suggest that the party cue is indeed picked up by potential voters, leading them, depending on their initial partisan preference, to either express support for the Labour candidate or Party in the PCC Election, or to refuse to reveal who they will support to the volunteer delivering the message.

We now return to testing the effectiveness of both the GOTV phone messages that mention the candidate's Labour affiliation and of those that do not for each partisan group separately. We use the differences-in-proportions estimator, which subtracts the turnout rate among individuals in the control group from the turnout rate among individuals in the relevant treatment group. This will provide us with an estimate of the Intent-To-Treat effect (ITT).[16]

The ITT effect of both treatment conditions for each partisan group is shown in the top two rows of Table 3. Multiplied by 100, it can be interpreted as the percentage-point difference between the average turnout in the treatment and control groups. We estimate p-values and confidence intervals using randomization inference. Following standard

[15] Due to differential attrition between the two treatment groups on this question, we display Lee trimming bounds in the Online Appendix. Both the lower and higher Lee bounds for the combined sample, and for the group of Labour supporters are significantly different from zero (Table A2).

[16] Formally, we can write this as: $ITT_Y = \frac{1}{N} \sum_{i=1}^{N} (Y_i(z = 1) - Y_i(z = 0))$, where $Y_i(z = 1)$ is the potential outcome for individual *i* under the treatment condition and $Y_i(z = 0)$ is the potential outcome for *i* under the control condition.

randomization-inference procedure (Gerber and Green 2012; Aronow and Samii 2012), we first calculate the differences-in-proportions from our experimental data set. In order to see if estimated differences are simply due to sampling variability, we then simulate 10,000 alternative randomization outcomes under the assumption that there was no treatment effect for any subject, and compare the differences-in-proportions estimate that we received from our actual randomization to the sampling distribution of all test-statistics. Since we can actually observe both potential outcomes under the "sharp-null" because the missing outcome is equal to the observed outcome, we can now calculate the probability of obtaining an estimated ITT "at least as large as the one we obtained from our actual experiment if the treatment effect were in fact zero for every subject " (Gerber and Green 2012: 62). The randomization inference-based confidence interval estimates rely on the assumption that within each block, the treatment effect is constant for every unit (see Gerber and Green 2012: 67).[17]

In order to improve the precision of our estimates of the ITT effect, we include a number of pre-treatment covariates that are likely to predict some of the unobserved variance in turnout. We include covariates for turnout in the previous seven elections, for gender, age and for the electoral ward in which an individual resides. The re-estimated ITT effects are shown in the next two rows of Table 3. Covariate-adjusted, the effect sizes of the baseline GOTV treatment range from 3 percentage-points for Labour and for rival party supporters to 4 percentage-points for subjects without reported partisan attachments (significant with p-values $< .05$, $< .05$ and $< .001$ respectively).

Turnout among all subjects that were assigned to receive the party affiliation message was also around 3 (2 without covariate-adjustment) percentage-points higher than among those in the control group. This difference is significant with a p-value $< .01$. The

[17] All randomization-based inference is performed using the statistical package RI for R (Aronow and Samii 2012).

size of the treatment effect varies by partisan group though. After adjusting for covariates, the party affiliation treatment increases turnout among Labour supporters by 4, and among subjects without reported partisan attachments by 3, percentage-points (both significant with a p-value < .01). We find no substantial or significant increase in turnout among rival party supporters.

The Intent-to-Treat (ITT) effects are conservative estimates of the impact of the GOTV campaign on the treated, as we failed to administer the treatment to around 60% of our treatment groups. To take account of this non-compliance, we calculate the Complier Average Causal Effect (CACE) for each of the three partisan groups separately. The CACE is the average treatment effect for so-called 'compliers' (those who were successfully contacted by the campaign) only.[18] We define compliers as subjects who answered the phone when called by party volunteers. To obtain an estimate of the CACE, we divide the ITT effect by the contact rate. We estimated the CACE for each of the three partisan groups separately by running a two-stage least squares regression of turnout on contact, using treatment assignment as the instrumental variable.[19] The results are shown in the lower half of Table 3.[20]

For the whole sample we find a CACE of 7 without covariate-adjustment and 8 percentage-points with covariate-adjustment as a result of the GOTV message that does not refer to the candidate's Labour Party affiliation, and a CACE of 5 and 6 percentage-

[18] The ITT effect of treatment assignment (z) on actually receiving the treatment (*d*) equals the proportion of compliers in the sample. The CACE estimator can be defined as: $\widehat{CACE} = \frac{\widehat{ITT}}{\widehat{ITT_D}}$, where $\widehat{ITT_D} = \hat{E}[d_i(1)]$.

[19] If Y_i is turnout, D_{1i} is contact under the party affiliation treatment, D_{2i} is contact under the basic candidate treatment (the two endogenous variables), Z_{1i} is treatment assignment to the party affiliation treatment and Z_{2i} is treatment assignment to the basic candidate treatment, then the model we estimate can be formally written as: $Y_i = \beta_0 + \beta_1 D_{1i} + \beta_2 D_{2i} + \mu_i$, in which $D_{1i} = \gamma_0 + \gamma_1 Z_{1i} + \varepsilon_{1i}$ and $D_{2i} = \delta_0 + \delta_1 Z_{2i} + \varepsilon_{2i}$.

[20] On the last day of telephone canvassing, volunteers were instructed to leave messages on answering machines in case subjects could not be reached during the second round of calling. Our CACE estimator assumes that those messages had negligible effects on subjects. If we assume instead that the messages left on answering machines had an effect as large as speaking to subjects in person, the size of the CACE reduces by around one fourth since the overall contact rate increases from .40 to .56 (.56 for Labour and rival supporters, and .53 for unattached voters).

points, respectively, as a result of the message that does.[21] The latter numbers in particular, however, mask substantial differences between the three partisan groups. For Labour partisans and individuals unattached to any party we find a statistically significant increase in turnout of 10 and 8 percentage-points as a result of the campaign that mentions party affiliation. For rival party supporters the increase is only around 1 percentage-point and statistically insignificant.

We can also discuss the effectiveness of the campaign in terms of the number of successful contacts necessary to obtain one additional vote. To mobilize one additional Labour supporter with a message including the candidate's party affiliation required 10 successful contacts. To mobilize one additional rival party supporter with the same message however, required as many as 77 successful contacts. To obtain the same result using the baseline GOTV message required 15 successful contacts among Labour, and 16 successful contacts among rival party supporters.[22]

Although we find no evidence that messages containing a candidate's party affiliation mobilize supporters of rival parties, as consistent with the acceptance-rejection mechanism, we might wonder whether the effect of the party affiliation treatment differs in statistical terms by partisanship. To answer this question, we model turnout as a function of contact, partisan group and the interaction between the two, using assignment to the party affiliation treatment as opposed to the control group as the instrumental variable in a two-stage least squares regression.[23] The results of the estimation of this model are shown

[21] The covariates are included in both estimation equations of the two-stage least squares regression.
[22] The equivalent numbers for unattached voters are 13 and 10 contacts per vote, and for all groups combined 16 and 12 contacts per vote. These numbers are calculated by taking the number of successful contacts in each group (reported in Table 1) and dividing it by the same number of successful contacts multiplied by the increase in turnout for the same group (reported in the bottom two rows of Table 3).
[23] If Y_i is turnout, D_i is contact (the endogenous variable), Z_i is treatment assignment to the party affiliation treatment, X is a n-by-k covariate matrix and B is a k-by-1 vector of coefficients, then the model we estimate can be formally written as:
$Y_i = \beta_0 + \beta_1 D_i + \beta_2 Labour_i + \beta_3 Unattached_i + \beta_4 (D * Labour)_i + \beta_5 (D * Unattached)_i + X * B + \mu_i$, in which

in Table 4. The reference group reflects the effect of the party affiliation treatment for rival party supporters. The interaction effect (covariate-adjusted) between partisanship and the treatment shows that Labour partisans are around 8 percentage-points more likely to turn out when treated with the party affiliation message than supporters of rival parties. The interaction effect is statistically significant with $p < .10$ using a two-tailed hypothesis test.[24] We also find the interaction effect between unattached voters and the treatment to be around 6 percentage-points.

Table 3: Intent-to-Treat (ITT) effect and Complier Average Causal Effect (CACE) by partisanship, without and with covariates.

	Labour	Rival Party	Unattached	Combined
ITT estimates				
No covariates:				
Party affiliation	.028†	.006	.021*	.018*
treatment	[-.011, .065]	[-.028, .040]	[-.000, .042]	[.001, .034]
Baseline GOTV	.023	.029†	.031**	.028***
treatment	[-.015, .061]	[-.005, .064]	[.009, .053]	[.011, .046]
Covariates:				
Party affiliation	.044**	.008	.025**	.025**
treatment	[.008, .080]	[-.028, .034]	[.006, .045]	[.009, .041]
Baseline GOTV	.033*	.028*	.037***	.034***
treatment	[-.003, .069]	[-.004, .060]	[.016, .059]	[.018, .050]
CACE estimates				
No covariates:				
Party affiliation	.063†	.012	.061*	.045**
treatment	[-.027,.153]	[.062,.087]	[.006,.128]	[.002,.088]
Baseline GOTV	.050	.068†	.089***	.071***
treatment	[.037,.137]	[-.015,.150]	[.024,.154]	[.026,.116]
Covariates:				
Party affiliation	.096*	.013	.076***	.061***
treatment	[.009,.183]	[-.056,.082]	[.016,.136]	[.021,.101]
Baseline GOTV	.066†	.062†	.104***	.084***
treatment	[.016,.149]	[-.016,.140]	[.044,.164]	[.042,.125]

† p-value < .10; * p-value < .05; ** p-value < .01; *** p-value < .001 (based on one-tailed test of sharp 0 hypothesis conducted using randomization-inference). 95%-confidence intervals in brackets. Note: Covariates are turnout in seven previous elections, gender, age and electoral ward.

$D_i = \gamma_0 + \gamma_1 Z_i + \gamma_2 Labour_i + \gamma_3 Unattached_i + \gamma_4 (Z * Labour)_i + \gamma_5 (Z * Unattached)_i + X * B + \varepsilon_i$.

Both equations include the same set of exogenous covariates.

[24] We use a two-tailed test because, formally, we are testing the null hypothesis that there is no difference – neither positive nor negative– in the treatment effects, conditional on partisan group. Using a less conservative, one-tailed test we find that the interaction effect for Labour supporters is significant with a p-value < .05, and that the interaction effect for unattached voters is significant with a p-value < .10.

Furthermore, we directly test whether messages containing the party affiliation of the candidate were significantly more effective at mobilizing potential voters than the messages avoiding party labels. The first two columns in Table 5 show the estimates from regressing turnout on a dummy variable indicating whether a subject complied with the party affiliation treatment (coded 1) or with baseline GOTV treatment (coded 0), interacted with partisan group. Since the treatment groups have been randomly assigned, we can assume that the compliers do not systematically differ between both treatment groups. Therefore, we can compare the compliers directly to each other, using a linear probability model for estimation.[25]

Table 4: Complier Average Causal Effect (CACE) for partisan affiliation treatment versus control group interacted with partisanship, without and with covariates.

	No covariates	**Covariates**
Party affiliation treatment	.012	.017
	(.030)	(.028)
Rival Party	Reference	Reference
Labour	.006	-.002
	(.014)	(.014)
Unattached	-.085***	-.002
	(.013)	(.012)
Rival Party x party treatment	Reference	Reference
Labour x party treatment	.050	.078†
	(.046)	(.043)
Unattached x party treatment	.049	.062
	(.045)	(.041)
N	8357	8357

† p-value < .10; * p-value < .05; ** p-value < .01; *** p-value < .001 (based on two-tailed t-test). Standard errors in parentheses. Note: Covariates are turnout in seven previous elections, gender, age and electoral ward.

Based on the estimation results presented for all compliers who either received the party affiliation treatment or the baseline GOTV treatment in Table 5, our best guess is that

[25] Formally, we can write the OLS model, including covariates, in which Y_i is turnout, W_i indicates that a complier received the party affiliation treatment (as opposed to the baseline GOTV treatment), X is a n-by-k covariate matrix and B is a k-by-1 vector of coefficients, as:
$Y_i = \beta_0 + \beta_1 W_i + \beta_2 Labour_i + \beta_3 Unattached_i + \beta_4 (W * Labour)_i + \beta_5 (W * Unattached)_i + X * B + \mu_i$.
Results are robust to using the logit instead of the OLS estimator (available from the authors upon request).

employing a message including the candidate's party affiliation versus a similar message without such information reduced turnout by around 4 percentage-points among subjects who support a rival party and by around 2 percentage-points among subjects unattached to a political party, while increasing turnout by around 1 percentage-point among Labour supporters. While substantively interesting, these differences are not statistically significant.

Table 5: Complier Average Causal Effect (CACE) for partisan affiliation treatment versus baseline GOTV treatment interacted with partisanship, without and with covariates.

	All subjects	
	No covariates	**Covariates**
Party affiliation treatment	-.037	-.039
	(.035)	(.033)
Rival	Reference	Reference
Labour	.008	.014
	(.037)	(.035)
Unattached	-.065†	.018
	(.034)	(.033)
Rival x party treatment	Reference	Reference
Labour x party treatment	.029	.053
	(.052)	(.049)
Unattached x party treatment	.007	.023
	(.048)	(.045)
N	1619	1619

† p-value < .10; * p-value < .05; ** p-value < .01; *** p-value < .001 (based on two-tailed t-test). Standard errors in parentheses. Note: Covariates are turnout in seven previous elections, gender, age and electoral ward.

We have shown that Labour partisans, and arguably also unattached voters, are more likely to turn out to vote as a result of the message that includes party affiliation than rival party supporters. This is not the result of that message being more effective than a message without references to party affiliation in mobilizing Labour supporters and unattached voters. Rather, it seems due to the fact that the party affiliation message fails to mobilize rival party supporters (Table 3), although the difference in the impact of the two messages fails to reach statistical significance (Table 5).

The most plausible reason for why the differential effect of adding party affiliation to the baseline GOTV message is not statistically significant for the full sample of subjects is related to the established findings that heuristics are more useful to some voters than to others (Arceneaux 2008), and that the effectiveness of GOTV messages depends on which voters are located close to the turnout threshold (Arceneaux and Kolodny 2008). We might therefore expect a non-linear effect of mentioning party cues on turnout. As previous studies have shown, individuals who never vote are unlikely to vote in a low-turnout election, even when they are subject to mobilization attempts (Arceneaux and Kolodny 2009). This expectation is confirmed in our sample since no habitual non-voter turned out in the PCC election.[26] Moreover, it is debateable whether party cues are useful for habitual voters. Apart from their higher probability of voting even in the absence of a GOTV intervention, the literature does not agree on whether heuristics are more or less useful to highly engaged citizens. Some argue that highly engaged citizens, because they are better informed about politics, should be less reliant on cues to form opinions (Arceneaux and Kolodny 2009), while others argue that only politically sophisticated citizens can use heuristics effectively (Lau and Redlawsk 2001).

We therefore introduce propensity to vote *P*, indicating how many times an individual voted in all previous elections included in our records (7), as a measure of engagement. We add two three-way interactions of turnout propensity and its squared term with the treatment indicator (*W*), and the individual's party affiliation to our previous turnout model.[27] Table A1 in the Appendix shows that the three-way interactions between the party affiliation treatment, Labour partisanship, and both vote propensity and its

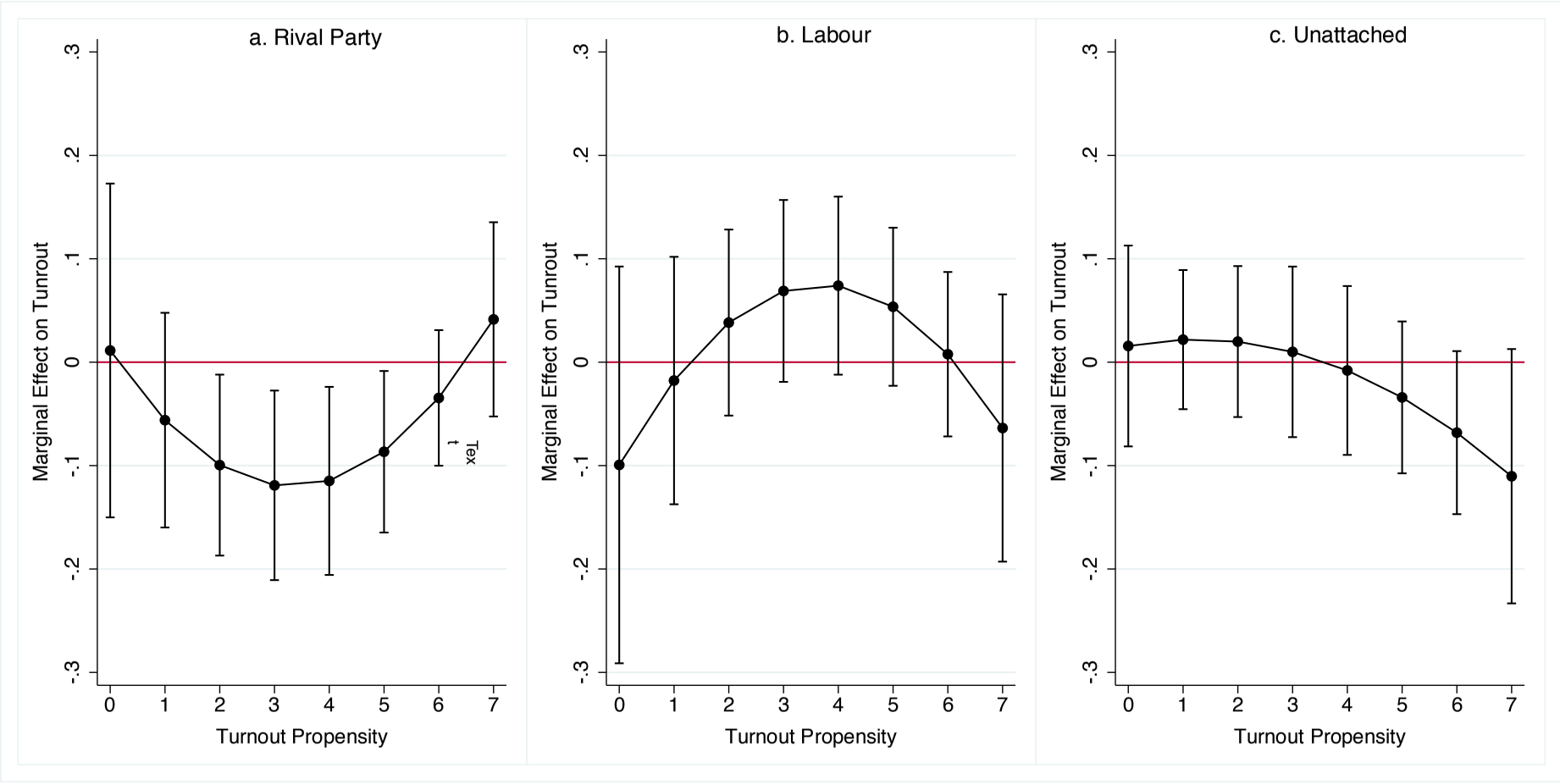
[26] Figure A1 in the Online Appendix displays the ITT effects of both GOTV treatments on turnout, conditional on turnout propensity.

[27] The OLS model can be written as: $Y_i = \beta_0 + \beta_1 W_i + \beta_2 Labour_i + \beta_3 Unattached_i + \beta_4 P + \beta_5 P^2 + \beta_6 (W * Labour)_i + \beta_7 (W * Unattached)_i + \beta_8 (P * Labour)_i + \beta_9 (P * Unattached)_i + \beta_{10} (P^2 * Labour)_i + \beta_{11} (P^2 * Unattached)_i + \beta_{12} (W * P) + \beta_{13} (W * P^2) + \beta_{14} (P * Labour * W)_i + \beta_{15} (P * Unattached * W)_i + \beta_{16} (P^2 * Labour * W)_i + \beta_{17} (P^2 * Unattached * W)_i + X * B + \mu_i$.

squared term are statistically significant with $p < .05$. For ease of interpretation, Figure 3 graphs the conditional ITT Effects and demonstrates that the effect of partisan cues on turnout is curvilinear and in opposite directions for Labour supporters and for supporters of rival parties. The negative effect of mentioning party affiliation in a GOTV message (versus not mentioning it) on turnout for rival party supporters who have shown a moderate rate of voting amounts to more than 10 percentage-points. Unattached voters are no more (or less) likely to turn out once they receive information about the party affiliation of the candidate, in contrast to what we would expect if they used party labels to reduce the costs of voter decision-making. Moreover, regardless of partisanship, habitual voters and habitual non-voters who receive the party affiliation message prove no more or less likely to vote than those who receive the non-party affiliated campaign message.

Our findings support the acceptance-rejection mechanism, which implies that while personalized mobilization messages are equally effective at increasing turnout when they are not associated with a party as when they are associated with a party, this does not apply to rival party supporters, who simply reject messages associated with the Labour Party. We only find this to be the case, however, among that majority of irregular voters. Our finding of the mobilizing impact of Labour-affiliated GOTV messages is in line with findings from previous non-partisan GOTV studies conducted in the UK (John and Brennan 2008; Cutts et al. 2009; Fieldhouse et al. 2012). In contrast, we find little support for the idea that providing potential voters with information about a candidate's party affiliation increases turnout by reducing decision-making costs or by activating partisan identities, irrespective of whether voters support or oppose the candidate's party.

Figure 3: Change in the probability of turnout resulting from mentioning the candidate's party affiliation in the GOTV message.



Note: Bars denote 90% confidence intervals.

Discussion and Conclusions

With this randomized field experiment, we set out to test competing theoretical expectations of how a candidate's party affiliation in voter mobilization campaigns influences the turnout decisions of voters. According to one perspective, party cues function as a mental shortcut for candidate positions, reducing the cost of the voting decision. Alternatively, party cues might prompt individuals to vote in order to derive the expressive or social benefits associated with supporting their partisan team. Both perspectives suggest that revealing a candidate's party affiliation would result in an additional boost in turnout above and beyond the impact of non-partisan GOTV efforts. In contrast, a third perspective suggests no such additional impact as individuals use party cues only to evaluate the message source and to decide whether to accept or reject the message's content. By providing new, field experimental, evidence on how citizens use party affiliation as a heuristic in election campaigns, our study significantly contributes to this debate.

We find that contact with a campaign that avoids mentioning a candidate's party affiliation is very effective at increasing turnout, and at least as effective as a message that explicitly mentions a candidate's affiliation. These results suggest that the positive effects on turnout associated with political parties contesting an election might be due to increased GOTV activities by party volunteers as opposed to the pure informational content of party labels. Simply having party volunteers contact citizens on the phone, and encourage them to vote might constitute the most important elements of a party's GOTV message. At the same time, we find that mentioning a candidate's party affiliation decreases the share of 'don't knows', when volunteers ask subjects who they intend to vote for. It significantly increases the likelihood of expressing support for the candidate or the party among party supporters and unattached voters. So although party cues might not mobilize voters, they

do seem to aid in short-term opinion formation. These results allow for a better understanding of the role of parties in election campaigns.

Our experimental results also suggest that mentioning a candidate's party affiliation in GOTV messages matters because it is used by individuals to judge the entire content of a GOTV message. Party affiliation appears to signal to individuals whether to approve of the source of a campaign message: If individuals approve of the source or are neutral to it, they are likely to accept the subsequent information contained in the mobilization message and behave accordingly; if they disapprove they are more likely to reject the content without considering the information contained in it. Interestingly, otherwise useful information such as polling date, polling place and candidate's policy positions, is therefore simply discounted. Our results thus largely support the theoretical expectations formulated by authors who emphasize the role of the messenger rather than the content of the message (Kuklinski and Hurley 1994; Michelson 2005) and who attribute important conditioning effects to partisanship (Zaller 1992; Iyengar and Simon 2000). Furthermore, the results of this paper qualify Arceneaux and Kolodny's (2009) important finding about the heuristic function of issue-based endorsements in GOTV campaigns, suggesting that party affiliation might play a different role due to its direct relationship to longstanding partisan predispositions.

The fact that the PCC Election, the context of our experiment, was the first election of its kind allowed for a strong test of whether or not associating a candidate with a particular political party leads to higher turnout among supporters of different parties. Arguably, if in the context of this low-information election awareness of a candidate's party affiliation does not reduce decision-making costs and increase turnout, it is unlikely to do so in an election in which information about candidates and their policy positions is abundant. At the same time the electoral context raises some questions about the

generalizability of some of our findings. Although, as with any experiment, replication would be beneficial, this might be difficult in elections in which information is more widely available, as the majority of voters would be more familiar with the party affiliation of the major candidates.

Results from our randomized field experiment suggest that GOTV telephone campaigns as commonly conducted by political parties in local constituencies in the United Kingdom are very effective in mobilizing their target population –namely those individuals who have previously been identified as likely to vote for the party and unaffiliated voters. Their effectiveness is not principally due to the use of party labels, but to other elements of a volunteer GOTV call. The CACE estimates of around 10 percentage-points for Labour partisans and 8 percentage-points for unattached voters are the more impressive considering the low turnout context of the election. Given a turnout rate in the Labour partisan control group of 17.3 percent, a 10 percentage-point turnout increase among the treated means that those Labour supporters who received a partisan message were 58 percent more likely to vote. Although these impressive results contrast with the null findings reported by some partisan GOTV telephone experiments in the US (McNulty 2005; Panagopoulos 2009) and by observational studies of political parties' local campaigns in Britain (Whiteley and Seyd 1994; Pattie and Johnston 2003), they match the large effect sizes reported in non-partisan GOTV studies in the UK. John and Brannan (2008) find a significant increase in turnout in the 2005 British General Election as a result of their non-partisan telephone campaign in Manchester. They report an ITT effect of 3.5 and a CACE of 7.3 percentage-points. A second study conducted in the UK, by Fieldhouse et al. (2012) during the 2010 British General Election finds an ITT of 3 percentage-points and a CACE of 4 percentage-points, but no increase in turnout in the European Election of 2009, as a result of a nationally-representative non-partisan GOTV telephone campaign.

Our results are therefore consistent with the mobilization effects found in two studies of non-partisan telephone mobilization during British General Elections. As the first randomized field experiment to evaluate the effectiveness of political parties' GOTV campaigns in the United Kingdom, these results are clearly important in a comparative perspective and might update our prior assumptions about the effectiveness of local GOTV campaigns run by political parties.

References

Alford, Robert R., and Eugene C. Lee. 1968. "Voting Turnout in American Cities." *The American Political Science Review* 62(3): 796-813.

Aldrich, John H. 1995. *Why Parties? The Origin and Transformation of Party Politics in America.* Chicago: University of Chicago Press.

Aldrich, John H. 2006. *Why Parties? A Second Look*. Chicago: University of Chicago Press.

Arceneaux, Kevin. 2008. "Can Partisan Cues Diminish Democratic Accountability?" Political Behavior 30(2): 139-160.

Arceneaux, Kevin, and Robin Kolodny. 2009. "Educating the Least Informed: Group Endorsements in a Grassroots Campaign." *American Journal of Political Science* 53(4): 755–770.

Arceneaux, Kevin, and David Nickerson. 2009. "Who is Mobilized to Vote? A Re-Analysis of 11 Field Experiments. *American Journal of Political Science* 53(1): 1–16.

Aronow, Peter M., and Cyrus Samii. 2012. ri: The R package for performing randomization-based inference for experiments. R package version 0.9. http://CRAN.R-project.org/package=ri.

Bailey, Michael A., Daniel J. Hopkins, and Todd Rogers. 2013. "Unresponsive and Unpersuaded: The Unintended Consequences of Voter Persuasion Efforts." Available at SSRN: http://ssrn.com/abstract=2307631.

Bartle, John, and Paolo Bellucci (eds.). 2009. *Political Parties and Partisanship: Social Identity and Individual Attitudes*. London: Routledge.

Barton, Jared, Marco Castillo, and Ragan Petrie. 2013. "What Persuades Voters? A Field Experiment on Political Campaigning." *The Economic Journal* 124: F293-F326.

Bochel, John M. and David T. Denver. 1971. Canvassing, Turnout and Party Support: An Experiment. *British Journal of Political Science* 1(3): 257-269.

Campbell, Angus, Philip E. Converse, Warren E. Miller, and Donald E. Stokes. 1960. *The American Voter*. Chicago: University of Chicago Press.

Carmines, Edward G., and James H. Kuklinski. 1990. Incentives, Opportunities, and the Logic of Public Opinion in American Political Representation. In: John A. Ferejohn and John. H. Kuklinski (Eds), *Information and Democratic Processes,* pp. 240-268. Urbana, IL: University of Illinois Press.

Chaiken, Shelly. 1980. "Heuristic Versus Systematic Information Processing and the Use of Source Versus Message Cues in Persuasion." *Journal of Personality and Social Psychology* 39(5): 752-766.

Chaiken, Shelly, and Durairaj Maheswaran. 1994. "Heuristic Processing Can Bias Systematic Processing: Effects of Source Credibility, Argument Ambiguity, and Task Importance on Attitude Judgement." *Journal of Personality and Social Psychology* 66(3): 460-473.

Conover, Pamela J. 1981. "Political Cues and the Perception of Candidates." *American Politics Research* 9(4): 427–448.

Cutts, David, Edward Fieldhouse, and Peter John. 2009. "Is Voting Habit Forming? The Longitudinal Impact of a GOTV Campaign in the UK." *Journal of Elections, Public Opinion and Parties* 19(3): 251-263.

Delli Carpini, Michael X., and Scott Keeter. 1996. *What Americans Know about Politics and Why It Matters*. New Haven: Yale University Press.

Druckman, James N. 2001. "Using Credible Advice to Overcome Framing Effects." *Journal of Law, Economics, and Organization* 17: 62–82.

Electoral Commission. 2013. *Police and Crime Commissioner elections in England and Wales. Report on the administration of the elections held on 15 November 2012.* London: The Electoral Commission.

Fieldhouse, Edward, David Cutts, Paul Widdop, and Peter John. 2012. Do impersonal mobilisation methods work? Evidence from a nationwide Get-Out-the-Vote experiment in England. *Electoral Studies* 32(1): 113-123.

Garland, Jess and Terry, Chris. 2012. How not to run an election. The Police and Crime Commissioner Elections. Report and Analysis. London: The Electoral Reform Society: www.electoralreform.org.uk/images/dynamicImages/How%20not%20to%20run%20an%20election.pdf.

Gerber, Alan S. 2004. "Does Campaign Spending Work?: Field Experiments Provide Evidence and Suggest New Theory." *American Behavioral Scientist* 47(5): 541-574.

Gerber, Alan S., and Donald P. Green. 2012. *Field Experiments. Design, Analysis and Interpretation*. New York: Norton and Company.

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): 33-48.

Gerber, Alan S., Daniel P. Kessler and Marc Meredith 2011. "The Persuasive Effects of Direct Mail: A Regression Discontinuity Based Approach," *Journal of Politics* 73(1): 140-155.

Green, Donald P., Peter M. Aronow, and Mary C. McGrath. 2013. "Field Experiments and the Study of Voter Turnout." *Journal of Elections, Public Opinion & Parties*: published online.

Green, Donald P., and Alan S. Gerber. 2008. *Get Out the Vote: How to Increase Voter Turnout*. Washington: Brookings Institution Press.

Green, Donald P., Bradley Palmquist, and Eric Schickler. 2002. *Partisan Hearts and Minds: Political Parties and the Social Identities of Voters.* New Haven: Yale University Press.

Green, Donald P., and Ian Shapiro. 2004. *Pathologies of Rational Choice Theory: A Critique of Applications in Political Science*. New Haven: Yale University Press.

Greene, Steven. 2004. "Social Identity Theory and Party Identification." *Social Science Quarterly* 85(1): 136-153.

Holbrook, Thomas M., and Karen Kaufmann. 2012. Partisan Context and Political Behavior in U.S. Cities. Paper presented at the 2012 meeting of the Elections, Public Opinion and Parties (EPOP) Conference, University of Oxford, 7-9th September 2012.

Huckfeldt, Robert and John Sprague. 1992. "Political Parties and Electoral Mobilization: Political Structures, Social Structures and the Party Canvass". *American Political Science Review* 86(1):70-86.

Iyengar, Shanto and Adam F. Simon. 2000. "New Perspectives and Evidence on Political Communication and Campaign Effects". *Annual Review of Psychology* 51: 149-169.

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(3): 565-574.

Johnston, Ronald J., and Charles J. Pattie. 2003. "Do Canvassing and Campaigning Work? Evidence from the 2001 General Election in England. *British Elections & Parties Review* 13(1): 248-273.

Karnig, Albert K., and B. Oliver Walter. "Decline in Municipal Voter Turnout: A Function of Changing Structure." *American Politics Research* 11(4): 491-505.

Kuklinksi, James H. and Norman L. Hurley. 1994. “On Hearing and Interpreting Political Messages: A Cautionary Tale of Citizen Cue-Taking”. *The Journal of Politics*, 56(3): 729-51.

Lau, Richard R., and David P. Redlawsk. “Advantages and Disadvantages of Cognitive Heuristics in Political Decision Making." *American Journal of Political Science* 45(4): 951-971.

Lupia, Arthur. 1994. “Shortcuts Versus Encyclopedias: Information and Voting Behavior in California Insurance Reform Elections.” *American Political Science Review* 88(1): 63-76.

Lupia, Arthur, and Mathew D. McCubbins. 1998. *The Democratic Dilemma: Can Citizens Learn What They Need to Know?* Cambridge: Cambridge University Press.

Michelson, Melissa R. 2005. “Meeting the Challenge of Latino Voter Mobilization”.The ANNALS of the American Academy of Political and Social Science 601(1): 85-101.

McGuire, William J. 1969. "The Nature of Attitudes and Attitude Change." In Lindzey Gardner and Elliot Aronson. *Handbook of Social Psychology,* 2nd ed., Reading: Addison-Wesley.

McNulty, John E. 2005. “Phone-Based GOTV--What’s on the Line? Field Experiments with Varied Partisan Components, 2002-2003.” *The ANNALS of the American Academy of Political and Social Science* 601(1): 41–65.

Mondak, Jeffery J. 1993a. “Public Opinion and Heuristic Processing of Source Cues.” *Political Behavior* 15(2): 167-192.

Mondak, Jeffery J. 1993b. “Source Cues and Policy Approval: The Cognitive Dynamics of Public Support for the Reagan Agenda.” *American Journal of Political Science* 37(1): 186-212.

Mutz, Diana C. 2006. *Hearing the Other Side: Deliberative versus Participatory*

Democracy. Cambridge: Cambridge University Press.

Nickerson, David W. 2005. "Partisan Mobilization Using Volunteer Phone Banks and Door Hangers." *The ANNALS of the American Academy of Political and Social Science* 601(1): 10-27.

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): 85-97.

Page, Benjamin I., and Robert Y. Shapiro. 1992. *The rational public: Fifty years of trends in Americans' policy preferences*. Chicago: University of Chicago Press.

Panagopoulos, Costas. 2009. "Partisan and Nonpartisan Message Content and Voter Mobilization: Field Experimental Evidence." *Political Research Quarterly* 62(1): 70-76.

Pattie, Charles J., and Ronald J. Johnston. 2003. "Hanging on the Telephone? Doorstep and Telephone Canvassing at the British General Election of 1997." *British Journal of Political Science* 33(2): 303-322.

Petty, Richard E., and John T. Cacioppo. 1986. *Communication and Persuasion: Central and Peripheral Routes to Attitude Change*. New York: Springer-Verlag.

Pons, Vincent. 2013. Does Door to Door Canvassing Affect Vote Shares: Evidence from a Countrywide Field Experiment in France. Working paper: http://economics.mit.edu/files/9000.

Popkin, Samuel. 1994. *The reasoning voter: Communication and persuasion in presidential campaigns*. Chicago: University of Chicago

Rahn, Wendy M. 1993. "The Role of Partisan Stereotypes in Information Processing about Political Candidates". *American Journal of Political Science* 37(2): 472-496.

Roberts, Marcus. 2013. *Labour's Next Majority. The 40% strategy*. London: Fabian Society.

Rolfe, Meredith (2012) Voter Turnout: A Social Theory of Political Participation. Political Economy of Institutions and Decisions. Cambridge: Cambridge University Press,

Schaffner, Brian, F, Matthew Streb, and Gerald Wright. 2001. "Teams Without Uniforms: The Nonpartisan Ballot in State and Local Elections." *Political Research Quarterly* 54(1): 7–30.

Schaffner, Brian F., and Matthew J. Streb. 2002. "The Partisan Heuristic in Low-Information Elections." *The Public Opinion Quarterly* 66(4): 559–581.

Secretary of State for the Home Department. 2012 *Choose my PCC. Candidates for the West Midlands Police Force Area.* Basingstoke: Bang Communications.

Sinclair, Betsy. 2012. The Social Citizen: Peer Networks and Political Behavior. Chicago: University Of Chicago Press.

Sniderman, Paul M., Richard A. Brody, and Philippe E. Tetlock. 1991. *Reasoning and choice:*

Explorations in political psychology. Cambridge: Cambridge University Press.

Snyder, James M. Jr., and Michael M. Ting. 2002. "An Informational Rationale for Political Parties." *American Journal of Political Science* 46(1): 90-110.

Terry, Deborah J., and Michael A. Hogg. 1996. "Group Norms and the Attitude-Behavior Relationship: A Role for Group Identification." *Personality and Social Psychology Bulletin* 22(8): 776-793.

Travis, Alan. 2011. Labour conference: party may boycott police commissioner elections. The Guardian Monday 26 September 2011: http://www.theguardian.com/politics/2011/sep/26/labour-boycott-police-commissioner-election.

Whiteley, Paul and Patrick Seyd. 1994. Local Party Campaigning and Electoral Mobilization in Britain. *The Journal of Politics* 56(1): 242-252.

Zaller, John. 1992. *The nature and origins of mass opinion*. Cambridge: Cambridge University Press.

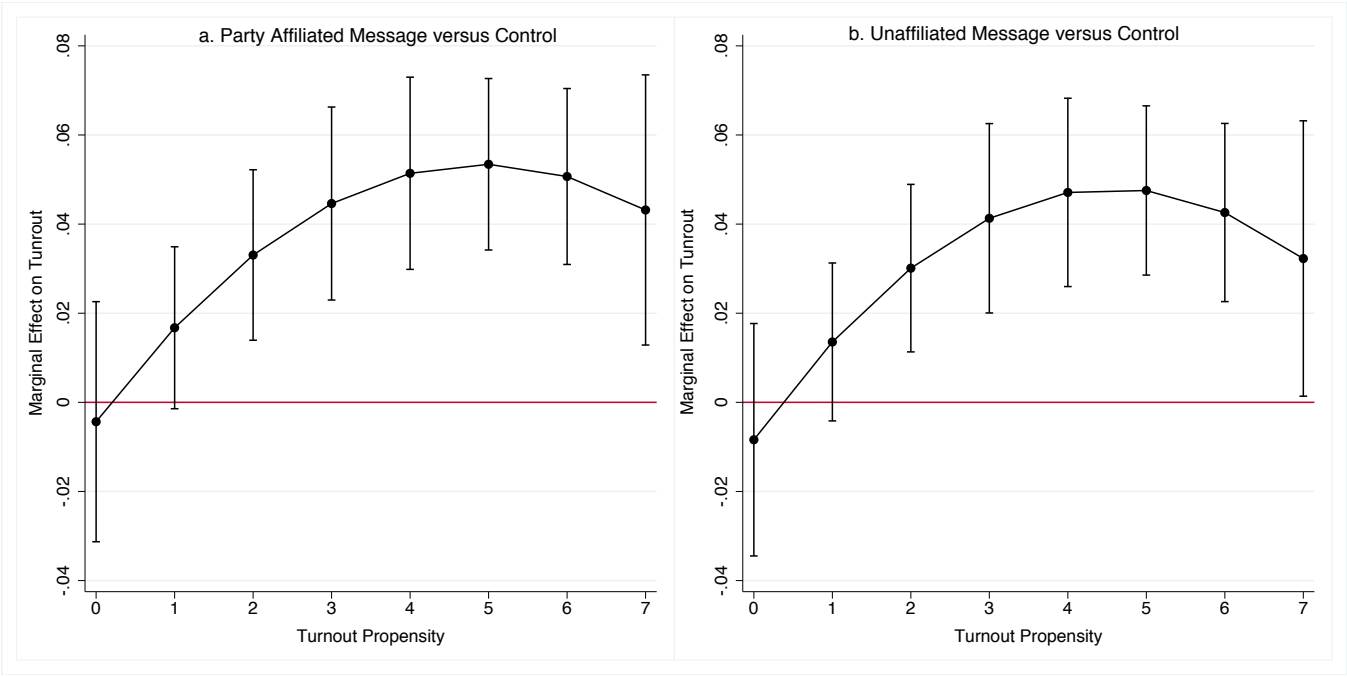
Appendix

Table A1: Complier Average Causal Effect (CACE) for partisan affiliation treatment versus baseline GOTV treatment interacted with partisanship.

Party treatment	.011
	(.098)
Rival	Reference
Labour	.057
	(.115)
Unattached	.012
	(.082)
Vote propensity	.032
	(.042)
Vote propensity squared	.003
	(.005)
Vote propensity x party treatment	-.079
	(.058)
Vote propensity squared x party treatment	.012
	(.007)
Rival x vote propensity	Reference
Labour x vote propensity	-.042
	(.065)
Unattached x vote propensity	-.043
	(.054)
Rival x vote propensity squared	Reference
Labour x vote propensity squared	.006
	(.008)
Unattached x vote propensity squared	.008
	(.007)
Rival x party treatment	Reference
Labour x party treatment	-.111
	(.152)
Unattached x party treatment	.004
	(.114)
Rival x party treatment x vote propensity	Reference
Labour x party treatment x vote propensity	.174*
	(.088)
Unattached x party treatment x vote propensity	.089
	(.074)
	Reference
Rival x party treatment x vote propensity squared	-.025*
	(.011)
Labour x party treatment x vote propensity squared	-.016
	(.010)
Unattached x party treatment x vote propensity squared	
N	1619

† p-value < .10; * p-value < .05; ** p-value < .01; *** p-value < .001 (based on two-tailed t-test). Standard errors in parentheses. Note: Covariates are gender, age and electoral ward.

Figure A1: ITT Effect of Party Affiliated and Unaffiliated Message versus Control, conditional on Turnout Propensity



Note: Bars denote 90% confidence intervals.

2012 PCC Elections Calling Script

Pool 1 Party affiliation message

When you make a call tick a box in the "Call Attempted" row. Do not leave an answer phone message unless it is the fifth call and we have not made contact yet. Do not call again if a contact has been made (i.e. the larger Question boxes have been filled in), similarly don't write anything in these boxes unless you make a contact or establish that it is a wrong number.

We have a message that we would like you to deliver. You can do it in a conversational manner but please do try and hit all the talking points in the message.

Please do make sure to mention that [candidate name] is the Labour party candidate. This is to preserve the integrity of this experiment which will greatly help us in the long run.

Message:
"Hello, my name is …. I am phoning from your local Labour Party. I just wanted to remind you to go out and vote for Labour candidate [candidate name] in the Police and Crime Commissioner Election on Thursday. Your local polling station is located at … during the usual opening hours from 7am to 10pm.
Have you heard of the Labour candidate [candidate name]?

Labour's [candidate name] is determined to fight the Tory cuts to frontline policing that will hit Birmingham hard if a Conservative is elected. The Conservatives have sacked Police Officers and closed down Police Stations. In contrast, the Labour Party put more Police Officers on the ground and will protect police numbers.

- Are you going to vote for a Police Commissioner?
- Which candidate/party are you going to support in this election?
- If there was a General Election tomorrow, which party would you support?

Thanks a lot for taking the time to talk to me.

Voice message: On the 5th attempt leave a voice message with the above content but without the ending questions.

Fill in the boxes 1 to 5 according to the criteria laid out on the next page:

Q1	Q2	Q3	Q4	Q5

2012 PCC Elections Calling Script

Pool 2 Baseline GOTV message

When you make a call tick a box in the "Call Attempted" row. Do not leave an answer phone message unless it is the fifth call and we have not made contact yet. Do not call again if a contact has been made (i.e. the larger Question boxes have been filled in), similarly don't write anything in these boxes unless you make a contact or establish that it is a wrong number.

We have a message that we would like you to deliver. You can do it in a conversational manner but please do try and hit all the talking points in the message.

Please do NOT mention that [candidate name] is the Labour party candidate unless the contact brings it up or asks you which party he represents. This is to preserve the integrity of this experiment which will greatly help us in the long run.

Message:

"Hello, my name is …. I am phoning to remind you to go out and vote for [candidate name] in the Police and Crime Commissioner Election on Thursday. Your local polling station is located at … during the usual opening hours from 7am to 10pm. Have you heard of [candidate name]?

[Candidate name] is a candidate for Police and Crime Commissioner and he is determined to fight the cuts in frontline policing. As [former role] [candidate's first name] has a strong record in reducing crime and protecting our Police Force. [Candidate name] has been fighting for the victims of crime for over 30 years.

- Are you going to vote for a Police Commissioner?
- Which candidate/party are you going to support in this election?

Thanks a lot for taking the time to talk to me.

Voice message: On the 5th attempt leave a voice message with the above content.

Fill in the boxes 1 to 5 according to the criteria laid out on the next page:

Q1	Q2	Q3	Q4	Q5

Filling in Q1-5:

Q1

tells us the status of the call so that we can analyse if and how contact was made. Use the following codes to indicate this status:

1. Conversation with the specific individual i.e. you spoke to them and they didn't ask you to "call back later"
2. Voice message left; do not leave a message unless it is the fifth attempt to contact
3. Wrong number i.e. number is for a different address/family or the specific individual has moved away
4. Number not recognised i.e. line is dead or is a fax/modem line

Q2

tells us whether the message was delivered in full; please use the following codes:

1. Full message delivered
2. Individual ends the conversation before you can deliver the full message and does not ask you to "call back later"; if you are asked to call back later leave all of the question boxes blank and we will call through the list again later
3. Individual has already voted i.e. postal voter

Q3

tells us if the individual is interested in which party [candidate name] represents

1. Individual asks you which party [candidate name] represents
2. Individual knows and mentions that [candidate name] represents Labour

Q4

tells us how the person will vote in the PCC election. Please use the following codes:

L	Labour	**B**	UKIP / [candidate name]
A	Against Labour	**I**	Independent / [candidate names]
D	Don't Know	**Z**	Not voting in PCC elections
X	Won't say	**J**	Will vote for [candidate name] specifically (rather than just the Labour candidate)
T	Conservative / [candidate name]	**O**	Will vote against [candidate name] personally (rather than just generally Against Labour [A])
S	Lib Dem / [candidate name]		

Q5

IF the individual mentions that [candidate name] is Labour or asks what party he represents, please finish by asking which party they would support if there was a General Election tomorrow and use the following codes:

L	Labour	**A**	Against i.e. not Labour
T	Conservative	**D**	Don't Know
S	Lib Dem	**X**	Won't Say
G	Green	**V**	BNP
B	UKIP	**Z**	Won't vote

Impact of the Treatment on Expressed Candidate Support

To establish that the baseline GOTV treatment and the party affiliation treatment were perceived by subjects as providing different party cues, we estimate the impact of the treatment on answering either the Labour candidate or Party to the question –posed at the end of the GOTV phone call– which party the subject would support in the PCC Election.

Attrition in the outcome variable is a more serious issue when estimating treatment effects on expressed candidate support than when examining turnout data, which is collected from public records and hence does not rely on individuals answering their phones and being willing to reply to personal questions about their candidate preferences (Bailey et al. 2013). In order to test if attrition was a function of treatment assignment, we regress a variable indicating if outcome data was missing on whether compliers received the party affiliation or the baseline GOTV treatment and extracted the resulting F-statistics. In the next step we simulate blocked random assignment to control and treatment groups 10,000 times and compare the resulting mean of all F-statistics to the F-statistics extracted from the realized data set. The p-value of .002 indicates that we must reject that attrition is unrelated to treatment assignment. To correct for this differential attrition across treatments, we follow Bailey et al. (2013) and estimate so-called Lee (trimming) bounds. Lee bounds give us a lower and an upper bound for our treatment effect by trimming off either the highest or lowest values until both treatment groups have an equal number of respondents (Gerber and Green 2012).

Table A2 shows the CACEs of the party cue on expressing support for either the Labour candidate or Party (versus any other response) by partisan group, and combined as a weighted average. We can estimate the CACE of the party affiliation versus the baseline GOTV treatment under the assumption resulting from random assignment that the shares of compliers and never-takers do not systematically differ between the two treatment groups. Our design does not allow us to estimate ITT effects because we are missing data for never-takers, those subjects who would not answer the phone when assigned to

treatment. We display both unadjusted and covariate-adjusted differences-in-proportions estimates without correcting for sample attrition ("naïve estimates"), as well as upper and lower Lee bounds.

Table A2: Complier Average Causal Effect (CACE) on Labour candidate or Party support (1 = Candidate/Labour; 0 = other) by partisanship, without and with covariates.

	Labour	Rival Party	Unattached	Combined
Naïve estimates				
No covariates	.273***	.031	.121**	.140***
	[.175, .368]	[-.043, .110]	[.050, .194]	[.092, .192]
Covariates	.273***	.036	.123**	.140***
	[.176, .369]	[-.042, .113]	[.049, .198]	[.093, .192]
N	388	394	475	1257
Lee trimming bounds:				
Lower bound	.244***	.016	.014	.086*
	[.140, .348]	[-.984, .130]	[-.095, .122]	[.017, .156]
Upper bound	.324***	.034	.162***	.172***
	[.211, .436]	[-.046, .115]	[.078, .246]	[.112, .233]

† p-value < .10; * p-value < .05; ** p-value < .01; *** p-value < .001 (based on two-tailed hypothesis tests). 95%-confidence intervals in brackets. Note: Covariates are turnout in seven previous elections, gender, age and electoral ward. Combined estimate controls for the partisan blocks.

As we would expect, mentioning the Labour Party affiliation of the candidate makes Labour supporters 27 percentage-points more likely to volunteer that they would vote for the Labour candidate. Correcting for sample attrition does not alter the size or the statistical significance of our CACE estimate, with a lower Lee bound of 24 and an upper Lee bound of 32 percentage-points. In contrast, we consistently find a null effect for supporters of rival parties. They are no more or less likely to express that they would vote for the candidate when his affiliation with the Labour Party is mentioned. Moreover, there is also some evidence that unattached voters are more likely to express the intention to vote for the candidate when they receive a message including his Labour Party affiliation, albeit the lower Lee bound of the CACE is not significantly different from zero. Across all partisan groups we find a robust positive effect of 14 percentage-points of mentioning the Labour Party affiliation of the candidate on expressing the intention to support him in the PCC Election.

HOUSEHOLD PARTISAN COMPOSITION AND CAMPAIGN MOBILIZATION:

Investigating Experimental Spillover Effects Between Partisans

Florian Foos & Eline de Rooij[1]

Abstract

Individuals who share partisan preferences with their spouse or other close members of their household are more likely to vote in elections. One explanation for this common finding is that individuals in heterogeneous partisan households will abstain from discussing electoral campaigns because they want to avoid partisan conflict. As a result, household members will be less likely to mobilize each other during election campaigns. However, existing work that tries to identify mobilization resulting from interactions within heterogeneous and homogeneous partisan networks has relied on observational data, raising questions about self-selection into networks. Mobilization effects stemming from social interaction are likely to be confounded with unobserved and time-variant individual differences and unequal probabilities of being contacted by election campaigns. We address the self-selection issue by exploiting the spillover effects from a previously conducted campaign experiment in a major UK city. During the experiment, we assigned one household member to receive a GOTV phone message with partisan content, a message without partisan content, or no message. We use data from the official electoral register to test whether the messages' spillover effects on the turnout of the non-experimental household member varies conditional on the partisan composition of the household and the partisan content of the message. Our downstream experiment suggests that regardless of partisan content, spillover of campaign messages is stronger in heterogeneous partisan households than in homogeneous households. We also find that adding partisan content to a campaign message increases spillover effects in heterogeneous partisan households, while it slightly decreases indirect mobilization within homogeneous partisan households.

[1] Simon Fraser University, Vancouver.

Introduction

Voters are increasingly viewed as social beings, whose decisions on whether or not to participate in elections and who to support are influenced by their spouses, friends and neighbors (Huckfeldt 1995; Cutts and Fieldhouse 2009; Zuckerman 2005; Zuckerman et al 2007; Rolfe 2012; Sinclair 2012). Rolfe argues, for instance, that turnout is essentially the result of mobilization by election campaigns, with much of this mobilization "occurring indirectly" through our social networks (2012: 121). The experimental literature on indirect campaign mobilization has given support to this assertion. Randomly assigned contact with campaign volunteers mobilized not only the individual who was contacted, but also individuals who lived together with the contacted person (Nickerson 2008; Sinclair et al. 2012). Still, our knowledge about spillover effects is exclusively based on non-partisan campaign interventions. Yet, most political contacts are made by partisan campaigns. This omission of the partisan element in indirect mobilization is potentially consequential because the growing literature on the partisan composition of social networks claims that indirect mobilization effects should be conditioned by the partisan composition of the social network (McClurg 2003; Mutz 2002; 2006; Belanger 2007; Nir 2011). In a series of papers, Diana Mutz (2002; 2006) emphasizes the association between heterogeneous partisan environments and the likelihood of political abstention. The basic human desire to fit in and to avoid disharmony is thought to result in increased political participation among the politically like-minded on the one hand, and political disengagement among spouses and friends who disagree in their evaluation of political parties and political issues, on the other.

This claim should hold true for within-household mobilization because the household is our most personal network and, as such, is often described as the most important context of interpersonal political influence (Zuckerman et al. 2007; Nickerson

2008; Cutts and Fieldhouse 2009; Fieldhouse and Cutts 2012; Sinclair 2012),

Although there have been a number of observational studies of how political heterogeneity within households affects political involvement and electoral participation, results have been mixed. Some studies question the causal nature of the negative association between partisan heterogeneity and participation (Scheufele et al. 2004; 2006; Huckfeldt et al. 2004; Fitzgerald and Curtis 2012; Klofstad et al. 2013).

The controversy shows no sign of abating because researchers have struggled to overcome the well known 'reflection problem', outlined at length in Manski (1993). People who are close to each other might behave in similar ways because they share the same background characteristics, and not because they exert social or political influence on each other. Researchers who have addressed this question using observational data acknowledge these problems at length (Mutz 2006: 115f; Belanger 2007: 851). Lately, a few studies have been published that exploit the benefits of panel data for causal inference (Fitzgerald and Curtis 2012; Klofstad et al. 2013; Bello and Rolfe 2013). However, while an improvement on cross-sectional data, observational panel data do not allow a researcher to overcome the selection problem related to time-variant confounders, and reverse causation. These issues are particularly prominent when studying political communication within personal networks because individuals who are more politically interested may be more likely to both talk about politics, and decide to vote over the course of an election campaign, independently from each other (Arceneaux 2010). Homogeneous partisan households might also have higher probabilities of being contacted by election campaigns than heterogeneous partisan households because they provide more campaign targets.

We are the first to use the downstream effects from a randomized field experiment, which we previously conducted in co-operation with a local Labour Party in the city of Birmingham (UK), to overcome some of these problems of causal inference (Green and

Gerber 2002). Spillover experiments were previously used to detect whether external shocks are passed on to other members of defined social networks (Nickerson 2008; Sinclair et al. 2011; Arceneaux et al. 2010; Sinclair 2012). However, so far no experiment has attempted to test whether within-household mobilization effects depend on the partisan makeup of the household. Relying on randomized experiments has the advantage that households who are contacted are in expectation the same as households who are not contacted. Differences in turnout should hence be due to social influence, and not due to compositional differences. Moreover, campaign contact is assigned with equal probabilities. Therefore, heterogeneous households will have the same probability of being contacted as homogeneous households.

In order to test spillover effects between partisans, we randomly sampled one individual per household to be part of the experiment, and then randomly assigned him or her to be contacted by telephone either with a campaign message that emphasized partisan competition, that excluded all party political content, or to a control group that did not receive any phone call. We use data from the official voter register to test whether the messages' spillover effects on the turnout of other household members are stronger in households where both partners share the same partisanship than in heterogeneous, mixed, partisan households. The ability to manipulate the partisan intensity of a campaign message in combination with information about the partisanship and prior turnout behavior of both household members provides some leverage to start opening the "black box" (Nickerson 2008) and conduct an 'implicit test' (Gerber and Green 2012) of the conflict-avoidance hypothesis.

In what follows, we first develop a theoretical framework that integrates the literature on partisan disagreement within social networks and the experimental literature on campaign spillover effects in order to understand how campaign contact with a partisan

campaign affects turnout dynamics within partisan households. We then describe the research design and present the results. Our downstream experiment suggests that regardless of partisan content, spillover of campaign messages is strong./er in heterogeneous partisan households than in homogeneous households. We also find that adding partisan content to a campaign message increases spillover effects in heterogeneous partisan households, while it slightly decreases indirect mobilization within homogeneous partisan households. We discuss the implications from our findings in the concluding section.

A journey into the 'black box': Partisan dynamics within the household

In many ways, what happens within the household during partisan election campaigns is a black box for experimentalists (Nickerson 2008; Sinclair et al. 2012; Fowler et al. 2011). As Nickerson emphasizes when discussing the results of his well-known spillover experiment: "the experiments cannot provide a mechanism for the contagion itself" (2008: 55). Theoretical accounts seem to be torn between a mechanism relying on social conformity on the one hand, and on the benefits of talk and discussion about an unfolding campaign on the other hand. Rolfe (2012: 122) spells out the communication mechanism when she writes that "localized mobilization efforts will increase discussion of an election among those reached by a candidate", and that "higher levels of political discussion […] increase the likelihood of voting. The discussion of political campaigns between household and family members has long been seen as central to how social influence shapes politics (Zuckerman 2005). For instance, Pattie and Johnston write that "[c]onversation is the forgotten context of electoral analysis" (2001: 18). Political conversation between family members is thought to increase the salience of the election, and reduce voting costs (Klofstad et al. 2013; Rolfe 2012).

However, Mutz argues that there is a dark side to discussing partisan politics: "In mixed company, there is no way to please the members of one's network and assure social harmony" (2006: 107). Partisan political discussion can of course lead to unwelcome arguments, and according to Mutz (2006), individuals will therefore try to avoid them. The mechanism at the heart of her theory is the human desire to avoid conflict, the "political angst that leads to a withdrawal from civic life", as Huckfeldt and his colleagues (2004: 4) put it. According to the conflict avoidance hypothesis, family members of opposing political views will be less likely to discuss politics with each other, and stop participating in politics. Withdrawal or exit is one way of avoiding cognitive dissonance (Festinger 1959) and partisan cross-pressure (Hirschman 1970). According to this perspective, individuals behave alike because they want to "fit in with desired friends" (Greene 1994: 92). Coleman (2004) interprets this type of behavior as a sign of 'social conformity'. Hence, according to the conflict-avoidance hypothesis, the desire to avoid unpleasant conversations about partisan politics plays a particularly important role in explaining why partisans in heterogeneous households should be less likely to mobilize each other.

Huckfeldt et al. (2004: 7) also point out that while individuals might use their household members as a low-cost source of information about the upcoming election, the source of this information must be agreeable to the individual. An opposing partisan identity can, of course, make a source politically disagreeable to the individual. The source cue effect has been widely documented in the literature on campaign communication (Zaller 1992; Iyengar and Simon 2000). Following these considerations, an individual who shares the same partisanship with someone who was contacted by a GOTV campaign should be more likely to vote than an individual who lives with someone who was contacted by the GOTV campaign and whose partisanship differs. Nevertheless, this perspective on disagreement disregards that partisanship is not the only attribute that

matters to individuals who are personally close to each other such as family members and spouses: indeed, "respondents will be more likely to retain spouses, family members and other close ties as political discussants, regardless of political agreement" (Bello and Rolfe 2013: 136). From a source cue perspective, we could therefore conclude that spouses are trustworthy sources, regardless of whether they share the same partisanship because personal proximity might out-rank party attachments when judging the credibility and trustworthiness of a source.

Moreover, political discussion often occurs as a by-product of sharing the same household and is not as easy to avoid as unwelcome TV channels or online news sites. Bello and Rolfe's (2013: 141) study of the 2010 UK General Election Campaign shows that respondent are more likely to continue discussing politics during a campaign with spouses who hold different political views when compared to spouses who hold the same view. They also find some evidence that respondents who report disagreeing with their spouse during an election campaign are more likely to change their vote intention. However, we should not assume that being persuaded by one's spouse to change one's vote in one election brings an end to partisan disagreement. Changing one's vote is not the same as changing one's party identification. On the contrary, partisans might choose to vote for the opposition party in one election and come back to their 'natural' party in the next contest. Moreover, discussing a campaign does not have to result in persuasion, or a change in vote intention. Some partisans might just as well withstand the pressure to conform to their spouse. In this case, partisan politics might serve to increase the interest of household members in the election campaign, despite the fact that household members will vote for different parties. To use an allegory, we might still discuss sports with our household members even if we support different teams. The personal relationship hypothesis therefore suggests that if individuals are very close to each other, partisan

disagreement should not lead to political withdrawal. On the contrary, within-household mobilization as a result of campaign contact might be higher in heterogeneous than in homogenous partisan households for two reasons: Either interactions might be more intense because there is more to argue about and household members might try to persuade each other (at least at the beginning of a relationship), or heterogeneous households have already resisted the pressures of persuasion or conformity for some time. While 'agreeing to disagree' might hence be a result of conflict avoidance, it might just as well be a reflection of being at ease with disagreement. This does not exclude the possibility, that overall, partisan content may have a negative effect on intra-household mobilization, but only that this effect may not be particularly pronounced amongst individuals who have either just moved in with each other and are trying to influence the other person, or already got used to disagreeing with each other on party.

Research Design

The aim of this paper is to address the question whether intra-household mobilization is conditioned by the partisan composition of the household. Similar to other experimental studies of household spillover effects (e.g. Nickerson 2008) we rely on a previously conducted randomized campaign experiment to identify mobilization effects between household members. We conducted this campaign experiment in the British city of Birmingham in the context of the West Midlands Police and Crime Commissioner (PCC) election. The PCC election was held on November 15th 2012 and was the first of its kind in England and Wales. This campaign experiment combines two design innovations that make it especially well suited for the study of mobilization-effects between partisans.

First, it was conducted in co-operation with a local Labour Party. Existing field experimental studies of interpersonal influence within households have so far exclusively

relied on identifying the spillovers of non-partisan GOTV interventions. Working from within a party's campaign will, therefore, inject more realism into a study that tries to exogenously manipulate partisan mobilization within households. Moreover, the campaign agreed to instruct canvassers to use two different messages: One campaign message tried to avoid fuelling partisan passions by refraining from mentioning the Labour candidate, or mentioning the Conservative government as being responsible for budget cuts to frontline policing. The second message, however, emphasized partisan politics and mentioned both the Labour affiliation of the candidate multiple times, as well as the candidate's opposition to the "Tory cuts" in police numbers. The message, therefore, allows for the causal identification of how direct partisan appeals affect intra-household communication between partisans.

Indeed, the second advantage is that the campaign provided us with detailed information about the pre-treatment partisanship of all household members. Our measure of individual partisanship is based on the extensive data collected by the local Labour Party during prior canvassing efforts. The party contacts every individual living in the parliamentary constituency at least once a year and asks individuals which party they support. The party identification question asked is similar to questions used in election surveys. Phone bank volunteers were blind to individuals' pre-treatment partisanship. Moreover, the database included information on gender, year of birth, whether individuals were registered as postal voters, and individuals' turnout histories (as available from public records). Since we are partially dealing with treatment-covariate interactions, this extensive set of covariates allows us to show that our conclusions are robust to covariate adjustment.

Sampling

To fulfill the non-interference assumption (SUTVA), we randomly selected one individual per household to be part of the original experiment. The household members of the experimental subjects were, thus, excluded from the original experiment. In order to achieve this, we relied on a list, derived from the Labour Party Campaign's electoral database, of 26,827 individuals who lived in a select number of electoral wards in the British city of Birmingham and who were registered to vote.[2] Next, we excluded all individuals for whom a mobile phone number was provided in the database for a separate experiment leaving a total of 13,065 primary experimental subjects with a known landline number.[3] Of these 13,065 individuals, 5,190 lived in a household that included two registered voters, according to the Labour Campaign database.[4] These secondary, non-experimental subjects in these two-voter households are the focus of this study. For random assignment, households were first grouped into three blocks based on the latest recorded partisanship of the experimental subject: 'Labour', 'rival party' supporters (i.e. Conservative, Liberal Democrat, Green, BNP, Respect, UK Independence Party, or who otherwise indicated to oppose Labour), and those who were 'unattached' to a party or whose latest recorded partisanship was unknown (i.e. 'won't say', 'don't know', 'non-voter' or missing). Next, within each group, primary subjects were randomly allocated to either one of two treatment groups or to a control group. An overview of the allocation according to the primary subject's partisanship of two-voter households into treatment and control groups in the experiment is shown in Figure 1.

[2] We used a fully anonymized version of the database in which names and contact details were removed and replaced by a unique personal identifier code by the Labour campaign.
[3] Individuals for whom the database listed a mobile phone number, either instead of or in addition to a landline number, were the subjects in a second experiment.
[4] Individuals living in households with more than two registered voters are excluded from the subsequent analyses.

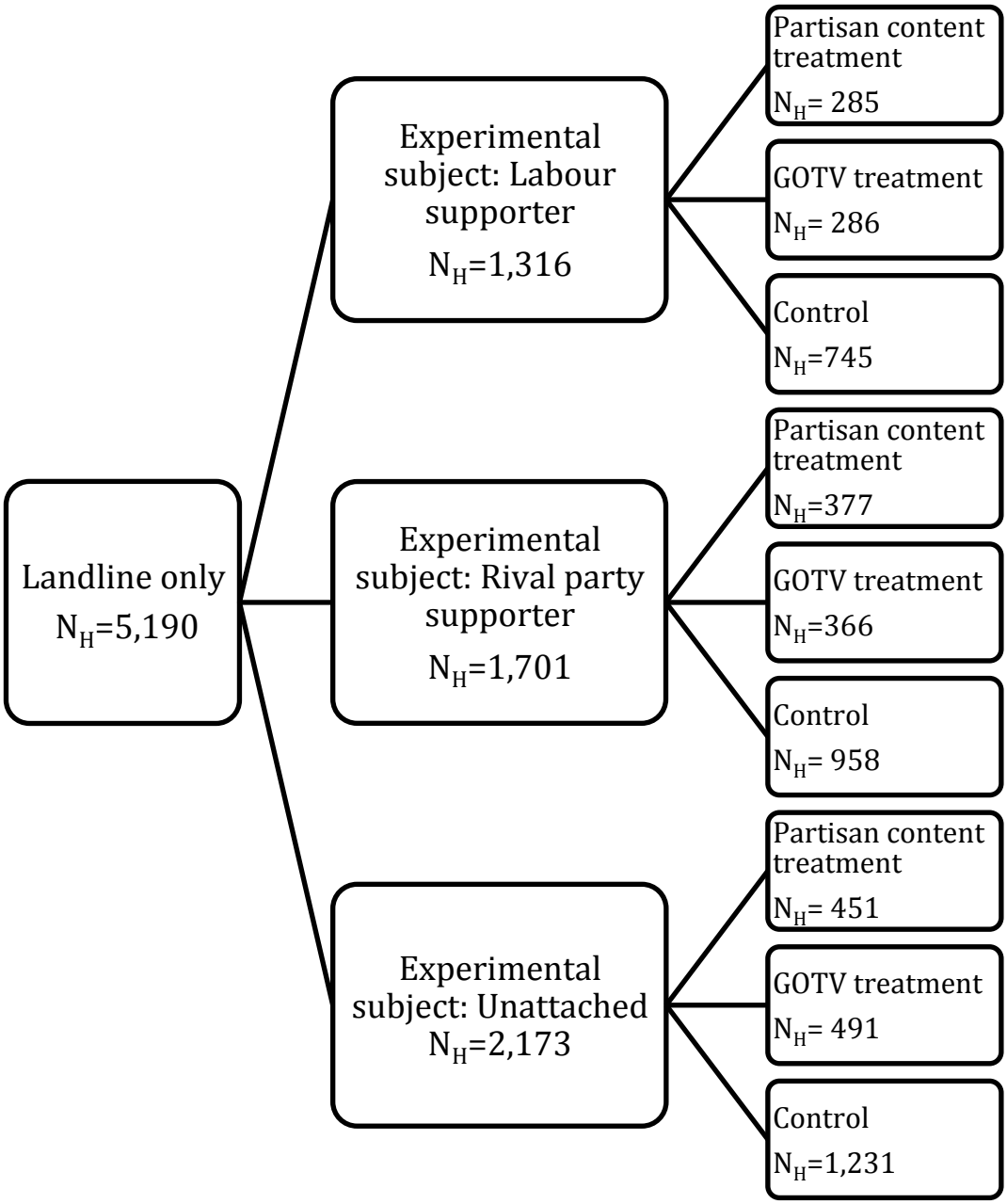
We followed standard randomization inference procedure (Gerber and Green 2012) to check if any covariate imbalance in our restricted sample of two-voter households was larger than what we would expect from random sampling variability alone.[5] Given that we did not detect any problems with covariate imbalance in the original experiment among experimental subjects, we were confident that we would reach the same conclusion for secondary experimental subjects. Indeed, the resulting p-value of .41 indicates that we cannot reject the sharp null hypothesis that the covariates taken together are not systematically related to treatment assignment of any secondary subjects.[6] A detailed balance table can be found in the Appendix (Table A.1)

In order to estimate the within household spillover effects from the treatment, we measure turnout for Labour partisans, unattached individuals, and rival party supporters, who either live with individuals who were assigned to one of the two treatment conditions, or who live with those assigned to the control condition. As Fowler and colleagues write, following this procedure, "systematic differences in network behaviors or attitudes can be attributed to diffusion of the treatment or outcome through the network" (2011: 463). In doing so, our spillover model relies on the common assumption that effects spill over within the household, but not between neighboring households. This assumption has been tested and confirmed in previous field experiments (Sinclair et al. 2012). We exploit the fact that both the partisanship of the experimental subject and the partisanship of their household member are known to test if spillover effects are conditioned by the partisan composition of the household.

[5] We use multinomial logit regression to regress treatment assignment on a n-by-k matrix of the following pre-treatment covariates: turnout in seven previous elections, gender, age and electoral ward, extract the log-likelihood, and compare it to the mean of the log-likelihood statistics we receive after simulating random assignment 10.000 times under the sharp null hypothesis that the covariates taken together do not predict treatment assignment for any subject.

[6] The p-value of randomization inference-based balance check is .31 amongst primary experimental subjects who live in two-voter households.

Figure 1: Blocked random assignment of number of two-voter households (N_H) to treatment and control groups.



The random assignment of partisan content to one of our two treatments and the availability of pre-treatment information on partisanship of both household members also allows a tentative look at possible mechanisms, which are thought to drive mobilization within partisan household. We can explore whether spillover effects of the treatment on turnout among secondary subjects are smaller when a message appeals to individuals'

partisan identities. Moreover, we can assess whether the partisanship of the primary subject, or, alternatively, the partisanship of the secondary subject, is more important in conditioning campaign spillover effects. While the former may provide some (tentative) support for the idea that it is the contacted person who, based on her partisanship, decides whether or not to pass on a campaign message to her household member, the latter may suggest that it is the person who was not contacted by the campaign who decides how to behave in response to a mobilization attempt. Following this second mechanism, potentially conflicting partisan information would not be withheld from the other household member, but household members pass on campaign messages and decide how to respond based on their individual political predispositions (Huckfeldt et al. 2004).

Treatments

Primary subjects in the two treatment groups were called by telephone by Labour Party volunteers in the week leading up to the election (November 10th – 15th), and were encouraged to vote in the West Midlands PCC election on November 15th 2012. Phone bank volunteers were instructed to ask to speak to the person in the experimental group before delivering the messages. In case the person was unavailable, volunteers were instructed to politely finish the conversation without revealing the intent of the call. In formulating the message scripts, we worked closely with the local Labour Party in order to ensure the messages paralleled normal campaign efforts. Both messages provided practical information to subjects on the election date and their local polling station, and encouraged them to vote for the same candidate, providing information about the candidate's background and his most important policy stance (opposition to cuts in police numbers). However, the content of the scripts varied drastically across the two treatment groups. The 'party content treatment' was formulated in a strongly partisan tone, and explicitly referred

to the candidate's own Labour party affiliation in an approving way, while disapproving of the main rival party, the Conservative Party. In contrast, the 'baseline GOTV treatment' message avoided all statements about party competition, and mentioned neither the candidate's party affiliation, nor the rival party. Both message scripts are displayed below.

Box 1: GOTV Message without Partisan Content.

"Hello, my name is …. I am phoning to remind you to go out and vote for [candidate name] in the Police and Crime Commissioner Election on Thursday. Your local polling station is located at … during the usual opening hours from 7am to 10pm.

Have you heard of [candidate name]? [Candidate name] is a candidate for Police and Crime Commissioner and he is determined to fight the cuts in frontline policing. As [former role] [candidate's first name] has a strong record in reducing crime and protecting our Police Force. [Candidate name] has been fighting for the victims of crime for over 30 years.

Thanks a lot for taking the time to talk to me."

Box 2: GOTV Message with Partisan Content.

"Hello, my name is …. I am phoning from your local Labour Party. I just wanted to remind you to go out and vote for Labour candidate [candidate name] in the Police and Crime Commissioner Election on Thursday. Your local polling station is located at … during the usual opening hours from 7am to 10pm.

Have you heard of the Labour candidate [candidate name]? Labour's [candidate name] is determined to fight the Tory cuts to frontline policing that will hit Birmingham hard if a Conservative is elected. The Conservatives have sacked Police Officers and closed down Police Stations. In contrast, the Labour Party put more Police Officers on the ground and will protect police numbers.

Thanks a lot for taking the time to talk to me."

Campaign volunteers were asked to fill in a form on whether contact with the targeted individual was made, any reasons for why contact had failed (i.e. answering machine, no answer, hang-up, etc.), and the number of call-backs made. After the election the local Labour Party provided us with an updated version of the (anonymized) database which included turnout data for the PCC election from the marked electoral register for the primary subjects in the experiment, as well as for the household members (secondary subjects) who were also included in the database. Before the analysis, all two-voter

households that included a postal voter and all two-voter households with missing outcome data were removed from the sample. Table A.2 in the Appendix provides an overview of all remaining households listed according to primary and secondary subject partisanship.

Contact Rates and One-sided Noncompliance

The phone bank volunteers were not able to speak to all individuals who were assigned to the treatment groups. In some instances, phone lines were dead; in other instances individuals had moved away or were otherwise not reachable on a first or second attempt. Unfortunately, our manpower was not sufficient to call back more than once. We took great care in the parallel administrating of both treatments. The total contact rate -- the percentage of individuals in the respective treatment groups who were actually spoken to by volunteers -- among individuals living in two-voter households is 44% for both treatment groups. Thus, we failed to administer the treatment to 56% of those assigned to be treated. Table A.3 in the Appendix shows the contact rates by treatment and the partisan group of both household members in two-voter households. This rate of noncompliance is similar to the rate recorded in other Get-Out-the-Vote studies in the UK (John and Brannan 2008). We will return to the issue of non-compliance in the next section.

Analysis and Results

Before we delve into how the partisan composition of the household conditions spillover effects, we start off by establishing whether either of the two campaign messages were effective at mobilizing subjects who were assigned to be contacted, and which proportion of this effect was passed on to their household members. Table 1 presents the turnout rates in the control group and in the two treatment groups, for both experimental subjects and their household members.

Table 1: Turnout Rates in Treatment and Control Groups

	By Experimental Subject Partisanship			
Treatment	**Labour**	**Rival Party**	**Unattached**	**Combined**
		Experimental Subject		
Control	.174	.173	.069	.128
GOTV call	.220	.194	.142	.178
Partisan content	.223	.187	.115	.167
		Non-Experimental Subject		
Control	.170	.177	.079	.133
GOTV call	.238	.208	.150	.190
Partisan content	.218	.201	.124	.174
N	1028	1302	1746	1028

Note: Numbers exclude postal voters and individuals with missing turnout data.

At around 13 percent, turnout in the control group was low for British standards. As would be expected thanks to random assignment to the experimental and the non-experimental sample, turnout rates in the control group among experimental and non-experimental household members are similar. Table 1 also shows that for both contacted and non-contacted household members, turnout in the two treatment groups was much higher than in the respective control group. Indeed, for experimental subjects, turnout in the GOTV phone call treatment group was 39% higher than in the control group, while turnout in the partisan phone call treatment group was around 31% higher than in control. Similar increases in turnout can be observed for non-experimental subjects who lived with a subject who was assigned to receive a telephone call.

In order to see if the turnout rates amongst subjects who were assigned to treatment and subjects who lived with someone who was assigned to treatment are significantly different from the turnout rates in the control groups, we first estimate the Intent-to-Treat Effects (ITT). To estimate the ITT, we calculate the differences-in-proportions between the respective treatment and the control group, once without adjusting for pre-treatment covariates and once when adjusting for age, electoral ward, gender and turnout in seven previous elections. In both cases, we rely on the standard randomization-inference procedure (Gerber and Green 2012; Aronow and Sami 2012) to test if we can reject the

sharp null hypothesis of no treatment effect for any subject. We compare the difference in proportions using the outcome data from our experiment to the mean of all differences in proportions that we receive after simulating random assignment 10.000 times under the assumption that the treatment effect is zero for all subjects.

Table 2: Intent-to-Treat (ITT) effect and Complier Average Causal Effect (CACE) for assigned subject, by partisanship of assigned subject and including covariates.

	Labour	**Rival party**	**Unattached**	**Combined**
ITT estimates		*Randomization-Inference*		
No covariates				
GOTV treatment	.046 [-.014, .106]	.021 [-.030, .072]	.073*** [.041, .107]	.050*** [.024, .077]
Partisan content treatment	.049 [-.010, .109]	.015 [-.038, .068]	.046* [.014, .080]	.037** [.010, .063]
Covariate-adjusted				
GOTV treatment	.040 [-.018, .101]	.025 [-.023, .074]	.079*** [.048, .111]	.054*** [.029, .079]
Partisan content treatment	.044 [-.014, .101]	.015 [-.032, .064]	.051*** [.021, .083]	.039** [.013, .064]
CACE estimates		*IV Regression*		
No covariates				
GOTV treatment	.088† (.053)	.046 (.053)	.190*** (.043)	.112*** (.028)
Partisan content treatment	.112† (.064)	.029 (.047)	.121** (.041)	.089** (.029)
Covariates-adjusted				
GOTV treatment	.078 (.052)	.054 (.049)	.200*** (.040)	.120*** (.027)
Partisan content treatment	.103 (.064)	.029 (.044)	.133*** (.039)	.089** (.027)

† p-value < .10; * p-value < .05; ** p-value < .01; *** p-value < .001 (based on two-tailed hypothesis tests). Randomization inference-based 95%-CIs in brackets, standard errors in parentheses. Note: Covariates are turnout in seven previous elections, gender, age and electoral ward.

The ITT effects of both treatment conditions for each partisan group are shown in the upper halves of Tables 2 and 3. Moreover, in the last column of Tables 2 and 3 we also display the treatment effect for the entire experimental sample averaged over all partisan blocks and weighted by the inverse of the probability of assignment to treatment in each partisan block. There is little difference between the covariate-adjusted and the unadjusted ITT estimates. In the following discussion we therefore concentrate on the unadjusted estimates.

As the upper part of Table 2 shows, assignment to direct campaign contact significantly increased turnout compared to the control group by 5 and 3.7 percentage-points respectively (both increases are statistically significant with $p < 0.01$ and $p < 0.1$). However, these overall increases in turnout mask some differences depending on the targeted partisan group. The treatment without partisan content is most effective among unattached voters, and least effective for rival party supporters, yet the differences are not statistically significant. Again, the message that included partisan content had a negligible effect on turnout among rival party supporters (1.5 percentage-points), but boosted turnout among Labour Party supporters and unattached voters by 4.9 and 4.6 percentage-points (the latter is significant with a two-tailed p-value of $< .05$). We similarly estimate the ITT for the non-experimental household members separately for each partisan block and as a weighted average across all partisan blocks using inverse probability weights.

The results reported in Table 3 mirror the results for the experimental subjects reported in Table 2. Among the household members (secondary experimental subjects) of the experimental subjects who were assigned to be treated with the GOTV message without partisan content, turnout significantly increased on average by 5.7 percentage-points compared to the turnout rate among secondary experimental subjects whose household member was included in the control group.

Table 3: Intent-to-Treat (ITT) effect and Complier Average Causal Effect (CACE) for household member, by partisanship of assigned subject and including covariates

	Labour	Rival party	Unattached	Combined
ITT estimates		*Randomization-Inference*		
No covariates				
GOTV treatment	.068* [.009, .129]	.031 [-.021, .084]	.071*** [.037, .106]	.057*** [.031, .085]
Partisan content treatment	.048 [-.009, .108]	.025 [-.028, .079]	.044* [.011, .080]	.039** [.009, .064]
Covariate-adjusted				
GOTV treatment	.074* [.018, .132]	.035 [-.015, .085]	.069*** [.037, .102]	.060*** [.035, .085]
Partisan content treatment	.049† [-.008, .106]	.023 [-.025, .074]	.041* [.009, .074]	.037** [.012, .063]
CACE estimates		*IV Regression*		
No covariates				
GOTV treatment	.129* (.054)	.067 (.054)	.183*** (.043)	.129*** (.029)
Partisan content treatment	.110† (.063)	.048 (.048)	.117** (.043)	.090** (.029)
Covariates-adjusted				
GOTV treatment	.141** (.051)	.077 (.050)	.180*** (.041)	.136*** (.027)
Partisan content treatment	.112† (.061)	.045 (.045)	.106** (.040)	.085** (.027)

† p-value < .10; * p-value < .05; ** p-value < .01; *** p-value < .001 (based on two-tailed hypothesis tests). Randomization inference-based 95%-CIs in brackets, standard errors in parentheses. Note: Covariates are turnout in seven previous elections, gender, age and electoral ward.

The effect sizes vary from 3.1 percentage-points for household members of rival party supporters to 6.8 and 7.1 percentage-points for household members of Labour party supporters and unattached voters (significant with p-values of < .05 and < .001 respectively). Among secondary experimental subjects whose household member was assigned to be treated with the partisan content message, turnout increased by 3.9 percentage-points overall. Whereas turnout increased by 4.8 percentage-points for household members of Labour Party supporters and by 4.4 percentage-points for household members of unattached voters, we found that household members of rival party supporters were only 2.5 percentage points more likely to turn out.

Since we failed to administer the treatment to around 56% of our treatment groups, the ITT effects are conservative estimates of the impact of the treatment on the treated. To take account of non-compliance, we calculate the Complier Average Causal Effect (CACE) for each of the three partisan groups separately. The CACE is the average treatment effect for so-called 'compliers' (those who were successfully contacted by the campaign) only. The ITT effect of treatment assignment (z) on actually receiving the treatment (d) equals the proportion of compliers in the sample. The CACE estimator can be defined as: $\widehat{CACE} = \frac{\widehat{ITT}}{\widehat{ITT_D}}$, where $\widehat{ITT_D} = \hat{E}[d_i(1)]$.

We define compliers in the experimental group as subjects who answered the phone when called by party volunteers, and in the non-experimental group, as subjects who lived with someone who answered the phone when contacted. To obtain an estimate of the CACE, we divide the ITT effect by the contact rate. We estimated the CACE for each of the three partisan groups separately by running a two-stage least squares regression of turnout on contact, using treatment assignment as the instrumental variable.[7] The results

[7] If Y_i is turnout, D_{1i} is contact under the partisan content treatment, D_{2i} is contact under the GOTV treatment (the two endogenous variables), Z_{1i} is treatment assignment to the partisan content treatment and Z_{2i} is treatment assignment to the GOTV treatment, then the

are shown in the lower half of Tables 2 and 3. Overall, we register a substantially large and statistically significant increase in turnout of 11.2 and 8.9 percentage points as a function of directly receiving either the treatment with or without partisan content. We register an impressive increase of 19 percentage-points in turnout for unattached voters and an increase of 8.8 percentage-points in turnout for Labour voters as a result of the GOTV message that does not include any partisan content (significant with a p-value of < .001 and < 0.10). For those who receive the partisan content treatment, Labour partisans and individuals unattached to any party are 11.2 and 12.1 percentage-points more likely to vote as a result of the campaign that mentions the candidate's Labour Party affiliation. For rival party supporters the increase is statistically insignificant and amounts to only 2.9 percentage-points.

In sum, we find statistically significant and substantial CACE effects on turnout averaged across all partisan groups of 11.2 and 8.8 percentage-points as a result of the GOTV message without partisan content and the partisan content message. The effect of the partisan campaign message on the turnout rate of rival party supporters is not only non-significant, but also small in size.

The indirect impact of the GOTV message without partisan content on the turnout rate among household members living with experimental subjects who actually received the message, is on average 12.9 percentage-points, ranging from 6.7 for those living with rival party supporters, over 12.9 for those living with Labour Party supporters to 18.3 for those living with unattached voters (the latter two are statistically significant with p-values of < .05 and < .001 respectively). The indirect impact of the partisan content message on the turnout rate among household members living with experimental subjects, who actually

---

model we estimate can be formally written as: $Y_i = \beta_0 + \beta_1 D_{1i} + \beta_2 D_{2i} + \mu_i$, in which $D_{1i} = \gamma_0 + \gamma_1 Z_{1i} + \varepsilon_{1i}$ and $D_{2i} = \delta_0 + \delta_1 Z_{2i} + \varepsilon_{2i}$. Whenever specified, we included covariates in both estimation equations of the two-stage least squares regression.

received the message, is on average 9.0 percentage-points. Again, although we find a significant indirect impact on the turnout rate of those living with Labour Party supporters (11.0 percentage-points, significant with a p-value of < .10) and those living with unattached voters (11.7 percentage-points, significant with a p-value of < .01), we find no significant indirect impact on the turnout rate of those living with rival party supporters.

We can now calculate the percentage of the direct treatment effect that is passed on to the second household member. Following Nickerson (2008), we calculate this by dividing the ITT estimates reported in Table 3 by their counterparts reported in Table 2. Thus, across all partisan groups, 114% of the effect of the GOTV treatment without partisan content is passed on to secondary household members, and 105% of the partisan content treatment is passed on. This means that we find strong and consistent support for the presence of intra-household mobilization effects: Individuals two-voter households who live with someone who was contacted by the Labour campaign are more likely to vote than individuals who live with someone who was not contacted. Table A3 in the Appendix breaks down the estimates by the partisanship of the non-experimental subject. The spillover effects across all partisan groups are of course identical to those reported in Table 3. Table A3 shows that in contrast to Labour Party supporters and unattached voters, supporters of rival parties are not mobilized by their household members as a result of the Labour Party phone call. As we argue in a previous paper, this pattern of behavior is consistent with the idea that party functions as a source cue that signals to partisans whether to accept the information contained in campaign communication. If individuals support or are neutral to the source, they are likely to accept the subsequent information contained in the mobilization message and behave accordingly by turning out to vote; if they disapprove – as in the case of rival party supporters -- they are more likely to reject the content without considering the information contained in it.

Table 4: Logistic and IV regression of turnout of household member on treatment of assigned subject, conditional on partisan household composition

	Model 1	Model 2	Model 3
ITT			
Treatment	.834*	.947*	1.236
	(.373)	(.412)	(.756)
Heterogeneous	Reference	Reference	Reference
Homogenous	.170	-.066	-.101
	(.296)	(.320)	(.324)
Unattached	-.459	-.141	-.134
	(.314)	(.340)	(.345)
Heterogeneous x Treatment	Reference	Reference	Reference
Homogenous x Treatment	-.707†	-.767†	-.712
	(.394)	(.434)	(.437)
Unattached x Treatment	-.258	-.316	-.327
	(.395)	(.436)	(.438)
CACE			
Treatment	.316**	.306**	.216
	(.118)	(.110)	(.186)
Heterogeneous	Reference	Reference	Reference
Homogenous	.025	-.003	-.003
	(.043)	(.040)	(.040)
Unattached	-.081	-.018	-.026
	(.042)	(.418)	(.040)
Heterogeneous x Treatment	Reference	Reference	Reference
Homogenous x Treatment	-.276*	-.259**	-.274**
	(.123)	(.115)	(.117)
Unattached x Treatment	-.164	-.157	-.137
	(.124)	(.116)	(.117)
Covariates	No	Yes	Yes
Covariates x Treatment	No	No	Yes
N		4076	

† p-value < .10; * p-value < .05; ** p-value < .01; *** p-value < .001 (based on two-tailed t-test). Standard errors in parentheses. Note: Covariates are turnout in seven previous elections, gender, age and electoral ward.

In the next step, we estimate conditional treatment effects. We divide the sample conditional on the partisan composition of the household. Households that combine one Labour partisan and one supporter of a rival party (mostly Conservative supporters) are classified as mixed or 'heterogeneous partisan households", while households where both members fall into the rival partisan (mostly Conservative supporters), or the Labour partisan category are classified as 'homogenous partisan households". Household combinations that include unattached voters are included as "unattached households". This subgroup analysis helps us to address the question whether intra-household mobilization is more pronounced in households in which members support the same political party – homogeneous households – than in households in which members support different political parties – heterogeneous households.

The first rows in Table 4 show the estimates resulting from a logistic regression of the turnout of the secondary subject on a dummy variable indicating whether her household member was assigned to any of the two treatments (coded 1) or to the control group (coded 0), interacted with the partisan composition of the household and weighted by the inverse of the probability that each household is assigned to treatment within a given experimental block. We estimate the model once excluding pre-treatment covariates, once including pre-treatment covariates, and once including covariates and their interactions with the treatment. Moreover, we can estimate the CACE using two-stage least squares regression, in which we model turnout as a function of contact, partisan composition, and the interaction between the two, using the household members' assignment to treatment as opposed to the control group as the exogenous variable and

including inverse probability weights.[8] The estimates resulting from these three models are shown in Table 4.

The upper part of Table 4 shows the results of the logistic regression, while the lower part shows the results from the instrumental variable regression, which is equal to the CACE. Column 1 displays the unadjusted, column 2 displays the covariate-adjusted treatment effects, and column 3 displays the covariate-adjusted treatment effect accounting for observable compositional differences of heterogeneous and homogeneous partisan households. The main effect always reflects the spillover effect for the reference group, secondary subjects in heterogeneous partisan households.

Moreover, modeling observable compositional differences between heterogeneous and homogeneous households by entering interactions between covariates and the treatment does not reduce the size of the estimates. This result suggests that within-household mobilization effects are stronger, and not weaker, in households where individuals' partisanship differs. Although we are dealing with a treatment-by-covariate interaction, the second and third columns show that this relationship holds net of individual and compositional differences in age, gender, location, and turnout history.

Figure 2 shows the marginal changes in predicted probabilities based on model 1. While indirect mobilization effects amount only to around 2.5 percentage-points among homogeneous households and to around 6 percentage-points among unattached households, spillover effects in heterogeneous households amount to around 14 percentage points. Figure 3 provides a more finely grained analysis that breaks down the homogeneous

[8] If Y_i is turnout, D_i is contact (the endogenous variable), Z_i is treatment assignment, then the model we estimate can be formally written as: $Y_i = \beta_0 + \beta_1 D_i + \beta_2 Homogenous_i + \beta_3 Unattached_i + \beta_4 (D * Homogenous)_i + \beta_5 (D * Unattached)_i + \mu_i$,
in which
$D_i = \gamma_0 + \gamma_1 Z_i + \gamma_2 Homogenous_i + \gamma_3 Unattached_i + \gamma_4 (Z * Homogenous)_i + \gamma_5 (Z * Unattached)_i + \varepsilon_i$.

partisan household category into Labour households and rival party households. Moreover, we divide the unattached category into households that include one unattached subject and one partisan and households where both members are unattached.

Figure 2: ITT effect of campaign message on household (unassigned) member, conditional on household partisan composition (90% CIs).

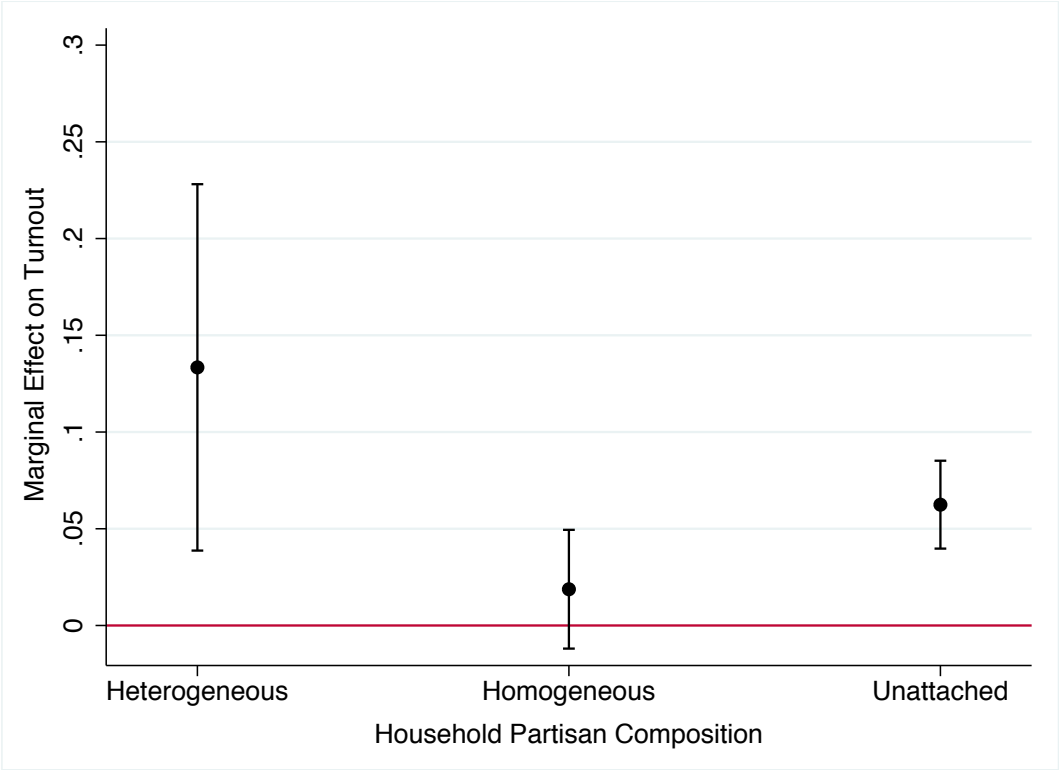
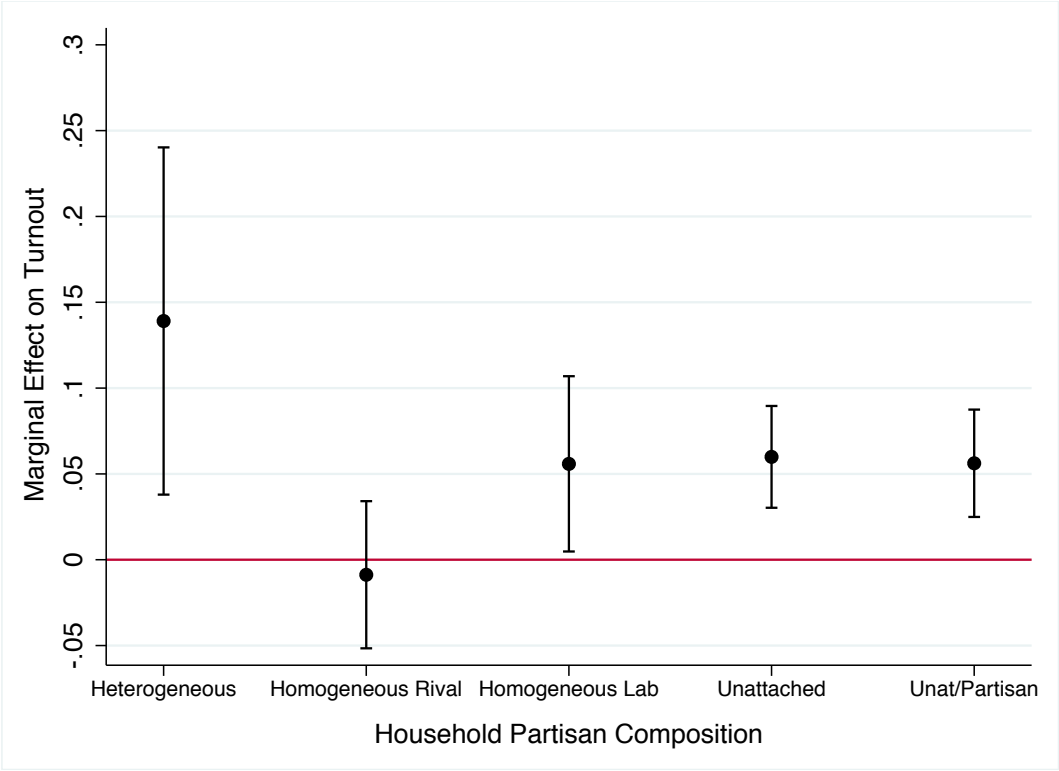


Figure 3 underlines that spillover effects of campaign contact are larger in heterogeneous partisan households than in homogeneous rival, Labour or unattached households, with a statistically significant difference for the homogeneous versus rival partisan comparison. The logistic regression analyses on which the predicted probabilities are based are shown in Table A4 in the Appendix. Moreover, we display the conditional treatment effects adjusted for observable compositional differences in previous turnout rates, age, gender and electoral ward in Figure A1 in the Appendix.

Figure 3: ITT effect of campaign message on household (unassigned) member, conditional on household partisan composition (90% CIs).



What do these results tell us? First of all, we find no evidence to support the hypothesis that individuals are more likely to mobilize a co-partisan than a household member who is supporting another party as a result of being contacted by the Labour campaign. Although we find that campaign messages spillover in homogenous households consisting of two Labour Party supporters or two voters who have not committed themselves to any party (unattached voters), mobilization effects are no greater than in heterogeneous partisan households. In contrast, we find some support for the idea that spillover is greater when partisans disagree on the party that they support.

An implicit test of the conflict-avoidance mechanism

The previous results demonstrate that within-household mobilization plays an important role in partisan election campaigns, and that in contrast to what the conflict-avoidance mechanism suggests, heterogeneous partisan households are conductive to indirect

campaign mobilization. We implicitly test the effects of raising the level of partisan conflict within the household by adding partisan content and party labels to the campaign message. Adding value-laden references to the Labour and Conservative Parties to a telephone message should result in a higher possibility of household members taking up the respective parties in conversation.

This manipulation of partisan content allows us to evaluate the conflict-avoidance hypothesis by testing whether a campaign message is more or less likely to spill over to another household member who does not share the same partisan preferences if it contains strong partisan content and is advocating in favor of voting for the party that initiates contact. Table 5 shows the effects of infusing a campaign message with partisan content on within household mobilization, conditional on the partisan composition of the household.[9]

Overall, the results in columns 1 and 2 show that the campaign message that was openly partisan resulted in lower within household mobilization than the message without partisan content. However, the difference-in-proportions was only slightly more than 4 percentage-points for secondary subjects who lived with a complier, and hardly significant is statistical terms (see Tables A6 and A7 in the Appendix).

One could therefore argue that while we find some support for the hypothesis that spillover effects are smaller if the primary experimental subject receives a message with partisan content than a message without partisan content, the evidence is not overwhelming. Moreover, columns 3, 4 and 5 in Table 5 display models that include interactions between the partisan composition of the household and the partisan content of the message.

[9] Since the two treatments have been randomly assigned to primary experimental subjects, we can assume the share of compliers to be equal in both treatment groups. Therefore, we can compare the household members of compliers directly to each other, using logistic regression for estimation.

Table 5: Logistic regression results of turnout of household member on compliance with partisan content vs no content treatment, conditional on partisan composition

	Model 1	Model 2	Model 3	Model 4	Model 5
Treatment	-.242	-.242	1.364†	1.368	2.363
	(.171)	(.190)	(.781)	(.858)	(1.555)
Heterogeneous	Reference	Reference	Reference	Reference	Reference
Homogeneous	-.353	-.817	.626	.139	.275
	(.374)	(.424)	(.659)	(.703)	(.714)
Unattached	-.493	-.507	.508	.461	.587
	(.418)	(.473)	(.685)	(.738)	(.750)
Heterogeneous x Treatment	-	-	Reference	Reference	Reference
Homogeneous x Treatment	-	-	-1.675*	-1.691†	-1.917*
			(.819)	(.897)	(.918)
Unattached x Treatment	-	-	-1.724*	-1.711†	-1.915*
			(.819)	(.903)	(.919)
Covariates	No	Yes	No	Yes	Yes
Covariates x Treatment	No	No	No	No	Yes
N	782				

† p-value < .10; * p-value < .05; ** p-value < .01; *** p-value < .001 (based on two-tailed t-test). Standard errors in parentheses. Note: Covariates are turnout in seven previous elections, gender, age and electoral ward.

The statistically significant interaction terms in Models 3 to 5 clearly show that adding partisan content had a more positive effect on turnout for subjects who lived in heterogeneous partisan households than on subjects who lived in homogeneous and unattached partisan households. While column 4 shows the estimates unadjusted for compositional differences *between* homogeneous and heterogeneous households, column 5 adds interactions between pre-treatment covariates and the treatment to the equation.

Table 5 also shows that subjects whose partisanship differed from the partisanship

of the person who was contacted were substantively more, not less likely to turn out when their partner was contacted with a partisan message than when she was contacted with a message without partisan content.

Figure 4: CACE effect of partisan content on mobilization of household member, conditional on household partisan composition (90% CIs).

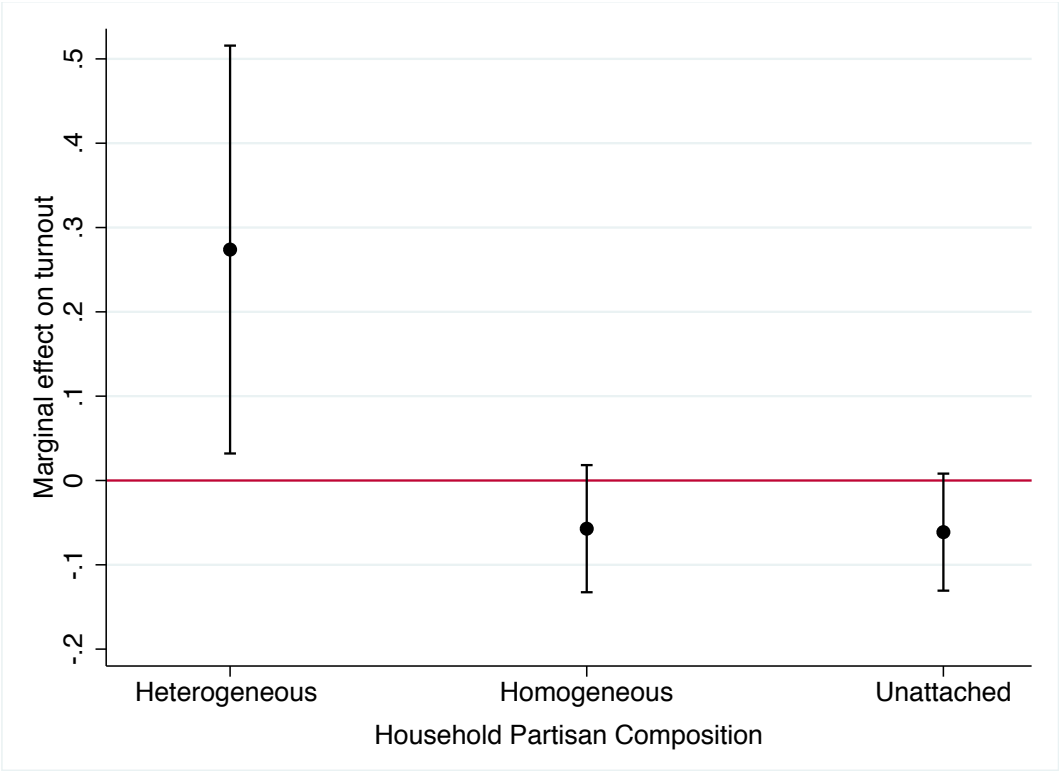
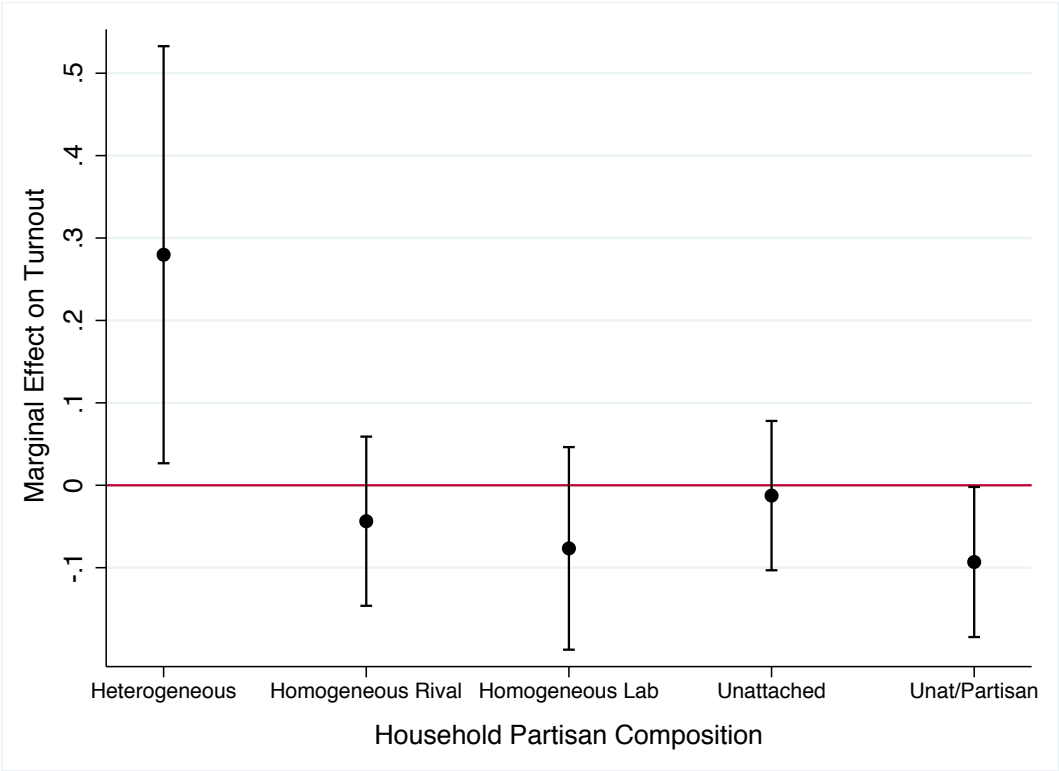


Figure 4 displays the marginal changes in predicted probabilities (from model 3 in Table 5) associated with a secondary subject's household member receiving the campaign message with partisan content as opposed to the message without partisan content, conditional on the partisan composition of the household. We see that experimental subjects living in homogeneous partisan households and subjects living in households that include at least one unattached voter are slightly less likely to mobilize their spouses and housemates once the campaign message can be identified as openly partisan. However, for both household types, the differential impact of receiving one of the two messages is not significantly different from zero. In contrast, Figure 4 shows that individuals in homogeneous partisan households respond positively to the introduction of partisan content. Again, this

interaction effect is not driven by the choice to group Labour and rival party households together in the same category as "homogeneous partisan households". As Figure 5 demonstrates, this effect is equally pronounced when comparing heterogeneous partisan households separately to homogeneous Labour households and households where both members support rival parties. Figure A2 in the Appendix displays the predicted probabilities net of observable compositional differences. They show that modeling observable differences between homogeneous and heterogeneous households does not reduce the size or statistical significance of the conditional treatment effects.

Figure 5: CACE effect of partisan content on mobilization of household member, conditional on household partisan composition (90% CIs).



Modeling partisan dynamics

The finding that campaign spillover effects are highly pronounced in heterogeneous partisan households is difficult to reconcile with the conflict avoidance mechanism. In what follows we try to further disentangle partisan dynamics within the household by looking at the three-way-interaction between the partisanship of the contacted subject, the

partisanship of her household member and the partisan phone call.

Figure 6 displays the marginal changes in predicted probabilities resulting from a logistic regression of turnout on the partisan phone call versus the control group, the partisanship of the experimental subject, the partisanship of her housemate and the two- and three-way interactions between treatment and partisan affiliations of both subjects. The model also includes pre-treatment covariates adjusting for individual observable differences between subjects, and for the interactions between covariates and the partisan phone call. The Rival, Labour and Unattached categories listed on the X-axis refer to the partisanship of the secondary subject, meaning the household member who was not exposed to direct campaign mobilization by the Labour Party. Figure 6a displays the indirect mobilization effects on the household member when the contacted subject was a rival partisan, 6b refers to situations where the contacted subject was a Labour partisan, and 6c displays indirect mobilization effect if the contacted subject was unattached to any party.

Figure 6: Partisan mobilization dynamics within Households (90% CIs).

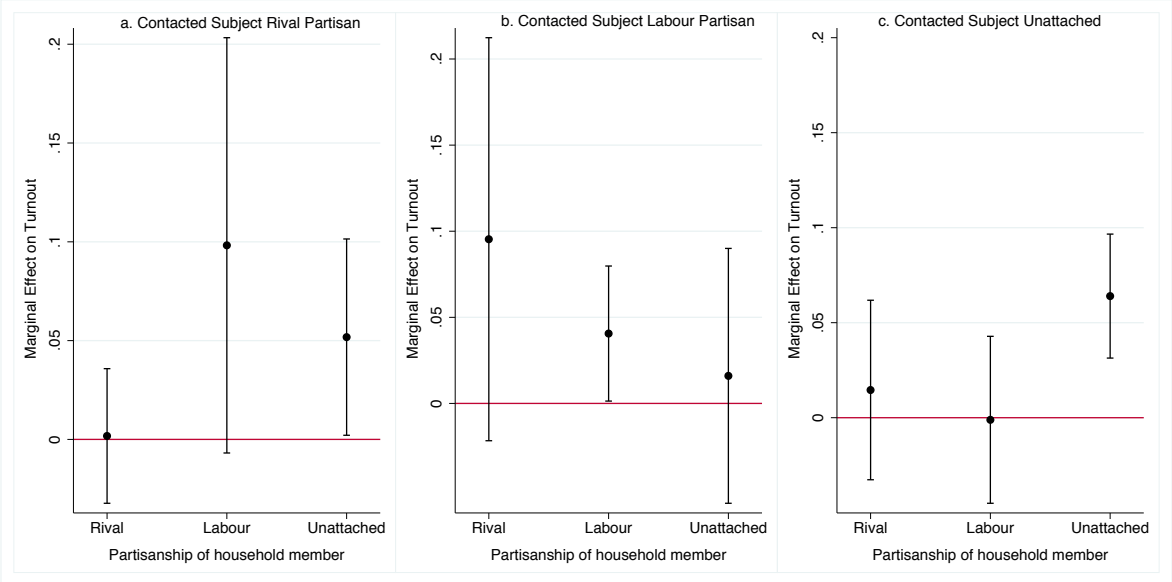


Figure 6a shows that rival partisans mobilize Labour voters if they are contacted by a Labour campaign, and 6b shows that Labour partisans mobilize rival partisans as a

consequence of Labour campaign contact. Both indirect mobilization effects amount to around 10 percentage-points. While the effects for heterogeneous households are noisy due to small sample sizes, they are substantially large, and seem even more pronounced than indirect mobilization effects between two Labour supporters (although the difference between the two effects is not statistically significant). Moreover, Figure 6c shows that campaign contact does not spill over if rival and Labour partisans live together with an unattached subject who is contacted by the Labour campaign. These results suggest that partisans pass on campaign messages to their household members even if a party they oppose initiates contact. Moreover, Figures 6a and 6b suggest that campaign messages are more effective if at least one household member supports the party that initiates contact.

Discussion and Conclusion

An increasing number of political scientists view turnout as a social activity. Spouses, friends and neighbors influence decisions about whether or not to participate politically and who to support (Huckfeldt 1995; Huckfeldt et al. 2004; Zuckerman 2005; Zuckerman et al. 2007; Rolfe 2012; Sinclair 2012). In this paper we look at a specific instance of social influence, namely that which occurs among household members as a result of election campaigns and underlies the decision to go out and vote.

Combining experimental manipulation of campaign contact with a unique set of data on household partisan attachments, we find support for those scholars who have questioned whether heterogeneous partisan environments really lead to lower intra-household mobilization (Huckfeldt et al. 2004; Klofstad et al. 2013) during election campaigns. We show that within household mobilization effects are larger, not smaller, as a result of campaign contact in households that consist of a Labour supporter and a supporter of a rival party. Moreover, we attempt to shed some light on the most important mechanism

that is hypothesized to lead to lower electoral participation: the desire to avoid partisan discussions and avoid conflict over partisan politics (Mutz 2006). We test this mechanism implicitly by adding and subtracting the partisan element from the treatment (Gerber and Green 2012). We assume that an openly partisan campaign message that endorses the Labour Party and blames the Conservative Party for cuts in police spending will result in a higher potential of partisan disagreement between a household member who supports the Labour Party and one who support the Conservative Party than a message that does not mention parties or partisan politics. Following this strategy, we find that partisan content does not negatively impact on mobilization within households where individuals hold opposing partisan attachments. Instead, partisan content seems to increase indirect mobilization in heterogeneous partisan households. Although this result might be unexpected there are good reasons to think that household members who are used to disagreeing with each other on partisan politics do so in good spirits and do not withhold relevant campaign messages from each other.

This paper also makes a novel contribution to the literatures on partisan GOTV campaigns (Arceneaux 2010). By investigating the direct and indirect mobilization effects of a party's field campaign, we find that more than 100% of the direct mobilization effect spills over to a second household member. This result provides new evidence for the hypothesis that partisan election campaigns have large indirect mobilization effects that have to be taken into account when looking at the cost-effectiveness of different voter mobilization methods. While the existence of an indirect effect has of course been demonstrated in nonpartisan GOTV experiments (Nickerson 2008; Sinclair et al. 2012), we are the first to show that they are also present in partisan campaigns. As Nickerson and colleagues (2006) caution, partisan and non-partisan campaigns differ in multiple ways, and intra-household partisan dynamics are widely expected to influence the effectiveness

of indirect campaign mobilization (Huckfeldt et al. 2004). The relatively large size of the direct mobilization effect that spills over to the household member, especially in comparison to studies using placebo designs (Nickerson 2008; Sinclair et al. 2012), might also be a function of the research design. Mann and Sinclair (2014) find in a recent paper that spillover effects are higher when the low-propensity voter is targeted than when the high propensity voter is targeted in a household. They suggest that the shock to a social network might be larger if household members act inconsistently with past behavior. As a function of our research design, the lower propensity voter had an equal probability of being assigned to contact than the higher propensity voter. In contrast, in Nickerson's (2009) placebo design the higher propensity voter might be more likely to self-select into opening the door when the canvasser knocks. Importantly, we also show that a party's campaign has the smallest indirect mobilization effects in households where both household members oppose the party that initiates contact. This intuitive finding suggests that the campaign message did not have the adverse effect of mobilizing individuals who are unlikely to vote for the party in the first place.

Of course, our findings might be context-dependent and are based on a relatively small sample of individuals living in heterogeneous partisan households. The results of this experiment, therefore, need replication using larger sample sizes in order to obtain more precise estimates. Moreover, our findings are limited to the household, and might not apply to social networks in which ties are weaker, and political discussion more easily avoided. By exploiting the downstream benefits of a fully randomized partisan campaign experiment to study within household mobilization, this study should nonetheless inform the ongoing debates about the political costs of partisan disagreement, and the behavioral consequences of heterogeneous political networks.

List of References

Arceneaux, Kevin, Alan S. Gerber, and Donald P. Green. 2006. "Comparing Experimental and Matching Methods using a Large-Scale Voter Mobilization Experiment". *Political Analysis* 14: 1-36.

Arceneaux, Kevin. 2010. "The Benefits of Experimental Methods for the Study of Campaign Effects." *Political Communication* 27(2): 199-215.

Arceneaux, Kevin, Alan S. Gerber, and Donald P. Green. 2010. "A Cautionary Note on the Use of Matching to Estimate Causal Effects: An Empirical Example Comparing Matching Estimates to an Experimental Benchmark". *Sociological Methods and Research* 39: 256-282.

Aronow, Peter M., and Cyrus Samii. 2012. ri: The R package for performing randomization-based inference for experiments. R package version 0.9. http://CRAN.R-project.org/package=ri.

Belanger, Paul and Munroe Eagles. 2007. "Partisan Cross-Pressure and Voter Turnout: The Influence of Micro and Macro Environments." *Social Science Quarterly* 88(3): 850-867.

Bello, Jason and Meredith Rolfe. 2013. "Is influence mightier than selection? Forging agreement in political discussion networks during a campaign." *Social Networks* 36: 134-146.

Cutts David and Edward Fieldhouse. 2009. "What small spatial scales are relevant as electoral contexts for individual voters? The importance of the household on turnout at the 2001 General Election." *American Journal of Political Science* 53(3): 726-739.

Robert M. Bond, Christopher J. Fariss, Jason J. Jones, Adam D. I. Kramer, Cameron Marlow, Jaime Settle and James H. Fowler "A 61 -Million-Person Experiment in Social Influence and Political Mobilization". *Nature* 489: 295 -298.

Coleman, Stephen. 2004. "The effect of social conformity on collective voting behavior." *Political Analysis* 12(1): 76-96.

Conover, Pamela, Don Searing, and Ivor Crewe. 2002. "The Deliberative Potential of Political Discussion." *British Journal of Political Science* 32(1): 21–62.

Festinger, Leon. *A Theory of Cognitive Dissonance*. Vol. 2. 1962. Stanford: Stanford University Press.

Fieldhouse, Edward and David Cutts. 2012. "The Companion Effect: Household and Local Context and the Turnout of Young People." *The Journal of Politics* 74(3): 856-869.

Fitzgerald, Jennifer and Amber K. Curtis. 2012. "Partisan discord in the family and political engagement: A comparative behavioral analysis". *The Journal of Politics* 74(1): 129-141.

Fowler, James, Michael Heaney, David Nickerson, John Padgett, and Betsy Sinclair. 2011. "Causality in Political Networks." *American Politics Research* 39(2): 437–80.

Green, Donald P., and Alan S. Gerber. 2002. "The Downstream Benefits of Experimentation". *Political Analysis* 10(4): 394-402

Gerber Alan S. and Donald P Green. 2012. *Field Experiments: Design, Analysis and Interpretation*. New York: WW Norton.

Hirschman, Albert O. *Exit, voice, and loyalty: 1970. Responses to decline in firms, organizations, and states*. Cambridge: Harvard University Press, 1970.

Huckfeldt, Robert R., Paul E. Johnson, and John D. Sprague. 2004. *Political disagreement: The survival of diverse opinions within communication networks*. Cambridge: Cambridge University Press.

Huckfeldt, Robert R. 1995. *Citizens, politics and social communication: Information and influence in an election campaign*. Cambridge University Press.

Iyengar, Shanto and Adam F. Simon. 2000. "New Perspectives and Evidence on Political Communication and Campaign Effects". *Annual Review of Psychology* 51:149-169.

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(3): 565-574.

Klofstad, Casey.A. 2010. "The Lasting Effect of Civic Talk on Civic Participation: Evidence from a Panel Study." *Social Forces* 88: 2353-2375.

Klofstad, Casey A. 2011. *Civic Talk: Peers, Politics, and the Future of Democracy*. Philadelphia: PA: Temple University Press.

Klofstad, Casey.A., Sokhey, A.E., & McClurg, S.D. 2013. "Disagreeing About Disagreement: How Conflict in Social Networks Affects Political Behavior." American Journal of Political Science 57: 120-134.

Manski, Charles. 1993. "Identification of Exogenous Social Effects: The Reflection Problem." *Review of Economic Studies* 60: 531–42.

Mann, Christopher B. and Betsy Sinclair. "Two Votes For the Price of One: Targeted Messages and Spillover Effects in Early Voting". Southern Political Science Association Annual Meeting, 2014.

McClurg, Scott D. 2003. "Social Networks and Political Participation: The Role of Social Interaction in Explaining Political Participation." *Political Research Quarterly* 56(4): 448–64.

Mutz, Diana C. 2002. "The Consequences of Cross-Cutting Networks for Political Participation." *American Journal of Political Science* 46(4): 838–55.

Mutz, Diana C. 2006. *Hearing the Other Side: Deliberative versus Participatory Democracy*. Cambridge: Cambridge University Press.

Nelson, Philipp J. 1994. "Voting and imitative behavior". *Economic Inquiry* 82: 92–102.

Nelson, Philipp J. and Greene, Kenneth V. 2003. *Signaling Goodness: Social Rules and Public Choice*. Ann Arbor: University of Michigan Press.

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): 85-97.

Nickerson, David W. 2008. "Is Voting Contagious? Evidence from Two Field Experiments." *American Political Science Review* 102(1): 49–57.

Nir, Lilack. 2011. Disagreement and Opposition in Social Networks: Does Disagreement Discourage Turnout? *Political Studies* 59: 674–692.

Pattie, Charles J. and Johnston, Ron J. 2001. "Talk as a political context: conversation and electoral change in British elections, 1992-1997". *Electoral Studies* 20(1): 17-40.

Rolfe, Meredith. 2012. *Voter Turnout: A Social Theory of Political Particip*ation. Cambridge University Press, Cambridge, UK.

Scheufele, Dietram, A., Matthew C. Nisbet, Dominque Brossard, and Erik C. Nisbet. 2004. "Social Structure and Citizenship: Examining the Impacts of Social Setting, Network Heterogeneity, and Informational Variables on Participation". *Political Communication*, 21(3): 315–38.

Scheufele, Dietram A., Bruce W. Hardy, Dominique Brossard, Israel S. Waismel-Manor, and Erik C. Nisbet. 2006. "Democracy Based on Difference: Examining the Links between Structural Heterogeneity, Heterogeneity of Discussion Networks and Democratic Citizenship", *Journal of Communication* 56(4): 728–53.

Sinclair, Betsy, Margaret McConnell, and Donald P. Green. 2012. "Detecting Spillover Effects: Design and Analysis of Multi-level Experiments." *American Journal of Political Science* 56: 1055-1069.

Sinclair, Betsy. 2012. *The Social Citizen: Peer Networks and Political Behavior*. Chicago:

University Of Chicago Press.

Zaller, John. 1992. *The nature and origins of mass opinion*. Cambridge: Cambridge University Press.

Zuckerman, Alan. S. 2005. *The Social Logic of Politics. Personal Networks As Contexts for Political Behavior.* Philadelphia: Temple University Press.

Zuckerman, A. S., Fitzgerald, J. and Dasovic, J. 2007. *Partisan Families*. New York: Cambridge.

Appendix

Table A.1: Covariate balance among non-experimental subjects

	Control	GOTV Call	Partisan
Mean Age	36.53	36.07	37.14
Age Missing	.613	.611	.612
Female	.434	.456	.446
Male	.461	.454	.451
Unknown	.105	.091	.104
Ward 1	.268	.267	.305
Ward 2	.127	.134	.138
Ward 3	.257	.268	.253
Ward 4	.348	.331	.304
Voted Local 2012	.497	.529	.507
Voted Local 2011	.584	.588	.583
Voted General 2012	.751	.775	.756
Voted Election 1	.419	.404	.396
Voted Election 2	.393	.370	.386
Voted Election 3	.475	.466	.500
Voted Election 4	.474	.463	.480

Note: Figures exclude postal voters and individuals with missing turnout data.

Table A.2: Frequencies of household partisanship combinations.

Experimental subject	**Non-experimental subject**			
	Labour	Rival party	Unattached	***Combined***
Labour	706	77	245	***1028***
Rival party	89	833	380	***1302***
Unattached	258	355	1133	***1746***
Combined	***1053***	***1265***	***1758***	***4076***

Note: Numbers exclude postal voters and individuals with missing turnout data.

Table A.3: Household contact rates by partisanship, in percentages.

Experimental subject	**Non-experimental subject**			
	Labour	Rival party	Unattached	***Combined***
Partisan content treatment				
Labour	48.43	50.00	27.78	***43.67***
Rival party	44.44	54.82	46.67	***51.24***
Unattached	32.65	46.67	36.21	***37.92***
Combined	***44.68***	***52.14***	***37.50***	***43.78***
GOTV treatment				
Labour	58.78	53.33	36.67	***52.47***
Rival Party	45.00	45.31	48.05	***46.02***
Unattached	32.89	45.24	38.20	***38.68***
Combined	***49.59***	***45.70***	***40.00***	***44.42***

Note: Numbers exclude postal voters and individuals with missing turnout data.

Table A4: Intent-to-Treat (ITT) effect and Complier Average Causal Effect (CACE) for experimental subjects, by non-experimental subject partisanship

	Labour	Rival party	Unattached	Combined
ITT estimates		*Randomization-Inference*		
No covariates				
GOTV treatment	.080* [.024, .139]	.030 [-.023, .083]	.062*** [.027, .098]	.057*** [.031, .085]
Partisan content treatment	.051† [-.006, .110]	-.004 [-.060, .051]	.062** [.028, .098]	.039** [.009, .064]
Covariate-adjusted				
GOTV treatment	.079** [.026, .134]	.035 [-.016, .084]	.064*** [.031, .098]	.060*** [.035, .085]
Partisan content treatment	.043 [-.01, .099]	-.007 [-.056, .046]	.055*** [.026, .092]	.037** [.012, .063]
CACE estimates		*IV Regression*		
No covariates				
GOTV treatment	.157** (.054)	.065 (.054)	.154*** (.043)	.129*** (.029)
Partisan content treatment	.114† (.061)	-.008 (.047)	.1667*** (.046)	.090** (.029)
Covariates-adjusted				
GOTV treatment	.157** (.051)	.076 (.052)	.161*** (.040)	.136*** (.027)
Partisan content treatment	.100† (.060)	-.014 (.043)	.146*** (.042)	.085** (.027)

† p-value < .10; * p-value < .05; ** p-value < .01; *** p-value < .001 (based on two-tailed hypothesis tests). Randomization inference-based 95%-CIs in brackets, standard errors in parentheses. Note: Covariates are turnout in seven previous elections, gender, age and electoral ward.

Table A5: Logistic regression results of turnout of household member on treatment assignment of experimental subject, conditional on partisan household composition

	Model 1	Model 2	Model 3
ITT			
Treatment	.831**	.924*	1.118
	(.266)	(.413)	(.760)
Labour/Rival Party	Reference	Reference	Reference
Rival Party/Rival Party	.209	-.086	-.109
	(.229)	(.351)	(.357)
Labour/Labour	.119	-.082	-.118
	(.233)	(.349)	(.354)
Unattached	-1.070***	-.052	-.038
	(.242)	(.396)	(.340)
Unattached/Labour/Rival	-.478	-.106	-.100
	(.231)	(.346)	(.350)
Labour/Rival Party x Treatment	Reference	Reference	Reference
Rival/Rival x Treatment	-.890**	-.961*	-.892†
	(.294)	(.454)	(.454)
Lab/Lab x Treatment	-.497†	-.502	-.482
	(.296)	(.458)	(.462)
Unattached x Treatment	-.128	-.075	-.105
	(.305)	(.472)	(.480)
Unattached/Lab/Rival x Treatment	-.360	-.413	-.415
	(.291)	(.451)	(.454)
Covariates	No	Yes	Yes
Covariates x Treatment	No	No	Yes
N		4076	

† p-value < .10; * p-value < .05; ** p-value < .01; *** p-value < .001 (based on two-tailed t-test). Standard errors in parentheses. Note: Covariates are turnout in seven previous elections, gender, age and electoral ward.

Table A6: Indirect Complier Average Causal Effect (CACE) of partisan content versus no content treatment by household partisan composition, no covariates

Experimental subject	Non-experimental subject			
	Labour	Rival party	Unattached	Combined
Labour	-.074	.250	.039	-.029
	[-.214, .071]	[-.188, .688]	[-.257, .308]	[-.149, .091]
N	164	16	37	217
Rival party	.278	-.044	-.002	-.006
	[-.122, .688]	[-.158, .074]	[-.207, .199]	[-.098, .090]
N	21	178	79	278
Unattached	-.298†	-.144	-.013	-.088†
	[-.541, .029]	[-.342, .050]	[-.119, .091]	[-.180, .004]
N	41	73	173	287
Combined	-.082	-.054	-.003	-.042
	[-.204, .038]	[-.154, .045]	[-.091, .091]	[-.100, .018]
N	226	267	289	782

† p-value < .10; * p-value < .05; ** p-value < .01; *** p-value < .001 (based on two-tailed tests of the sharp null hypothesis). Randomization inference-based 95%-CIs in brackets.

Table A7 Indirect Complier Average Causal Effect (CACE) of partisan content versus no content treatment by household partisan composition, covariate-adjusted

Experimental subject	**Non-experimental subject**			
	Labour	Rival party	Unattached	Combined
Labour	-.058	-.141	-.041	-.034
	[-.198, .083]	[-.903, .621]	[-.333, .2808]	[-.151, .086]
N	164	16	37	217
Rival party	.346	-.075	-.035	-.033
	(-.251, .943)	[-.195, .045]	[-.229, .158]	[-.130, .061]
N	21	178	79	278
Unattached	-.288*	-.163	.025	-.067
	[-.517, -.066]	[-.371, .045]	[-.084, .132]	[-.153, .021]
N	41	73	173	287
Combined	-.053	-.076	-.002	-.046
	[-.170, .064]	[-.171, .019]	[-.083, .086]	[-.100, .010]
N	226	267	289	782

† p-value < .10; * p-value < .05; ** p-value < .01; *** p-value < .001 (based on two-tailed tests of the sharp null hypothesis). Randomization inference-based 95%-CIs in brackets. Note: Covariates are turnout in seven previous elections, gender, age and electoral ward.

Table A8 Logistic regression results of turnout of household member on assignment of experimental subject to partisan phone call vs control, conditional on partisan household composition

	Model 1	Model 2	Model 3
Treatment	.356**	-.045	.672
	(.120)	(.222)	(.815)
Person 1 Rival	Reference	Reference	Reference
Person 1 Labour	.138	.143	.068
	(.180)	(.223)	(.447)
Person 1 Unattached	-.243	-.279	-.154
	(.149)	(.182)	(.265)
Person 2 Rival	Reference	Reference	Reference
Person 2 Labour	.032	-.094	.284
	(.183)	(.227)	(.462)
Person 2 Unattached	.166*	-.032	-.035
	(.148)	(.181)	(.262)
Person 1 Rival x Treatment	-	Reference	Reference
Person 1 Lab x Treatment		.017	1.025
	-	(.378)	(.791)
Person 1 Unat x Treatment		.160	.207
	-	(.314)	(.499)
Person 2 Rival x Treatment		Reference	Reference
Person 2 Lab x Treatment	-	.431	.911
		(.386)	(.693)
Person 2 Unat x Treatment	-	.647*	.649
		(.315)	(.451)
Person 1 Lab x Person 2 Lab	-	-	-.250
			(.644)
Person 1 Lab x Person 2 Unat	-	-	.151
			(.589)
Person 1 Unat x Person 2 Lab	-	-	-.710
			(.620)
Person 1 Unat x Person 2 Unat	-	-	-.058
			(.389)
Person 1 Lab x Person 2 Lab x Treatment	-	-	-1.469
			(1.049)
Person 1 Lab x Person 2 Unat x Treatment	-	-	-1.505
			(1.027)
Person 1 Unat x Person 2 Lab x Treatment	-	-	-1.171
			(1.041)
Person 1 Unat x Person 2 Unat x Treatment	-	-	.057
			(.690)
Covariates included	Yes	Yes	Yes
N		3171	

† p-value < .10; * p-value < .05; ** p-value < .01; *** p-value < .001 (based on two-tailed t-test). Standard errors in parentheses. Note: Covariates are turnout in seven previous elections, gender, age and electoral ward.

Figure A1 ITT effect of treatment assignment on mobilization of household member based on Table 4 Model 3, conditional on household partisan composition (90% CIs).

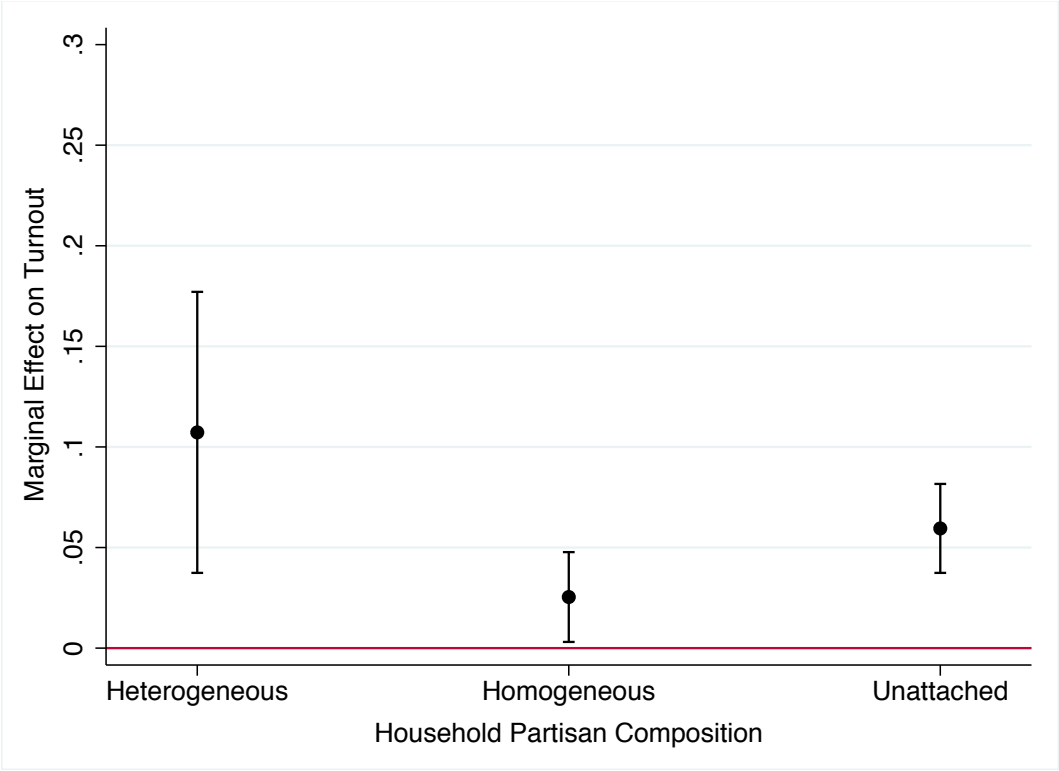
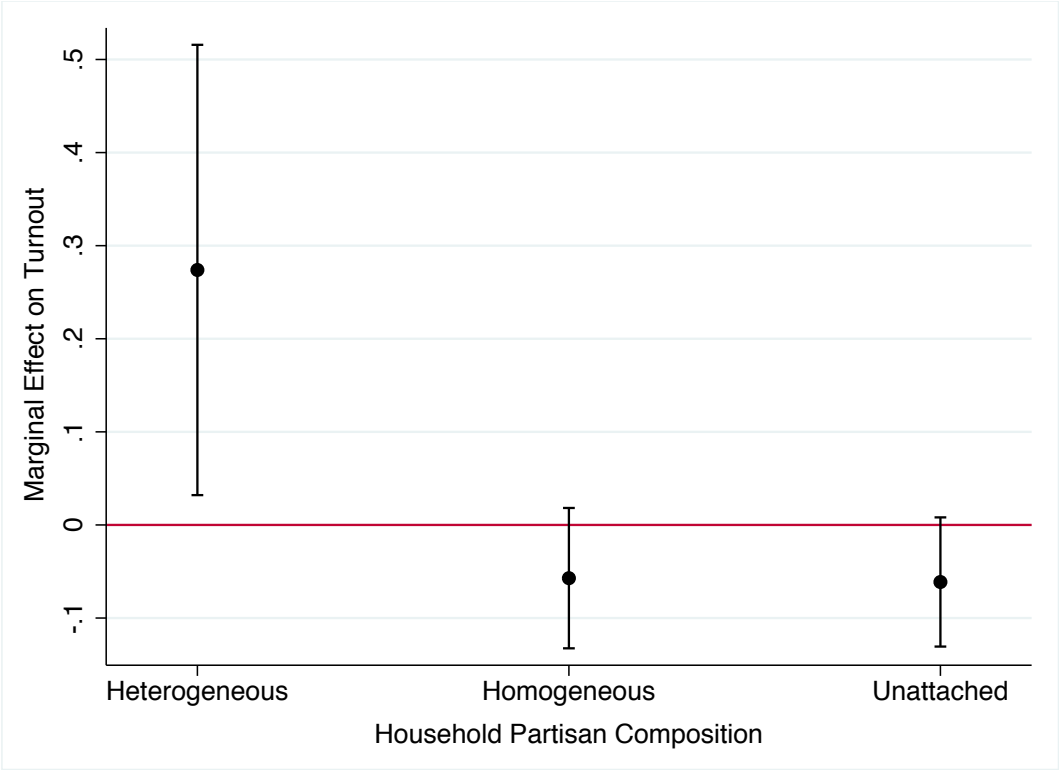


Figure A2 CACE of partisan content on mobilization of household member based on Table 5 Model 5, conditional on household partisan composition (90% CIs).



REACHING ACROSS THE PARTISAN DIVIDE:
A Randomized Persuasion Experiment Using Hand-Written Letters.

Florian Foos

This randomized persuasion experiment tests whether the act of signaling costly effort on the part of a UK Parliamentary Candidate persuades supporters of other parties to change their voting intentions. In co-operation with a first-time Parliamentary Candidate for the Labour Party, I randomly assigned whether constituents who support other parties, received a personalized letter, hand-written by the Candidate. In the letter, she introduces herself and invites subjects to a face-to-face meeting. Although only two experimental subjects took-up the invitation to meet the candidate, in a post-treatment survey I find some support for the hypothesis that the letter has a positive effect on a subjects' voting intentions and on how they evaluate the Candidate. However, the study also highlights important limitations of this persuasion-strategy: First, it shows that individuals who support rival parties were largely unwilling to personally engage with the Labour Party Candidate (only 1.3% replied to her letter). Secondly, 'persuadability' depends on partisanship. While the treatment proved effective with supporters of smaller parties, it was ineffective with those who support Labour's principal rival, the Conservative Party. This study constitutes the first randomized field experiment on political persuasion in the United Kingdom, and suggests that Candidates might be able to positively influence supporters of other parties, by signaling that they are willing to incur substantial costs to discuss their concerns.

Introduction

It is well understood that political campaigns seek to both mobilize and persuade. If successful, persuasion can be particularly powerful because it increases the vote share of a candidate while at the same time decreasing the vote share of the competitor (Bailey et al. 2013). Nevertheless, many scholars and practitioners are skeptical whether campaigns are able to change individuals' voting intentions. Individuals are thought to resist persuasion attempts (Bartels 1998; Arceneaux and Kolodny 2009; Bailey et al. 2013). Moreover, attempts to persuade supporters of other parties might prove counter-productive because they can backfire by reinforcing negative attitudes about out-party candidates, and by increasing the turnout of voters who are unlikely to support the candidate (Arceneaux and Kolodny 2009; Bailey et al 2013).

There are two major reasons why persuasion might prove costly and ineffective in real-world elections; first, most voters are 'coarse thinkers' (Mullainathan et al. 2008) who tend to think in categories rather than weigh the merit of each candidate and each issue. They rely on what psychologists call 'cues' and what economists refer to as 'signals' to form opinions about candidates (Conover and Feldman 1989; Lupia and McCubbins 1998; Potter and Gray 2008). Second, elections are not fought in a political vacuum but campaigns operate within a partisan universe. At the start of an election campaign large shares of the electorate are pre-committed to candidates because candidates are affiliated with political parties (Bartels 1998; Green et al. 2002). Therefore, 'party' is often described as the strongest signal that a candidate sends to voters during an election campaign (Jones and Hudson 1998). Party brands evoke partisan stereotypes (Green et al. 2002), and the signal activates long-standing partisan attachments in the electorate (Aldrich 1995). Therefore partisans might reject campaign communication sent by candidates affiliated with rival parties (Zaller 1992).

However, some of the literature on election campaigns suggests that not all partisans are immune to persuasion. Indeed, some supporters of other parties are thought to be "persuadable" (Hillygus and Shields 2008, Lavine et al. 2012; Nickerson and Roberts 2014) because they are "ambivalent" (Lavine et al. 2012) or "cross-pressured" (Hillygus and Shields 2008). Partisans are usually described as ambivalent or cross-pressured if their short-term evaluations of candidate performance or characteristics conflict with their long-standing partisan loyalties. If a candidate wants to win over voters who are not her natural supporters, Hillygus and Shields (2008) advise a candidate to prime these conflicting short-term evaluations.

It is however difficult for a candidate to know who those voters are. In a multi-party system such as the one in Britain, an obvious signal of persuadability is a voter's partisanship. For a Labour candidate supporters of smaller parties such as the Liberal Democrats or the UK Independence Party (UKIP) might be more 'persuadable' than supporters of the main rival party, the Conservative Party. This is because even if Conservative supporters are cross-pressured, they are less likely to defect to the main competitor and instead opt for a smaller party, such as the UK Independence Party (Ford and Goodwin 2014).

Party affiliation is not the only cue used by voters to maneuver the political environment (Lau and Redlawsk 2001). As Conover and Feldman (1989: 917) write "various cues can prompt voters to place candidates in different categories". Furthermore, signals might be especially important in defining a candidate who is relatively unknown (Conover and Feldman 1989). Building on these two insights and knowing that voters are coarse thinkers who rely on signals to form political opinions, letters containing verbal and non-verbal cues about a first-time parliamentary candidate are sent to subjects. In collaboration with the Labour parliamentary candidate for Southampton Itchen, Rowenna

Davis, randomly selected constituents who voted for parties other than the Labour in the 2010 General Election received a hand-written letter by the candidate. For the candidate the purpose of the letter was to introduce herself to the voter, and to invite her to an informal one-to-one meeting. According to the persuasion literature, the letter can serve as a signal that the voter is important to the candidate, that the candidate can be trusted and that she is willing to incur substantial costs to speak to the voter (Lupia and McCubbins 1998; Potter and Gray 2008). The letter emphasized that the candidate wanted to engage with the voters and know their opinions and concerns. It varied whether a business card, which was enclosed to the letter and included the candidate's contact details, contained a non-verbal cue, a photo of the candidate. The literature suggests that physical attributes might serve as cues about candidate quality (McDermott 1997; Banducci et al. 2008). While some see good looks to be beneficial for political careers (Verhulst et al 2010; Geys 2013), others question if they are an advantage to female leaders (Heilman and Stopeck 1985; Eagly 2007). Non-verbal cues might interact with verbal cues to either confirm or contradict stereotypical associations (Heilman and Stopeck 1985).

The results of this experiment show that supporters of rival parties who were randomly assigned to receive a hand-written letter from the Labour candidate were around 8 percentage points more likely to report that they intend to vote for her than those who were assigned to the control group. However the experiment also shows two important limitations of this indirect persuasion strategy: first, it is ineffective in winning over the supporters of the main competitor, the Conservative Party. Party is therefore still a dominant cue and candidates cannot persuade voters at will. Second, the results indicate that positive effects disappear if voters are shown a photograph of the candidate. This might be the case because the physical cue evokes negative associations based on the age and gender of the candidate.

Issue- versus candidate-based persuasion strategies

There are different persuasion strategies that campaigns commonly use to influence voters. Often campaigns try to persuade voters by talking to them about issues (Nickerson and Rogers 2014). Indeed, scholars have argued that partisans can be won over by emphasizing wedge issues (Hillygus and Shields 2008, van der Ward et al 2014). Experimental evidence whether issue-based persuasion succeeds in partisan election campaigns stemming from the United States (Nickerson 2005; Arceneaux 2007; Arceneaux and Nickerson 2009; Bailey et al. 2013) and Italy (Kendall et al. 2014) is mixed; Nickerson (2005), Arceneaux and Nickerson (2009) and Bailey and his colleagues (2013) report null or negative effects of issue-based messages on individual voting intentions. In contrast Arceneaux (2007) and Kendall and his colleagues (2014) report statistically significant, positive effects of issue-based messages on voting intentions at the individual level, and on vote shares at the aggregate level (Kendall et al. 2014). Interestingly the mechanism through which issue-persuasion is thought to work is unclear; Arceneaux (2007) and Lenz (2012) question whether priming issues which are favorable to the campaign has any effect on vote choice. While some scholars are skeptical whether campaigns are able to change political believes and attitudes (Arceneaux 2007; Arceneaux and Kolodny 2009), LaCour and Green (2014) show that high quality canvassing can indeed affect such change. But high quality persuasion of this kind is costly. Canvassers have to be trained to deliver the script smoothly. Effective conversations last for a long period of time (LaCour and Green 2014). Moreover, it has neither been shown that a change in issue positions affects vote choice, nor that partisans change positions if the persuader is affiliated with a rival political party. Furthermore, issue-based persuasion strategies can sometimes result in political dilemmas, especially in multi-party systems comparable to the UK. If a candidate wishes to appeal to supporters of smaller parties, which are based on the opposite ends of the political

spectrum, emphasizing an issue-position that pleases supporters of one party might put-off supporters of another party. This dilemma is very real in a seat such as Southampton Itchen, where Labour could appeal to supporters of smaller parties as diverse as UKIP, the Liberal Democrats or the Green Party. Therefore, a more effective strategy might be appeal to supporters of smaller parties on valence- or on personality grounds.

One important valence characteristic that might be valued by voters is to be a good constituency Member of Parliament (MP) (Norris 1997, Norton and Wood 1993; Heitshusen et al. 2005) and to represent the policy views of constituents (Carman 2006). Constituency service, dealing with casework, is an integral part of MPs' weekly routine. It has been shown that MPs in marginal seats and MPs that have just been elected spend more of their time on constituency service (Cain et al. 1984; Heitshusen et al. 2005). Moreover, British voters prefer their MPs to vote in line with constituency opinion and not in line with their party (Carman 2006). Hence, showing constituents that the candidate is interested in their opinions and that she would like to hear about their concerns might be a strong signal for how she will behave as an MP. Election campaigns can therefore try to affect voting intentions by using candidate-based persuasion strategies to overcome partisan pre-dispositions.

Although the most important signal remains a candidate's party affiliation, there are other competing signals that can be employed by candidates to affect voting intentions. According to Lupia and McCubbins (1998: 54) one particularly powerful signal that has the potential to overcome partisan resistance is "costly and observable effort" on the part of the persuader. Signaling that a candidate is willing to incur substantial costs can compensate for other characteristics and can persuade a voter to give her a hearing: "Actions speak louder than words. Someone who takes a costly action (i.e. exerts effort) reveals something to others about how much a particular outcome is worth to him or her."

(Lupia and McCubbins 1998: 58). As Lupia and McCubbins (1998: 55) elaborate, "as the magnitude of costly effort increases, the extent to which perceived common interests are required decreases." Costly effort might also serve as a cue to the quality of a candidacy (Potter and Gray 2008). If a candidate works hard to win over voters, the candidate might also work hard as an MP. Costly effort might therefore signal that the candidate scores high on valence or performance. This might fit the case of what Mullainathan et al. (2008) call "transference", where an attribute that is useful in one context (the election campaign) is used to infer how the candidate will perform in another context (representation). There are two campaign experiments that have tried to investigate whether costly signals make a difference when campaigns try to persuade voters (Barton et al. 2014; Potter and Gray 2008). Barton and colleagues (2014) find that personal contact with a candidate for local office has a large effect on vote choice. Moreover, Potter and Gray's (2008) results show that high quality signals are more effective than low quality signals in a low salience election.

Individuals have been shown to use non-verbal cues to form opinions about candidates (McDermott 1997; Banducci et al. 2008). One of the most-often discussed cues is physical appearance. Some see good looks as beneficial for political careers (Verhulst et al 2010; Geys 2013). Banducci et al. (2008) show for instance that attractive candidates are more likely to be perceived as winners. However, others question whether looks are an advantage or a disadvantage to female leaders (Heilman and Stopeck 1985; Eagly 2007, Mattes et al. 2010). British tabloids are known for their focus on the looks of female politicians (Mason 2014), and the candidate in question is a young journalist who was described as 'attractive' by the British media on multiple occasions (Daily Star 2012, Southern Daily Echo 2013). The discussion of Davis' looks is no exception to the treatment of female politicians in some of the British press, and might feed into the

stereotypes that some voters might hold against young, female candidates. Physical characteristics might prime attributes such as age and gender, which might in turn activate biases. The Southern Daily Echo (2013) used Davis' looks for instance to associate her with out-of town candidates that are championed by party leaders: "In the past we've had Blair's Babes and Cameron's Cuties, so I guess the attractive Ms Davis will be one of Mili's Fillies" (Southern Daily Echo 2013). The direction of the effect of physical appearance on voting intentions is thus unclear.

Experimental Design

The original experimental design had two stages. In the first stage it was randomly assigned whether a constituent who was on record to support a party other than Labour received a hand-written letter from the Labour candidate introducing herself and offering to meet the constituent for tea. The second stage consisted of a meeting between the candidate and the constituent. In between stages one and two, and again after the meeting had taken place, a telephone survey was arranged to interview the subjects in the treatment and the control groups about their candidate's evaluations and voting intentions. Moreover, turnout in the 2014 European and local Council elections which took place one month after the treatment was administered was measured using official records from the publicly-available electoral register.

Based on anonymous data included in the party's voter contact database, I selected my experimental sample of 597 households. There were three conditions that needed to be satisfied for a household to be included in the experimental sample. First, I restricted the sample to those households that had an available landline number. This restriction was important because subjects were interviewed by telephone in the month following the treatment in order to record their voting intentions and candidate evaluations (see the Post-

Treatment Survey in the Appendix). Second, I restricted the sample to those households that included information about household members' partisanship. This information was based on the local party's voter canvassing of constituents over the past years. As I have shown in the Birmingham experiment, voter ID data collected through canvassing is generally a valid and reliable, albeit a slightly noisy measure of partisanship in the United Kingdom. The post-treatment survey validates the partisanship of the respondent and shows a high correspondence between the partisanship recorded by canvassers and the answer to the party identification question that respondents gave to the post-treatment survey.[1] I then excluded all households that included a Labour partisan, or about whom there was no data on past voting behavior. This left me with supporters of four different parties: The Conservative Party who came very close to winning the constituency in the 2010 General election, the Liberal Democrats, The Green Party and the UK Independence Party. Moreover, the experimental sample includes those voters who would only volunteer that they are 'against Labour' and those who are currently undecided or do not volunteer their partisanship but are on record to have supported parties other than Labour in the past. In a second step, I randomly chose one person per household to be the experimental subject. I then stratified the sample based on the experimental subject's partisanship. This left me with seven different partisan blocks. Within these partisan blocks, I randomly assigned subjects with a probability of around .5 (depending on whether the numbers in each partisan block were odd or even) to either receive a letter or not to receive a letter

[1]Five subjects in the control group, and seven subjects in the treatment group answered that they were Labour identifiers to the last question on the post-treatment questionnaire. This difference is not statistically significant. Moreover, we cannot exclude that the treatment affected the answer of subjects to the party identification item. Since the vote intention question was asked before the party identification item in order to avoid biasing the results, subjects might have also brought their party identification in line with their vote intention. Nevertheless, since the pre- and post treatment party id classifications were based on different measures (canvassing data versus the BES question), differences could have occurred due to how the questions were formulated or due to how the data was gathered (Labour door-to-door canvassing versus unaffiliated telephone survey).

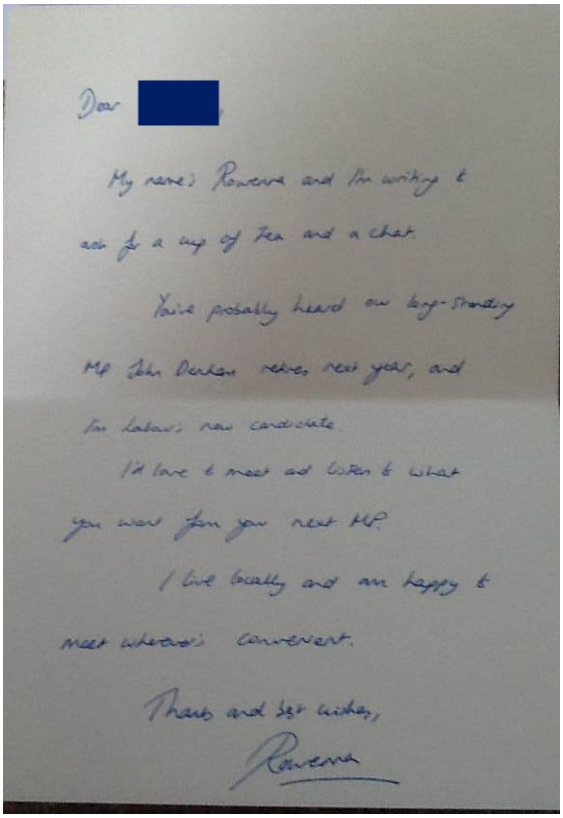
from the candidate. The subjects who were assigned to treatment were then reassigned using a 2x2 factorial design. Factor 1 varies the frame of the invitation used in the letter and factor 2 varied whether a business card, which was enclosed to all letters, included a colored photograph of the candidate. The random assignment therefore left me with 4 different combinations of letters and business cards. The treatment matrix is displayed in Table 1.

Table 1 Treatment Matrix

Treatments	Listening Frame	Discussion Frame	Control
Business card with photo	76	76	x
Business card without photo	76	76	x
Control	x	x	293

The different letters and business cards that were sent to respondents are shown in Figures 1-4. The letters in Figures 1 and 2 differ because they frame the proposed meeting differently. The first purpose is “to listen and see what you want from an MP” while the second purpose is “to discuss how we can build on our cities strength together”. Apart from this difference the letters had the same number of paragraphs and were written with the same ink and on the same plain, white paper. The letters were also sent on the same day and in similar envelopes.

Figure 1 'Listening' frame



Dear ,

My name is Rowenna and I'm writing to ask for a cup of tea and a chat.

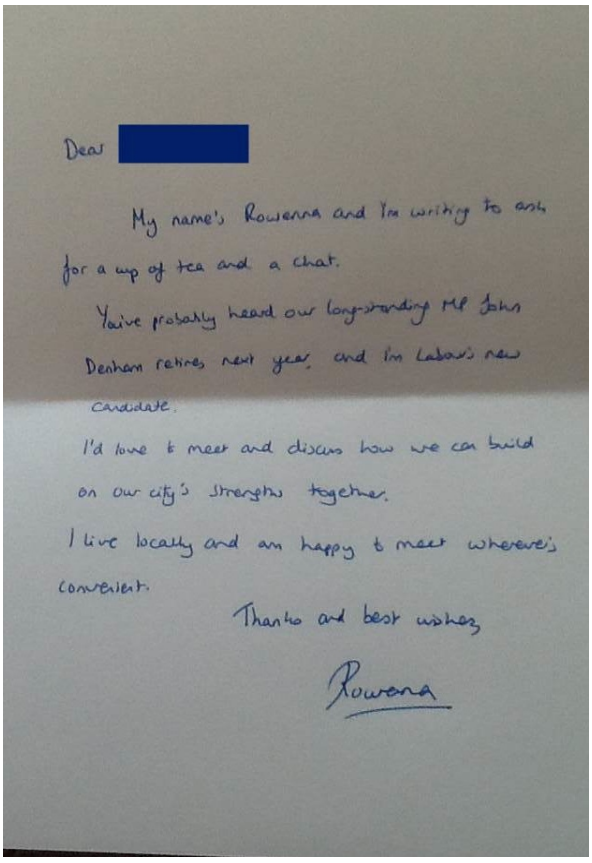
You've probably heard our long-standing MP John Denham retiring next year, and I'm the new Labour candidate.

I'd love to meet and listen to what you want from your next MP.

I live locally and am happy to meet whenever convenient.

Thanks and best wishes,
Rowenna

Figure 2 'Discussion' frame



Dear ,

My name is Rowenna and I'm writing to ask for a cup of tea and a chat.

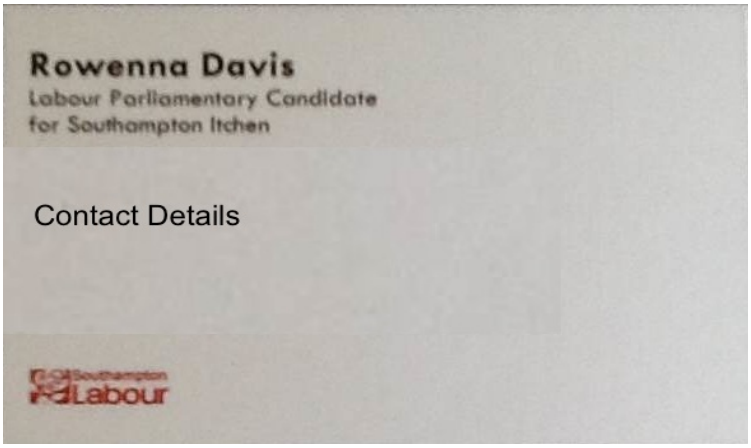
You've probably heard our long-standing MP John Denham retiring next year, and I'm the new Labour candidate.

I'd love to meet and discuss how we can build on our city's strength together.

I live locally and am happy to meet whenever convenient.

Thanks and best wishes,
Rowenna

Figure 3 Business card without photograph



Contact details have been removed.

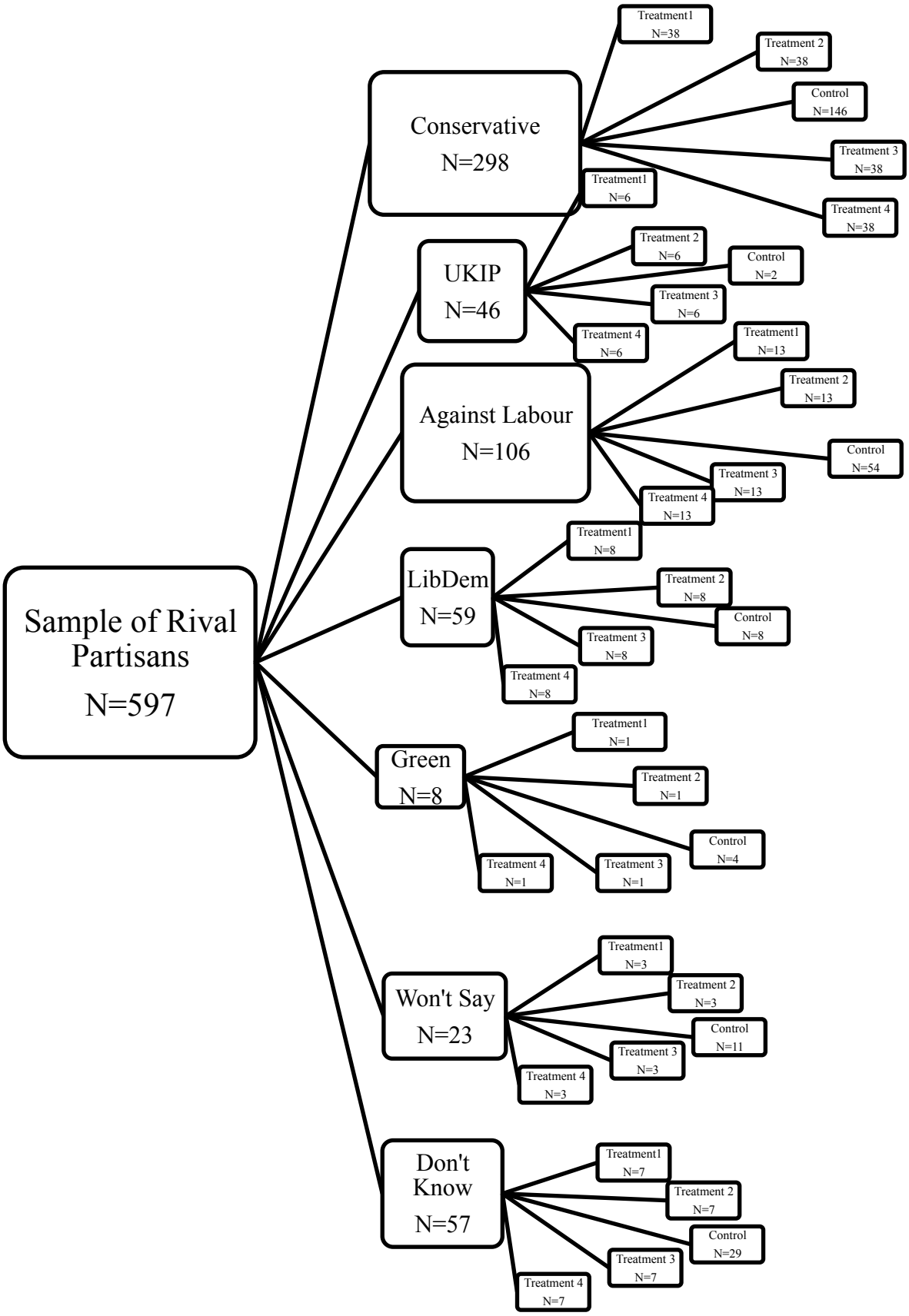
Figure 4 Business card with photograph



Contact details have been removed.

Figure 5 on the next page displays the random assignment to the control group and the four treatment groups within partisan blocks.

Figure 5 Blocked random assignment of experimental subjects to control and 4 treatment groups



Balance Check

Although it is impossible to conclusively prove that random assignment to treatment and control groups has indeed been successful, I use randomization inference to show that, as a result of random assignment, treatment assignment is unlikely to be systematically related to pre-treatment covariates. I include all available pre-treatment covariates: gender, age and electoral ward.

Following the standard randomization inference procedure outlined at length in Ho and Imai (2006) and in Gerber and Green (2012), I test whether I can reject the sharp null hypothesis that any existing covariate imbalances between the experimental groups could have occurred due to sampling variability alone. To do so I first extract the log likelihood resulting from a multinomial logit regression of treatment assignment on pre-treatment covariates. I then compare the extracted log likelihood to the mean of the log likelihoods that I obtain after simulating the random assignment of the pre-treatment covariates within each partisan block 10.000 times. The collection of log likelihoods "represents the exact sampling distribution under the null hypothesis that no covariates have any effect on the assigned treatment" (Gerber and Green 2012: 107). The resulting p-value of .55 indicates that we cannot reject the sharp null hypothesis that the covariates taken together are not systematically related to treatment assignment.

Missing Data

One of the big challenges facing researchers who study persuasion is sample attrition (Bailey et al. 2013). Studying voter persuasion at the individual level poses more challenges to the outcome data than studying the effects of election campaigns on voter turnout because we cannot rely on public records to provide outcome data on vote choice. I therefore have to rely on a telephone survey that was conducted by volunteers following

the administration of the treatments in May 2014. The survey was fielded about three weeks after the letters were sent on April 15th and was completed two days before the 2014 European and local elections on May 22nd. Callers managed to reach around 55% (327 subjects) of our sample, and around 21% (120 subjects) agreed to answer the survey. This response rate is comparable to other telephone surveys. Nickerson (2007) reports a response rate of 25.5%, while Barton et al. (2014) report a response rate of 13% to the vote choice question.

Of course sample attrition poses serious challenges for the identification of the treatment effect. First, it has been shown that some individuals become less likely to answer post-treatment surveys after they have been exposed to direct persuasion treatments (Bailey et al. 2013). Therefore, the persuasion treatments fielded in this study were specifically designed to be gentle, and not to put constituents off. In order to test if attrition is a function of treatment assignment, I regress a variable indicating if outcome data is missing on treatment assignment and extract the resulting F-statistic. In the next step I simulate blocked random assignment to control and treatment groups 10000 times and compare the resulting mean of all f-statistics to the f-statistic extracted from my realized data set. The P-value of .66 indicates that we cannot reject that attrition is unrelated to treatment assignment. However, that does not mean that attrition occurred at random, or independent of Potential Outcomes. Unfortunately, there is no credible way of checking that those who did not reply to the post-treatment survey would react in the same way to treatment than those who replied to the post-treatment survey. Strictly speaking, we can therefore only make inference about those subjects who reply to telephone surveys. But this limitation applies to all survey research that suffers from attrition and is not a problem specific to persuasion experiments. Moreover, we can check if attrition is systematically related to observable pre-treatment covariates. If we could not reject that attrition occurred

independently of observable covariates, we could be slightly more confident that our findings might apply to the entire experimental sample. I therefore use randomization inference to test if pre-treatment covariates are more predictive of sample attrition than would be expected from sampling variability alone. The resulting p-value of .50 indicates that I cannot reject that sample attrition occurred independently of observable background attributes.

Results

Table 2 lists the replies that the candidate received from constituents. In total, she only received four replies referring to her invitation to arrange a meeting. In other words, only 1.3 percent of constituents who usually support other parties replied to a personal letter addressed to them by the Labour Party Candidate. Moreover, only two out of 304 constituents who received a letter took up her offer to arrange a meeting. This result might be surprising given that the seat has been held by Labour seat since 1997, and that the Labour candidate might well be the next constituency MP.

Both meetings were held within a couple of days and both resulted in a positive outcome, according to the candidate. Both constituents pledged to support her and offered to volunteer on her campaign. Unfortunately, the small number of answers makes it impossible to analyze the replies quantitatively. However, the small number of replies is an interesting result in its own right. Before the experiment, the candidate was optimistic that many voters would take her up on the offer for a meeting even if they did not support her party. The results of this experiment underline that this is clearly not the case, and that outreach across the partisan divide is difficult.

Still, Table 2 can give a first idea about how constituents responded to the letter.

Table 2 Responses to the letter

Reply	Treatment	Partisanship	Tone	Medium	Meeting	Reasons
1	Listen x Photo	Green	Positive	Telephone	Yes	x
2	Listen x Photo	Conservative	Negative	Letter	No	x
3	Discussion x Photo	UKIP	Positive	Facebook	Yes	x
4	Discussion x No Photo	Conservative	Positive	Letter	No	Health

First, both constituents who requested a meeting are supporters of smaller parties, the Green Party and UKIP. The only negative reply that the candidate received, on the other hand, came from a Conservative supporter. Campaign volunteers were instructed to code a reply as positive in tone if the constituent explicitly thanked the candidate for the letter and expressed that she was pleased to receive it.

Analysis

There are two explanations for why subjects did not reply to the letter. The first explanation relates to a failure to perceive the message (Zaller 1992). It is conceivable that the letter was too subtle an intervention for constituents to notice, that letters were thrown away or that they forgot about the letter. The second explanation assumes that constituents recalled that contact was made, but that they did not want to meet the candidate. If the former were the case, we would not expect any effects of the letter on subjects' attitudes towards the candidate. The latter, however, opens the possibility that letters influenced subjects even if meetings did not occur on a scale large enough to affect the outcome of this experiment.

To shed light on these two possibilities, I test whether individuals who were assigned to receive a letter are more likely to recall being contacted by the Labour candidate than individuals who were not assigned to receive a letter. Although I know that all letters were posted to the addresses included in the database, I have of course no

possibility of knowing whether the subject opened and read the letter. Moreover, I am interested in whether the letter left a lasting impression on subjects. In fact this experiment is a difficult test of persuasion effects because a substantial period of time passed between treatment administration and the telephone survey, which amounted to a minimum of 26 and a maximum of 61 days. To assure the integrity of the experiment, volunteer callers followed the original list provided by the campaign. They were blind to the assignment of the experimental groups.

In order to minimize the association that subjects could make between the experiment and the consecutive telephone survey, callers only asked whether subjects could recall having been contacted by the Labour candidate. They did not mention any hand-written letters. I would therefore expect that there are substantial numbers of subjects who have received other type of campaign communication by the candidate or the Labour party. Nevertheless, thanks to random assignment we can identify the Intent-To-Treat Effect (ITT) on reporting contact with the Labour candidate. The ITT is defined as: $ITT = \frac{1}{N} \sum_{i=1}^{N} (Y_i(z = 1) - Y_i\,(z = 0))$, where $Y_i(z = 1)$ is the potential outcome for subject *i* if she is assigned to treatment, and $Y_i(z = 0)$ is the potential outcome for subject *i* if she is assigned to control.

27% report being contacted by the Labour candidate in the control group and 43% report being contacted in the treatment group. The difference in proportions between those who report contact in the treatment group and those who report contact in the control group is therefore around 16 percentage points. Due to the small sample size of 120 respondents in the post-treatment survey, I use randomization inference in order to estimate p-values and confidence intervals. Given small sample sizes, conventional methods such as t-test and regression analysis make unrealistic assumptions about the shape of the sampling distribution, which only approximates a normal distribution when *n* is large. Compared to

conventional methods, randomization-based inference is based on the realization that we can in fact know the exact sampling distribution of any test-statistic under the null hypothesis that the treatment has no effect on any unit (the sharp null-hypothesis). We know that the randomly assigned data in front of us is only one possible realization of the experimental data because random assignment could have given us an (almost) infinite number of datasets with different combinations of units in treatment and control groups. We create the exact sampling distribution by re-running the random assignment process within partisan blocks 10.000 times and compare the test-statistic in our data set to the mean of the test-statistics that we receive if the treatment effect is zero for every unit. In this case, the treated Potential Outcome $Y_i(1)$ is the same as the untreated Potential Outcome $Y_i(0)$. Using randomization inference I test the sharp null hypothesis and I find that this difference in proportions is statistically significant with a two-tailed p-value of .05. This means that subjects who were assigned to receive the letter are significantly more likely to recall contact with the Labour candidate than subjects who were assigned to the control group.

Although the candidate only received four replies, the higher recall of contact in the treatment group provides the basis to believe that the letters might have affected attitudes towards the candidate and electoral behavior independently of meeting the candidate. Table 3 looks systematically at the effect of the letter on voting intentions for the Labour candidate and for her main competitor, the Conservative candidate. Subjects were asked which candidate they indented to vote for in the 2015 General Elections. Respondents were given a list of declared candidates in the seat to choose from. Answers to the voting intention question were coded 1 if the respondent chose the respective candidate, and 0 if she chose any other option (including undecided and don't know). The specific question can be found in the Appendix.

Table 3 ITT Effects of Letter on Voting Intentions (binary) and Candidate Evaluations (5-points scale)

	Voting Intentions		Candidate Evaluations	
	Davis (Lab)	Smith (Con)	Davis (Lab)	Smith (Con)
Letter	.082†	-.096	.222	-.130
95%-Confidence Interval	[-.053, .169]	[-.236, .125]	[-.163, .616]	[-.593, .330]

Note: † $p < 0.10$, *$p < 0.05$ using 2-tailed test of sharp null hypothesis.

The voting intention column in Table 3 shows the ITT effect of being assigned to receive a letter on voting intentions for the Labour candidate Davis and the Conservative candidate Smith, while the column on the right shows the ITT effect of the treatment on how respondents evaluate the two candidates. P-values and 95% confidence intervals are estimated using the standard randomization inference procedure outlined in Gerber and Green (2012).

The ITT effect of receiving any type of letter on intentions to vote for the Labour candidate amounts to around 8 percentage-points, while the treatment has a negative effect of around 10 percentage-points on the challenger. The effect on the Labour vote is statistically significant with a p-value of .05 using a two-sided hypothesis test, while the treatment effect on the Conservative vote is not statistically significant by conventional standards (p=.12 using a two-sided test). That means that out-partisans who were assigned to receive one of the personalized, hand-written letters were around 8 percentage-points more likely to report that they intend to vote for the Labour candidate. Given the relatively small sample size of 110 respondents, the 95% confidence interval bracketing this marginal effect ranges from -5 to 17 percentage-points.

Answers to the candidate evaluation questions were coded on a 4-points scale ranging from strongly positive (1) to strongly negative (4). Respondents were also given the opportunity to volunteer that they did not know who the candidate was or where to

place her or him. The scale was recoded ranging from strongly negative (0) to strongly positive (4), with 'don't knows' imputed at the mean of the scale (2).

The ITT effect of treatment assignment on the evaluations of the Labour candidate is positive, and amounts to .22 on the 5-points candidate evaluation scale. However, the treatment effect is only statistically different from zero with a p-value of .27 using a two-sided test of the sharp null hypothesis. In contrast, the ITT effect of treatment assignment on the evaluations of the main challenger, the Conservative candidate, is negative, and amounts to -.13 on the 5-points candidate evaluation scale. For both Davis and Smith, I cannot reject the sharp null hypothesis that there is no effect for any subject. While there is some evidence that the treatment affected how subjects evaluate both candidates, the observed effects are not large enough, given the small sample size, to reach statistical significance.

Table 4 includes two 2x2 tables that present the disaggregated effects of framing and appearance cues on Labour voting intentions and candidate evaluations, compared to the control group. Column 1 displays a linear probability model regressing voting intentions on the treatment matrix, including inverse-probability-weights, and column 4 displays effects on candidate evaluations including inverse-probability weights. The letters varied whether a colored photograph is printed on the business card that was enclosed to the letter and whether the purpose of the proposed meeting is given as 'listening to' or as 'discussing with' the constituent.

No clear pattern emerges from the table, and these results do not suggest that adding the photo to the business card had any clear negative or positive effect. The most effective letters are those that either combine the photo with the discussion frame, or the listening frame and the blank business card. I did not expect to find this interaction.

Table 4 ITT Effects of Photo and Letter Content on Voting Intentions (binary) and Candidate Evaluations (5-points scale)

	Davis (Lab) Vote Intention		Davis (Lab) Candidate Evaluation	
Treatments	Listening Frame	Discussion Frame	Listening Frame	Discussion Frame
Business card with photo	-.047 (.129)	.289* (.121)	-.041 (.407)	.387 (.385)
Business card without photo	.197† (.112)	.016 (.110)	.826* (.352)	.161 (.342)

Note: †p < 0.10, *p < 0.05 using 2-tailed t-test.

The most plausible explanation for why the discussion frame is more effective with the photo, and the listening frame is more effective without the photo, relates to the choice of photo, which shows the candidate speaking in a talk show. It shows her talking, not listening. This interpretation also fits the pattern of results displayed in columns three and four. Again, the listening frame is most effective when it is not accompanied by a photo of the candidate that shows her talking. These results might suggest that constituents value consistency. If a candidate proposes to listen, her actions should match her words.

Moreover, I check whether the effect of the treatment on vote choice and candidate evaluations is conditioned by the pre-treatment partisanship of the respondent. Recall that I expect supporters of smaller parties to be more persuadable than supporters of the major rival party, the Conservative Party. Supporters of the Conservative Party should be more resistant to Labour persuasion efforts than Liberal Democrats, Greens and supporters of UKIP because the constituency is a Labour-Conservative marginal. Moreover, with the exception of the New Labour landslide in 1997, instances of Conservative to Labour conversion are rare in the British partisan universe (Zuckerman 2007). I therefore estimate the interaction effect between treatment assignment and partisanship, where partisanship is coded 1 if the respondent is a Conservative and 0 if she is a supporter of another party or if

she is undecided, but has previously supported another party. Formally, we can write the linear probability model, in which Y_i is either voting intention or candidate evaluation, Z_i indicates that a subject was assigned to receive the treatment, and B is a k-by-1 vector of coefficients, as:

$$Y_i = \beta_0 + \beta_1 Z_i + \beta_2 Conservative_i + \beta_3 (Z * Conservative)_i + B + \mu_i$$

Table 5 shows that there is a statistically significant interaction between treatment assignment and Conservative partisanship for voting intentions.[2] While the treatment has a statistically significant and positive effect on Labour voting intentions for supporters of smaller parties, Conservative supporters are significantly less likely to change their voting intentions as a response to the treatment. In fact, the effect of the treatment on voting intentions for Conservative supporters is slightly negative, but not statistically significant. A broadly similar pattern can be observed for candidate evaluations, although the interaction effect between the treatment and Conservative partisanship does not reach statistical significance.

Table 5 ITT Effects of Treatment on Voting Intentions (Labour/Other) and Candidate Evaluations (5-points scale), conditional on partisanship

	Voting Intention	Candidate Evaluation
Letter	.185* (.075)	.365 (.278)
Conservative	-.045 (.077)	-.379* (.289)
Letter x Conservative	-.218* (.109)	-.309 (.408)
N	110	116

Note: Standard errors in parentheses, † $p < 0.10$, * $p < 0.05$ using 2-tailed t-test.

[2] The model is robust to removing the small number of respondents who are thought to be undecided.

Finally, I check if the letters backfired, defined as mobilizing Conservative supporters without persuading them. Table 6 shows whether the treatment affected interest in the campaign (column 1), turnout intentions (column 2), and validated turnout (column 3). While validated turnout is available for the entire sample, campaign interest and turnout intentions are only available for the subsample of subjects that responded to the telephone survey. Interest in politics and turnout intentions for the Council and European elections are measured on a 4-points scale ranging from 1 (not at all interest/very unlikely to vote) to 4 (highly interested/very likely to vote). Interest is recoded to range from 0 to 3, and turnout intentions are recoded into a binary variable (very likely to vote versus everything else) to allow for a more intuitive comparison with the validated vote category.

Since we cannot know if subjects opened the envelope and read the letter, all treatment effects are Intent-To-Treat effects and will hence underestimate the true effect of the treatment on those subjects who actually read the letter. Table 6 displays the results from four linear probability models that regress the outcome variable on the binary treatment indicator (letter or no letter), Conservative partisanship and the interaction between partisanship and the treatment, weighted by the inverse of the probability of being assigned to treatment within each partisan group.

Table 6 ITT Effects of Letter on engagement, conditional on partisanship

	Interest (Survey Respondents)		**Turnout Intention** (Survey Respondents)		**Validated Turnout** (Full Sample)	
Letter	.104	-.273	.046	-.089	-.016	-.022
	(.166)	(.223)	(.079)	(.107)	(.040)	(.057)
Conservative	-	-.398	-	-.194†	-	-.012
		(.231)		(.111)		(.057)
Letter x Conservative	-	.808*	-	.290†	-	.013
		(.326)		(.157)		(.080)
N	120		120		597	

Note: † $p < 0.10$, *$p < 0.05$ using 2-tailed test of sharp null hypothesis.

Table 6 shows that Conservative supporters respond differently to receiving the invitation letter than supporters of smaller parties. Supporters of smaller parties are slightly less likely to report interest in the campaign and are less likely to indicate that they are very likely to vote in the upcoming elections. However, we cannot reject the null hypothesis that these estimates are different from zero in statistical terms. In contrast, it looks like Conservative supporters are fired up when they are contacted by a Labour candidate. Their interest in politics increases, and they are more likely than supporters of other parties to report that they are very likely to vote as a result of the campaign intervention. However, these results did not translate into higher electoral participation in the Council and European elections. While the interaction effect points into the same direction, it is not statistically significant, and Conservative supporters are no more likely to turn out as a result of the treatment. Thus, while it looks like the treatment increases psychological engagement, it has no effects on turnout.

At the same time the difference between psychological engagement and actual behavior might only be part of the explanation. Table 7 allows us to compare whether subjects who answered the telephone survey differ on unobserved characteristics from subjects who did not answer the phone or refused to complete the telephone survey. Recall that observed differences in age, gender and place of residence did not predict whether a subject answered the survey.

Table 7 shows that the subjects who answered the telephone survey were more responsive to the treatment than the subjects who did not answer the telephone survey. Although none of the effects reaches statistical significance in the sample of survey respondents, the direction and size of effects clearly matches the data on political interest and turnout intentions. Nevertheless, the effects between survey respondents and non-respondents are not significantly different.

Table 7 ITT Effects of Letter on engagement, conditional on partisanship

	Validated Turnout (Survey Respondents)		**Validated Turnout** (Non-respondents)	
Letter	-.060	-.136	-.007	.002
	(.092)	(.124)	(.044)	(.063)
Conservative	-	-.266*	-	.056
		(.129)		(.062)
Letter x Conservative	-	.163	-	-.018
		(.182)		(.088)
N	120		477	

Note: † $p < 0.10$, *$p < 0.05$ using 2-tailed test of sharp null hypothesis.

These findings, nevertheless, raise questions about our ability to extrapolate from the subsample of survey respondents to the initial experimental sample. It is likely that answering the post-treatment survey is a function of political interest or engagement in the campaign in general. It might hence be the case that the positive effects on voting intentions that I found in the telephone sample only apply to the subset of highly engaged subjects. Judging from the results on turnout, it is however unlikely that survey non-respondents would have responded negatively to the persuasion attempt. It is more likely that they would have been unaffected. This study clearly points to some of the problems that field researchers face when studying persuasion. We are likely to run into the problem of survey non-response and might hence exaggerate the inference that can be drawn from the intervention. Alternatively, researchers might choose to collect outcomes at the aggregate level. However, working at the aggregate level also has its pitfalls when it comes to studying mechanisms. Pons (2014) is for instance unable to distinguish between persuasion and (de-) mobilization effects since the recorded increases in the vote share of one party could result from changes to the composition of the electorate or from conversion. The best option is probably to combine outcome measurement at both the

individual and the aggregate level (Kendall et al. 2014). This solution, while it detects potential problems, also does not give us an answer about what to do when individual and aggregate evidence diverges. The use of other methods such as survey experiments, or field experiments combined with Internet panels (LaCour and Green 2014) to study persuasion might reduce the non-response rate, but it is doubtful if respondents who self-select into participating in these panels are more representative of the population of interest than those who answer telephone surveys.[3]

Discussion and conclusion

This paper reports the results from the first persuasion experiments conducted during a real world election campaign in the United Kingdom, and it might hence be too early to draw confident conclusions from this research. However, there are some interesting findings that ask for elaboration. Keeping in mind the limitations that researchers face when studying persuasion effects including non-response bias and a reliance on self-reported behavior, the results of the post-treatment survey suggest that hand-written letters had a positive effect on voting intentions. Although we might question whether this result is externally valid, meaning that it is doubtful whether we can extrapolate the findings to the entire experimental sample, it still contributes to our understanding of out-partisan persuasion. It is striking that letters inviting subjects to a meeting seem to have affected the voting intentions of some subjects even if they never responded to the invitation, and never met the candidate.

This finding suggests that there might be more subtle forms of political persuasion than hands-on efforts that challenge out-party voters to support a specific policy or party they initially dislike. As the old saying goes, "if all you have is a hammer, every problem

[3] The problem of survey non-response bias applies more generally to research that relies on outcome measures that cannot be collected from administrative records.

looks like a nail." There is ample evidence that the hammer has not been very successful when it comes to persuasion in the past (Arceneaux and Kolodny 2009; Bailey et al. 2013). Maybe subtle signals of good will that inspire confidence in a candidate and that contradict the partisan stereotypes that voters might hold dear (Green et al. 2002), are more promising strategies of reaching out across the partisan divide. This suggestion might apply especially to candidates that are faced with the strategic dilemma of winning over voters that are situated at the opposing ends of the political spectrum. One should not underestimate the difficulty of appealing to Green Party Voters, Liberal Democrats and UKIP supporters at the same time. An issue-based persuasion strategy might further complicate the strategic dilemma in which candidates find themselves. Instead, candidates can look for non-intrusive ways of communication to voters who are initially unlikely to support their candidacy.

The results from this experiment are also consistent with research on bounded partisanship that shows that partisans are usually unwilling to support the main opposition party, but more likely to switch from smaller to larger parties and back (Zuckerman 2007; Neundorf et al. 2011). The reluctance to cross the partisan divide might have different explanations ranging from strategic considerations, over partisan identities to explanations based on ideological proximity. The implications for the conduct of campaigns, however, are clear. Campaigns might be more successful in focusing their persuasion efforts on the supporters of smaller parties than on the supporters of the main rival party.

I will leave the reader with a few methodological considerations. Drawing on the experience from this experiment and other recent attempts to implement a two-stage experimental design that requires subjects to self-select into the second stage (see Sokhey and Mann 2014), first-stage treatments need to be sufficiently powerful. Although I hypothesized that a hand-written letter would be a powerful incentive to personally meet a

relatively high profile out-party candidate, this expectation was not validated. While we know a lot about how to turn out voters, we know very little about how to mobilize other forms of political participation such as attending a meeting. This needs to change. Moreover, since we are unlikely to find powerful first-stage treatments that do not violate the exclusion restriction (since they are powerful, they are likely to affect the outcome independently of participation in the second stage), we should think deeply about ways to account for this violation. One possibility of (partially) accounting for it would be to either measure outcomes repeatedly before and after the second stage, to randomly assign a similar letter without an invitation, or a combination of both strategies. Both strategies rely on assumptions, but they might be more defensible than the excludability assumption.

References

Aldrich, John H. 1995. *Why Parties? The Origin and Transformation of Party Politics in America.* Chicago: University of Chicago Press.

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* 2: 43-65.

Arceneaux, Kevin, and Robin Kolodny. 2009. "Educating the Least Informed: Group Endorsements in a Grassroots Campaign." *American Journal of Political Science* 53(4): 755–770.

Aronow, Peter M., and Cyrus Samii. 2012. ri: The R package for performing randomization-based inference for experiments. R package version 0.9. http://CRAN.R-project.org/package=ri

Bailey, Michael A., Daniel J. Hopkins, and Todd Rogers. 2013. "Unresponsive and Unpersuaded: The Unintended Consequences of Voter Persuasion Efforts." Available at SSRN: http://ssrn.com/abstract=2307631.

Banducci, Susan A. Jeffrey A. Karp, Michael Thrasher and Colin Rallings. 2008. "Ballot Photographs as Cues in Low-Information Elections." *Political Psychology* 29(6): 903-917.

Bartels, Larry M. (1998). Where The Ducks Are: Voting Power in a Party System. In Geer, John G. (ed). *Politicians and Party Politics.* Baltimore: John Hopkins University Press: pp. 43-79.

Barton, Jared, Marco Castillo and Ragan Petrie. 2014. "What Persuades Voters? A Field Experiment on Political Campaigning." *The Economics Journal* 124 (1): 293–326.

Cain, Bruce E., John A. Ferejohn, and Morris P. Fiorina. 1984. "The constituency service basis of the personal vote for US representatives and British members of parliament."

The American Political Science Review 78(1): 110-125.

Campbell, Rosie and Philipp Cowley. 2013. "What Voters Want: Reactions to Candidate Characteristics in a Survey Experiment." *Political Studies*. Published Online: doi: 10.1111/1467-9248.12048

Carman, Christopher Jan. 2006. "Public preferences for parliamentary representation in the UK: an overlooked link?." *Political Studies* 54(1): 103-122.

Conover, Pamela J. 1981. "Political Cues and the Perception of Candidates." *American Politics Research* 9(4): 427–448.

Conover, Pamela Johnston, and Stanley Feldman. 1989. "Candidate perception in an ambiguous world: Campaigns, cues, and inference processes." *American Journal of Political Science* 33(4): 912-940.

Cowley, Philip. 2013. "Why not ask the audience? Understanding the public's representational priorities". *British Politics* 8(2): 138–163.

Dewan, Torun, Macartan Humphreys and Daniel Rubenson. 2014. "The Elements of Political Persuasion: Content, Charisma and Cue." *The Economics Journal* 124 (1): 257–292.

Daily Star. 2012. Guido Fawkes: Plots, rumours and conspiracy. 10th September 2012, http://www.dailystar.co.uk/news/latest-news/271592/Guido-Fawkes-Plots-rumours-and-conspiracy

Eagly, Alice H. 2007. "Female leadership advantage and disadvantage: resolving the contradictions." *Psychology of women quarterly* 31(1):1-12.

Ford, Robert, and Matthew J. Goodwin. 2014: *Revolt on the Right: Explaining Support for the Radical Right in Britain*. London: Routledge.

Gerber, Alan and Donald P. Green. 2012. *Field Experiments: Design, Analysis, and Interpretation.* WW: Norton.

Geys, Benny. "Looks Good, You're Hired? Evidence from Extra-Parliamentary Activities of German Parliamentarians." 2013. *German Economic Review*, published online.

Green, Donald P., Bradley Palmquist, and Eric Schickler. 2002. *Partisan Hearts and Minds: Political Parties and the Social Identities of Voters.* New Haven: Yale University Press.

Heilman, Madeline E., and Melanie H. Stopeck. 1985. "Being attractive, advantage or disadvantage? Performance-based evaluations and recommended personnel actions as a function of appearance, sex, and job type." *Organizational Behavior and Human Decision Processes* 35(2): 202-215.

Heitshusen, Valerie, Garry Young, and David M. Wood. 2005. "Electoral Context and MP Constituency Focus in Australia, Canada, Ireland, New Zealand, and the United Kingdom." *American Journal of Political Science* 49(1): 32-45.

Ho, Daniel E. and Kosuke Imai. 2006. "Randomization Inference with Natural Experiments: An Analysis of Ballot Effects in the 2003 California Recall Election." *Journal of the American Statistical Association* 101(475): 888–900.

Iyengar, Shanto and Adam F. Simon. 2000. "New Perspectives and Evidence on Political Communication and Campaign Effects". *Annual Review of Psychology* 51:149-169.

Jones, Philipp and John Hudson. 1998. "The role of political parties: An analysis based on transaction costs." *Public Choice* 94: 175–189.

Kendall, Chad, Tomasso Nannicini and Francesco Trebi. 2014. "How Do Voters Respond to Information? Evidence from a Randomized Campaign." Working paper.

LaCour, Michael J., and Donald P. Green. 2014. "Messages, Messengers, and Diffusion of Support for Gay Equality: Results from Two Longitudinal Field Experiments." *Working paper.*

Lavine, Howard G., Christopher D. Johnston, and Marco R. Steenbergen. 2012. *The ambivalent partisan: How critical loyalty promotes democracy*. Oxford: Oxford University Press.

Lenz, Gabriel S. 2012. *Follow the leader?: How voters respond to politicians' policies and performance*. University of Chicago Press.

Lupia, Arthur and McCubbins Matthew. 1998. *The Democratic Dilemma*. Cambridge: Cambridge University Press.

Mason, Rowena. 2014. Daily Mail's 'Downing Street catwalk' condemned by MPs. The Guardian, Wednesday 16 July 2014: http://www.theguardian.com/media/2014/jul/16/daily-mail-downing-street-catwalk-female-ministers-outrage

Mattes, Kyle, Michael Spezio, Hackjin Kim, Alexander Todorov, Ralph Adolphs and R. Michael Alvarez. 2010. "Predicting election outcomes from positive and negative trait assessments of candidate images." *Political Psychology* 31(1): 41-58.

McDermott, Monika L. 1997. "Voting cues in low-information elections: Candidate gender as a social information variable in contemporary United States elections." *American Journal of Political Science* 41(1): 270-283.

Mullainathan, Sendhil, Joshua Schwarzstein and Andrei Shleifer. 2008. "Coarse thinking and persuasion." *The Quarterly Journal of Economics 123(2): 577-619.*

Nickerson, David W. and Todd Rogers. 2014. "Political Campaigns and Big Data." *The Journal of Economic Perspectives* 28(2): 51-73.

Nickerson, David W. 2007: "Quality Is Job One: Professional and Volunteer Voter Mobilization Calls," *American Journal of Political Science* 51(2): 269–282.

Neundorf, Anja, Daniel Stegmueller, and Thomas J. Scotto. "The individual-level dynamics of bounded partisanship." *Public opinion quarterly* 75.3 (2011): 458-482.

Norris, Pippa. 1997. "The puzzle of constituency service." *The Journal of Legislative Studies* 3(2): 29-49.

Norton, Philip, and David Michael Wood. 1993. *Back from Westminster: British Members of Parliament and their constituents*. Lexington, KY: University Press of Kentucky.

Page, Benjamin I., and Robert Y. Shapiro. 1992. *The rational public: Fifty years of trends in Americans' policy preferences*. Chicago: University of Chicago Press.

Potter, Philip B.K. and Julia Gray. 2008. "Does Costly Signaling Matter? Preliminary Evidence from a Field Experiment." BCSIA Discussion Paper 2008-09, Kennedy School of Government, Harvard University.

Pons, Vincent. 2013. Does Door to door canvassing affect vote shares: Evidence from a Countrywide Field Experiment in France. Working paper: http://economics.mit.edu/files/9000.

Rahn, Wendy M. 1993. "The Role of Partisan Stereotypes in Information Processing about Political Candidates". *American Journal of Political Science* 37(2): 472-496.

Snyder, James M. Jr., and Michael M. Ting. 2002. "An Informational Rationale for Political Parties." *American Journal of Political Science* 46(1): 90-110.

Sokhey, Anand Edward and Christoper B. Mann. 2014. Examining the Spread of Political Discussion: Results from a Field Experiment. Paper presented at the Annual Meeting of the American Political Science Association in Washington, DC.

Southern Daily Echo. 2013. Voting like sheep for Mili's Fillies. Friday 24th May 2013. http://www.dailyecho.co.uk/yoursay/blogs/ian_murray/10443968.print/

Tomz, Michael and Robert Van Houweling. 2009. The Electoral Implications of Candidate Ambiguity. *American Political Science Review* 103(1): 83–98.

Verhulst, Brad, Milton Lodge, and Howard Lavine. 2010. "The attractiveness halo: Why some candidates are perceived more favorably than others." *Journal of Nonverbal Behavior* 34(2): 111-117.

Zaller, John. 1992. *The nature and origins of mass opinion*. Cambridge: Cambridge University Press.

Zuckerman, A. S., Fitzgerald, J. and Dasovic, J. 2007. *Partisan Families*. New York: Cambridge.

Appendix

[INSTRUCTION: Please circle the numbers that correspond the answers]

1. First I'd like to ask you about how interested you are generally in what is going on in Politics? Are you

1. Highly interested
2. Somewhat interested
3. Somewhat uninterested
4. Not at all interested

2. I'd also like to ask you about your likelihood of voting in the local council election which is coming up in May. Some people plan to vote in the election and others do not. On a scale ranging from 1 to 4, are you

1 Very likely
2 Somewhat Likely
3 Somewhat Unlikely
or
4 Very Unlikely
to vote in the council election?

3. In case you decide to vote in the council election, which party would you vote for?

INTERVIEWER: DO NOT PROMPT

0 None/Will (probably) not vote
1 Labour
2 Conservative
3 LibDem
4 UKIP
5 Green Party
6 BNP
7 SNP
8 Plaid Cymru
9 Other (WRITE IN)
10 Undecided
11 Don't know
12 Refused

4. Now I'd also like to ask you about your likelihood of voting in the General Election in May 2015. Some people plan to vote in the election and others do not. On a scale ranging from 1 to 4, are you

1 Very likely
2 Somewhat Likely
3 Somewhat Unlikely
or
4 Very Unlikely

to vote in the 2015 General Election?

5. If you decide to vote in the General Election next May, which candidate will you vote for? I am now going to read out a list of the declared candidates in your seat:

1 Rowenna Davis for the Labour Party
2 Royston Smith for the Conservative Party
3 David Goodall for the Liberal Democrats
4 Alan Kebell for the UK Independence Party

96 Other (INTERVIEWER: DO NOT NAME OPTION, WRITE IN):
97 Undecided (INTERVIEWER: DO NOT NAME OPTION)
98 Don't know (INTERVIEWER: DO NOT NAME OPTION)
99 Refused (INTERVIEWER: DO NOT NAME OPTION)

6. Now, I would like to ask you how you view the two major General election candidates. On 4 points-scale ranging from strongly positive to strongly negative, how do you evaluate the Labour Candidate Rowenna Davis?

1 Strongly positive
2 Somewhat positive
3 Somewhat negative
4 Strongly negative

98 Don't know
99 Refused

7. On the same 4 points-scale ranging from strongly positive to strongly negative, how do you evaluate the Conservative Candidate Royston Smith?

1 Strongly positive
2 Somewhat positive
3 Somewhat negative
4 Strongly negative

98 Don't know
99 Refused

8. Have you already been contacted by any of the candidates running in the General election?

8a. Rowenna Davis

1. yes
2. no

8b. Royston Smith

1. yes
2. no

9a. And the last question: Some people think of themselves as usually being a supporter of one political party rather than another. Do you usually think of yourself as being a supporter of one particular party or not?

1 Yes
2 No

9b. IF 'yes'
Which party is that?
INTERVIEWER: DO NOT PROMPT

1 Labour
2 Conservatives
3 Liberal Democrats
4 United Kingdom Independence Party (UKIP)
5 Green Party
6 British National Party (BNP)
7 SNP
8 Plaid Cymru
9 Other (WRITE IN)

Thank you very much for taking the time to talk to me. Have a nice afternoon/evening!

IMPERSONAL, BUT NOTICEABLE?
New UK Evidence on the Effectiveness of Partisan Text Messages and Door Hangers

Florian Foos

Abstract:

This paper summarizes the results of two relatively small, randomized partisan GOTV experiments that test impersonal, but highly noticeable voter mobilization methods: text messages and door hangers. Both text messages and door hangers have the advantage that they are more noticeable than other forms of impersonal campaign communication. Both methods also have in common that relatively few experiments have evaluated their use, and that perceptions about their effectiveness are hence dependent on a few selected studies that have been conducted in highly specific electoral contexts, none of them outside of the United States. Results show that GOTV text messages sent by a local Labour Party proved overall ineffective, while there is some suggestive evidence that they increased turnout among the most active voters. Moreover, despite their mixed reputation, door hangers distributed by Labour activists produced large and positive treatment effects.

Introduction

One of the best-known and established findings in the GOTV literature is that impersonal GOTV methods are less effective at mobilizing voters than unmediated, personal interactions with campaign volunteers (Gerber and Green 2008). The establishment of a social connection between the campaign and the voter is therefore seen as an important stimulus of turnout. The relative effectiveness of personal canvassing as compared to impersonal mailers is often interpreted as evidence that personal contact is a prerequisite for successful mobilization (Bedolla and Michelson 2012; Sinclair 2012, Rogers et al. 2013). However, some social theorists of voter mobilization would not emphasize the initial contact with the campaign volunteer as the decisive social element in voter mobilization, but the subsequent transmission of messages within social networks (Rolfe 2012). One could also argue that in most election campaigns, some voters already feel connected to candidates because they support the same party (Rogers et al. 2013). In contrast, personal contact might just be a correlate, and not the cause of effectiveness. There is increasing evidence that noticeability (Dale and Strauss 2009) and the quality of campaign materials (Nickerson 2007; Potter and Gray 2008) play an important and independent role in voter mobilization. What is more, noticeable reminder theory (Dale and Strauss 2009, Malhotra et al. 2011) offers an alternative to personal connectedness in explaining the effectiveness of various GOTV methods at increasing turnout. While social theorists emphasize the importance of personal contact – the more personal, the more effective – Dale and Strauss (2009) argue that positive results of door-to-door canvassing or volunteer telephone calls are also consistent with theories of noticeability – the more noticeable, the more effective. In their text message experiment they disentangle the two theories by testing an intervention that is hard to ignore, a text message, but impersonal in character. They show that text messages are effective at increasing turnout, a result that has

been replicated (Malhotra et al. 2011). However, both studies are based only on one medium, the text message, and both have been conducted in non-partisan elections in the United States. Dale and Strauss (2009) explicitly mention that their theory should also apply to other impersonal, but noticeable mobilization methods such as door hangers, and there is no reason to assume that the theory should not apply outside of the US electoral contexts.

In this paper I report the results of two relatively small, randomized partisan GOTV experiments that test impersonal, but highly noticeable voter mobilization methods in two British elections: text messages and door hangers. Both text messages and door hangers have the advantage that they are highly noticeable – but door hangers are more costly to produce and to distribute than chain messages sent via mobile phones. Both methods also have in common that relatively few experiments have evaluated their use, and that perceptions about their effectiveness hence dependent on a few selected studies that have been conducted in highly specific electoral contexts, none of them outside of the United States.

In what follows I first outline factors that might condition the effectiveness of impersonal campaign appeals such as noticeability, perceived quality and the electoral context. I then describe the design of the two field experiments that were conducted in the UK and present the results. Based on the results of the partisan text message experiment, texts sent by a local Labour Party proved overall ineffective at mobilizing voters in the 2012 UK Police and Crime Commissioner elections. However, there is some suggestive evidence that they increased turnout amongst habitual voters. Moreover, despite the fact that they are rarely used in the UK and that their reputation of effectiveness is at best mixed, door hangers that were distributed by Labour activists in the 2014 European and Council elections produced large and positive treatment effects. Although the estimates are

too noisy to tell with confidence, the size of the estimates suggests that partisan door hangers might be an effective, impersonal method of increasing turnout among party supporters.

A Question of Noticeability, Quality and Context?

Although they are often discussed together, impersonal mobilization methods vary based on noticeability (Dale and Strauss 2009) and quality (Nickerson 2007), and their effectiveness is likely to vary conditional on the electoral context (Fieldhouse et al. 2012). Therefore, if we reject the assumption that impersonal campaign methods – per se – cannot work because they do not contain an element of direct social interaction, we can understand personal interaction as a variable amongst others that conditions the effectiveness of campaign appeals. Hence, there might be some impersonal GOTV methods that might prove more effective than others.

Compared to flyers and direct mail that have proved relatively ineffective in the US context (Green and Gerber 2008), both text messages and door hangers have the advantage that they are highly noticeable. But door hangers are more costly to produce and to distribute than chain messages sent via mobile phones. The door hangers that were distributed by Southampton Labour activists cost around £0.25 per item to produce, while the text messages sent by a local Labour Party in Birmingham were free thanks to mobile contracts that include unlimited text messages. Even if one assumes that a text message costs around 8p to send, the door hanger is still three times more expensive than the text message. At the same time, door hangers need to be delivered, a campaign operation that can take multiple hours and volunteers dependent on how many door hangers should be delivered. However, higher costs might also be a signal of campaign quality, particularly if they are invested in higher-quality materials. Experiments suggest that the quality of

campaign communication matters (Nickerson 2007, Potter and Gray 2008). Potter and Gray (2008) argue for instance that good quality materials signal that a candidate is serious about competing in an election.

Moreover, while text messages are short and have little space for additional information other than a short GOTV message, door hangers can potentially include useful information that might help people follow through on their vote intentions. Although some studies have found that the content of messages does not matter, evidence from social psychology would suggest that content might matter if it helps individuals to bridge the gap between a good intention and actual behaviour (Nickerson and Rogers 2010). Hence, while a simple text message that reminds an individual to vote might be noticeable, a noticeable reminder combined with explicit instructions that help voters find their polling station and plan their trip to the ballot box might be more effective.

Moreover, any method of voter mobilization might be dependent on contextual factors such as the baseline turnout rate, or the political context of an election (Malhotra et al. 2011; Fieldhouse et al. 2012; Fieldhouse et al. 2013). Although text messages might prove useful amongst casual voters in high-salience elections and create relatively large average treatment effects, they might be less effective in a very low-salience election where they only mobilize habitual voters (Malhotra et al. 2011). The partisan composition of the electorate might be particularly relevant for mobilization tools that are deployed in partisan campaigns: If the method of contact is impersonal, the label attached to the message might play a larger role because there is no personal connection to mediate it. We might therefore expect that impersonal, partisan methods of voter mobilization are entirely ineffective with individuals who do not share the same party affiliation.

Text Message Experiment

I designed the text message experiment together with Eline de Rooij[1], and conducted it during the 2012 Police and Crime Commissioner Elections in collaboration with a local Labour Party in Birmingham, UK. Based on the two non-partisan text message experiments conducted in the US electoral context by Dale and Strauss (2009), as well as Malhotra et al. (2011), text messages could maybe count as the most cost-effective mobilization tools. A chain message takes very little time to send, and many mobile phone contracts contain unlimited numbers of text messages. Hence, the only cost incurred would be the time spent assembling and maintaining a list with mobile phone numbers of supporters. Despite these apparent advantages of text messages as GOTV tools, there are not commonly used to increase turnout in a systematic way by UK election campaigns. The local party organizer I worked with in Birmingham saw text messages as an organizing tool to coordinate volunteers rather than a GOTV tool to remind a great number of supporters about an upcoming election. In collaboration with the Labour Campaign for Police and Crime Commissioner, Eline and I hence set out to test if text messages increased turnout in a competitive partisan election. We conducted the experiment at the same time as the Telephone Canvassing Experiment that is reported in Chapters 1 and 2.

Experimental Design

The campaign held records of 1884 mobile phone numbers of registered voters. These individuals were excluded from the telephone experiment. There were some individuals for whom only a mobile phone number was available and some for whom both a landline and a mobile phone number was available. Based on this information, we created two experimental blocks, and stratified further based on recorded partisan attachments –

[1] Simon Fraser University, Vancouver.

Labour partisans, supporters of rival parties, and unattached voters. This left us with a total of six experimental blocks. Within each block the rate of assignment to treatment and control groups was identical (0.50).

To check if data was missing at random, I regressed a variable indicating if outcome data was missing (1) or not (0) on treatment assignment and compared the F-statistic of this regression to the mean of the F-statistic over 10,000 simulated randomisations. The resulting p-value of 0.37 implies that attrition is unlikely to be a function of treatment assignment. I also decided to remove all postal voters from the sample based on a pre-treatment measure since the first text message was sent one day before the election when most postal votes would have already been mailed. Moreover, I used randomization inference to conduct a balance check for the text message experiment. The resulting p-value of .07 is not significant at the conventional .05 level. Since I am confident in our randomization procedure, existing imbalances should have occurred due to sampling variability alone. To account for slight imbalances in the distributions of covariates, I will present both unadjusted and covariate-adjusted treatment effect estimates.

Treatment

Individuals in the six different experimental blocks were randomly assigned to receive two similar text messages on their mobile phones – one on the day before the election and one in the morning of Election Day – encouraging them to vote in the West Midlands Police Commissioner election on November 15th 2012. These messages simply and briefly encouraged individuals to vote for the Labour candidate in the election, either on the next day or on the same day, explicitly stating the candidate's Labour Party affiliation and one major policy stance in partisan terms. The text message also included an 'unsubscribe option' and was sent from a valid mobile phone number. A small number of individuals

indicated after receiving the first text message that they did not want to receive a second message by texting 'STOP' to the campaign.

Box 1: 2012 PCC Election Text Message Scripts.

Send on Wednesday in the early afternoon:

"Msg from your local Labour Party: Vote for Labour's [candidate name] in the Police Elections tomorrow. Labour will fight the Tory cuts to policing! Txt stop to unsubscribe."

Send on Thursday morning:

"Msg from your local Labour Party: Vote for Labour's [candidate name] in the Police Elections today. Labour will fight the Tory cuts to policing! Txt stop to unsubscribe."

Both messages are displayed in Box A1. Similarly to the telephone experiment reported in Chapter 1, in this experiment I also test for the potentially differential impact of a party mobilization message on turnout among the three distinct partisan groups.

Results

Table 1: Percentage turnout for text message treatment and control groups, by partisanship.

	Labour	Rival party	Unattached	Combined
Mobile only:				
Labour text message	21.22	14.67	6.88	14.20
Control group	23.16	10.67	6.07	14.94
N Labour text message	278	75	247	600
N Control	285	75	247	607
Mobile and landline:				
Labour text message	27.53	18.52	18.18	23.22
Control group	29.41	21.82	14.14	23.00
N Labour text message	178	54	121	353
N Control	170	55	99	324
Total:				
Labour text message	23.63	16.29	9.80	17.38
Control group	25.54	15.36	9.11	17.92
N Labour text message	456	129	346	953
N Control	555	130	368	931

Note: Percentages are OLS estimates weighted by the inverse of the probability of treatment assignment within experimental blocks.

Table 1 shows the turnout rates for the treatment group and the control group across the six experimental blocks. Although turnout rates differ substantially depending on the experimental block, the table shows little variation across treatment and control groups within experimental blocks. Turnout among Labour supporters (across both the phone groups) in the partisan and control groups, for instance, was 24-25%. For other party supporters (across both the phone groups) turnout in both the partisan and control groups was 15-16%, and for unknown party supporters 9-10%.

Next I test whether any of these differences between treatment and control groups are significant. The ITT and the CACE estimates are shown in the top half of Table 2, together with the results of the one-tailed hypothesis tests. As a small number of text messages bounced back, we can calculate the Complier Average Causal Effect identifying those subjects as Compliers who had a working mobile number. The lower half of Table 2, finally, shows the results from a re-estimation of the ITT effects using covariate-adjustment based on several highly prognostic pre-treatment covariates included in the Labour election database, namely turnout in seven previous elections, gender, age, and the electoral ward in which the subject resides.

The results show no significant increase in turnout as a result of the Labour GOTV text message for any of the three partisan groups, or for the sample as a whole. Although I found heterogeneous effects of Labour volunteer phone messages on turnout conditional on subject partisanship in the telephone experiment, text messages mentioning a candidate's party affiliation seem to have no effect on turnout among any of the partisan groups. The estimates are not only statistically insignificant, but amount to less than a 1 percentage-point increase or decrease in turnout.[2]

[2] While in retrospect, the experiment was slightly underpowered and was not able to detect significant effects of that size, it would have had enough power to detect effects similar to what I found in the volunteer telephone experiment that was conducted during the same election, and effect sizes of the magnitude reported in Dale and Strauss (2009) using a one-tailed test.

Table 2: ITT and CACE of Labour text message, by partisanship without and with covariates.

	Labour	Rival party	Unattached	All
ITT				
No covariates:				
Labour text message	-.019	.009	.007	-.005
	[-.07, .04]	[-.08, .10]	[-.04, .05]	[-.04, .03]
Covariates:				
Labour text message	-.006	-.011	.003	-.006
	[-.06, .05]	[-.10, .07]	[-.04, .04]	[-.04, .03]
CACE				
No covariates:				
Labour text message	-.022	.010	.008	-.006
	(.032)	(.049)	(.025)	(.020)
Covariates:				
Labour text message	-.007	-.012	.004	-.004
	(.030)	(.046)	(.022)	(.018)
N	911	259	714	1884

† p-value < .10; * p-value < .05; ** p-value < .01; *** p-value < .001 (based on one-tailed test of sharp null hypothesis using randomisation inference). Randomization inference-based 95%-CIs in brackets. Note: Covariates are turnout in seven previous elections, gender, age and electoral ward.

In contrast to the US-based non-partisan text experiments our study did not find any overall effect of text messages on voter turnout. Following Malhotra et al. (2011) I therefore post-stratify on previous turnout history. It seems plausible that text messages are only effective as reminders for individuals who already intended to vote in the first place, but for whatever reasons were hesitant to follow through. Given that we are operating in a low-salience election with baseline turnout rates comparable to Malhotra and colleagues' (2011) study, we would expect that the message should be particularly effective with habitual voters. Table 3 shows the results for a limited sample of voters, who turned out in the seven previous elections for which records are available. All those elections had higher baseline turnout rates than the Police and Crime Commissioner Election.

Although the estimates hardly reach statistical significance, Table 3 offers some support for noticeable reminder theory.

Table 3: ITT and CACE of Labour text message on habitual voters, by partisanship without and with covariates.

	Labour	Rival party	Unattached	All
ITT				
No covariates:				
Labour text message	.091	-.133	.119	.045
	[-.11, .29]	[-.47, .21]	[-.21, .52]	[-.12, .20]
Covariates:				
Labour text message	.064	.028	.318†	.079
	[-.16, .29]	[-.37, .42]	[-.10, .73]	[-.149, .19]
CACE				
No covariates:	.104	-.143	.143	.050
Labour text message	(.120)	(.181)	(.227)	(.092)
Covariates:				
Labour text message	.072	.032	.399†	.091
	(.133)	(.224)	(.272)	(.099)
N	89	35	29	153

† p-value < .10; * p-value < .05; ** p-value < .01; *** p-value < .001 (based on one-tailed test of sharp null hypothesis using randomisation inference). Randomization inference-based 95%-CIs in brackets. Note: Covariates are turnout in seven previous elections, gender, age and electoral ward.

As a result of sending a partisan text message, Labour supporters who have voted in the previous 7 elections were around 10 percentage-points more likely to turn out. This result is only different from zero with $p < .24$ using a one-sided hypothesis test. Moreover, treatment effects are large for habitual unattached voters, but those are relatively rare and the estimate is therefore very noisy. The results can hence not provide any conclusive evidence for the theory. However, it would be worth conducting further tests in higher salience elections or amongst larger samples of habitual voters to test if text messages are effective at mobilizing voters in partisan elections.

Door Hanger Experiment

Door hangers are widely used to get-out-the-vote in US elections (Nickerson 2005, Nickerson et al. 2006), but are not part of the traditional British campaign inventory that relies heavily on door-to-door canvassing, leaflets and, since recently, telephone canvassing (Pattie and Johnston 2003). Although door hangers are very popular in the United States, only few studies have evaluated their use, and only one has done so on a relatively large scale (Nickerson 2005). Despite Nickerson's early encouraging findings, there is the belief that door hangers are relatively ineffective mobilization tools since they do not include any personal interaction between campaign volunteers and constituents (Green and Gerber 2008). In fact, door hangers seem to have earned a (partially undeserved) reputation of relative ineffectiveness. Sinclair and colleagues sum up the received wisdom in the following quote (2013: 43): "The existing literature on the effectiveness of door hangers suggests that they have an extremely small effect on turnout (Gerber and Green, 2000; Green and Gerber, 2008; Nickerson, 2005; Nickerson, Friedrichs, & King, 2006)." However, of the literature cited, only two experiments, one reported in three papers (Nickerson 2005; Nickerson et al. 2006, Green and Gerber 2008), and a much smaller non-partisan experiment reported in Green and Gerber (2008) actually provide results of door hangers experiments. Gerber and Green (2000) evaluate the use of leaflets, while Green and Gerber's (2008) conclusion that door hangers are relatively ineffective only rests on two cases and probably on the decision to discuss the results together with the much larger GOTV literature on leafleting. The two available partisan door hanger experiments both find statistically significant effects of door hangers on turnout. Nickerson reports a CACE of around 1.6 percentage points and concludes "the cost-effectiveness of door hangers is probably the most striking finding of this set of experiments" (2005: 15). Moreover, Lenoir (2013) reports an ITT of around 5 percentage-

points on a sample of voters that would usually be selected by election campaigns, and a null effect for a sample that includes all eligible voters in a Pennsylvanian mayoral election.

Door hangers, while more costly than text messages, are relatively inexpensive (max. 25 pence for a good quality hanger) and are easy to distribute, which makes their large-scale usage potentially appealing to campaigns. However, since they have not been part of the traditional UK campaign inventory and there is more uncertainty surrounding their effectiveness than that of canvassing or telephone calls, local party chapters are understandably reluctant to introduce them. Such was the situation in the Southampton Itchen parliamentary constituency in April 2014. The Labour Parliamentary candidate advocated for using door hangers to Get-Out-the-Vote in the upcoming 2015 General Election, but the local party was initially reluctant to fund the effort. Hence, it was agreed to subject the door hangers to a fair test in the 2014 European and local elections.

Research Design

To keep the effort in time and money for the volunteers who agreed to help the Parliamentary candidate on Election Day to a minimum, we agreed to distribute the door hangers in six Tower Blocks that were located in close proximity to each other and were easy to canvass. Moreover, we agreed to include houses located on the streets that canvassers had to pass to reach the Tower Blocks in the experiment. This left us with a potential sample of 664 flats and 218 houses. When dealing with impersonal mobilization methods such as door hangers or leaflets, treatments are usually assigned at the household-level. This happens for two reasons. First, if subjects share the same household, allocating treatments at the individual level would be a logistic impossibility since it would be impossible to assign one person in the household to receive a door hanger while her spouse

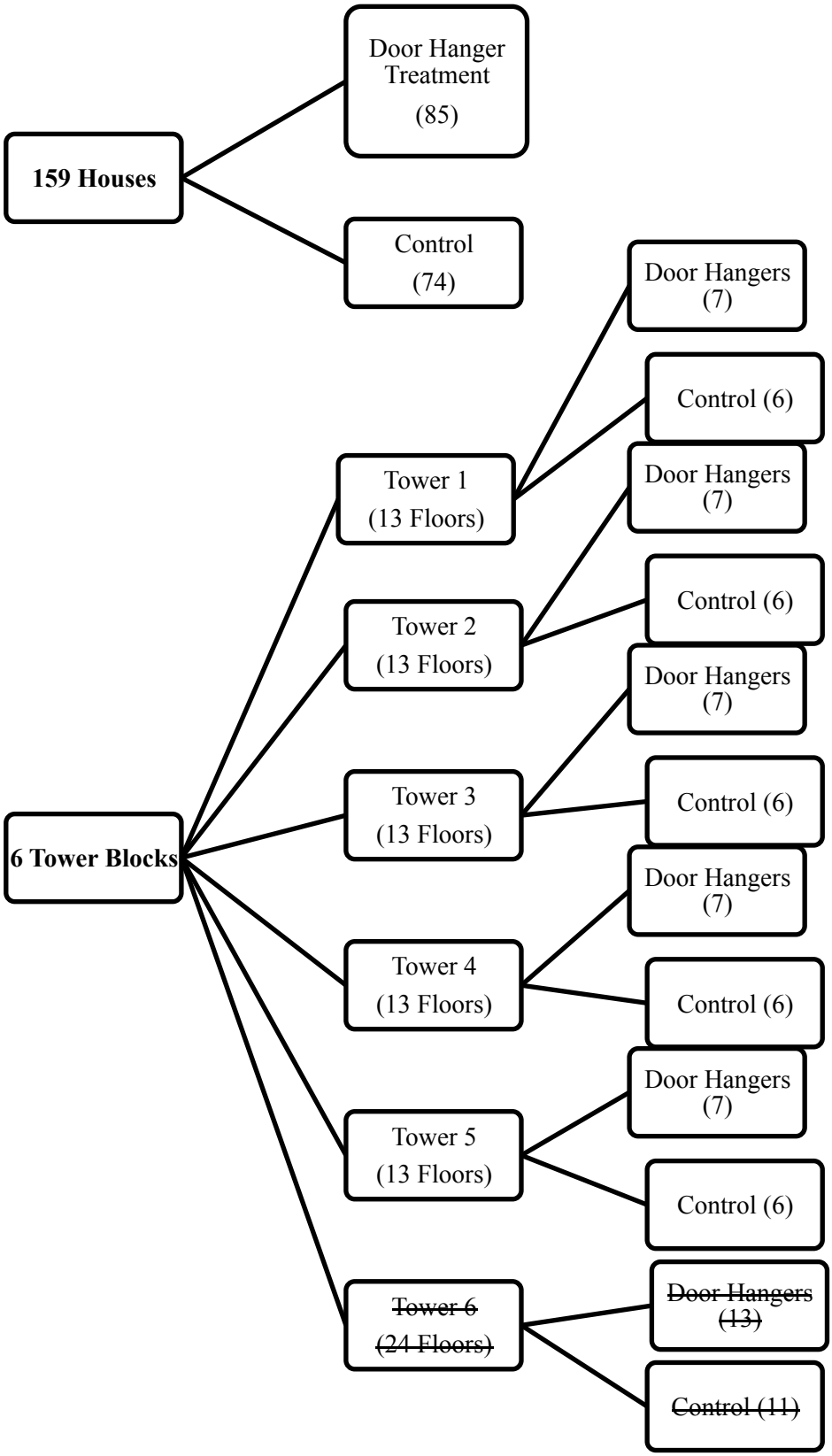
or family member would be assigned to control. Even if this sounds basic, there are recent experiments that have assigned door hangers at the individual level (Lenoir 2013). Doing so violates one of the key assumptions behind the identification of the treatment effect, the Stable-Unit-Treatment-Value-Assumption (SUTVA), which is also known as the non-interference assumption (Gerber and Green 2012). SUTVA assumes that an individual only responds to her own treatment (assignment), and not to the treatment (assignment) of others who interact with her. This assumption is violated whenever we can assume that subjects in control are affected by the treatment of others, or whenever treated subjects are influenced by the non-treatment of control units (which is less often the case). In order to avoid the introduction of bias in the estimator, we therefore need to be able to plausibly assume that subjects do not respond to the treatment of others. Although treatment assignment at the household level provides certainty that treatments do not spill over within the household, it obviously provides no guarantee that treatments do not spill over between households. This possibility was a particular concern for household in Tower Blocks that were located in very close proximity to each other. Control subjects would inevitably see a door hanger at their neighbors' door when they left the flats. To avoid spillovers it would therefore have been best to allocate treatments at the level of the tower block. It could be safely assumed that only very few subjects in control would hence be exposed to treatment. However, since the experimental sample was restricted to six Tower Blocks, doing so would have inevitably increased standard errors by a large margin. Given the small overall sample size available for experimentation, the experiment would have been unable to produce any reliable estimates. In negotiation with the campaign I therefore settled for the compromise of allocating treatments by tower floor. Although it is of course conceivable that some subjects visited people on other floors on Election Day, most subjects would use the elevator to exit or enter the building. Hence, they would not be

exposed to treatment if they were assigned to control. While this solution is not perfect, it is clearly an improvement compared to individual assignment or assignment at the apartment level. Given previous work on between-household spillovers (Sinclair et al. 2010), I was more confident that treatment effects do not spill over between neighbouring houses. Treatment assignment for individuals living in houses therefore happened at the level of the household.

For logistical reasons, I blocked the experiment by tower block. Moreover, all houses were included in a separate experimental block, which resulted de-facto in seven separate small experiments. However, I made sure that the probability of being assigned to treatment (.54) was similar in all experimental blocks. The decision to block the experiment proved useful in retrospect since some volunteers strayed from the experimental protocol in the 6th tower block by distributing door hangers to all floors because they felt bad that they had unused door hangers left over. Records show that the experimental protocol was adhered to in the remaining experimental blocks. Thanks to blocking, excluding one block from the experiment does not affect the unbiasedness of the estimators. However, it obviously increases the uncertainty surrounding the estimates. Since power calculations indicated that there was not much room for error, losing 144 households proved detrimental for detecting statistically significant effects amongst the sample of tower blocks. Before allocating houses to treatment and control I excluded all units that contained postal voters. Since the treatment happened in the early morning of Election Day, the treatment would have been ineffective for those households. The results are based on an anonymized version of the electoral database used by the local Labour Party that is updated with data from the publically available electoral register and contains information about a subject's partisanship, gender and age. The blocked and clustered

random assignment of apartment floors and houses to treatment and control are shown in Figure 1.

Figure 1: Blocked random assignment to treatment and control groups.



Balance Check

I use randomization inference to test whether I can reject the sharp null hypothesis that any existing covariate imbalance between the experimental groups could have occurred due to sampling variability alone. To do so I first extract the f-statistic resulting from an OLS regression of treatment assignment on pre-treatment covariates. I then compare the extracted f-statistic to the mean of the f-statistics I obtain after simulating the random assignment of the pre-treatment covariates within each experimental block 10.000 times.

By replicating the random assignment procedure, the simulated sampling distribution of test-statistics under the sharp null hypothesis fully accounts for the cluster- und blocked randomized experimental design. The collection of test statistics hence "represents the exact sampling distribution under the null hypothesis that no covariates have any effect on the assigned treatment" (Gerber and Green 2012: 107). The resulting p-value of .27 indicates that we cannot reject the sharp null hypothesis that the covariates taken together are not systematically related to treatment assignment.[3]

Treatment

On the day of the European and local elections, May 22nd 2014, the Labour parliamentary candidate and a number of Labour volunteers rose early to distribute the door hangers to the houses and apartment floors in the treatment group. The campaign operation lasted for three hours, between 5 am and 8 am in the morning. The aim of the early distribution on Election Day was that subjects would find the door hangers as they left for work. The door hangers were printed by a professional agency in colour on high-quality paper, and the front side contained the slogan "Vote Today. Vote Labour" in large print. This slogan was

[3] Outcome data was only missing for three subjects in the data set. While all three cases belonged to the control group, it seems very unlikely that outcome data was missing as a function of being assigned to the control group.

accompanied by the opening times of the polling station, a link to the local party website and the instruction to "turn over for map and directions". The back of the door hanger contained a coloured map that indicated the location of the local polling station and a barcode that could be scanned to transfer this information onto a mobile phone. The distributed door hanger therefore followed best practice in the layout of campaign information (Gerber and Green 2008). The front- and backsides of the door hanger are displayed in Figure 2.

Figure 2: Door Hanger Front and Back



After the election, the Labour Party collected the turnout data from the official electoral register, and identified whether individuals were allocated to treatment or control based on Apartment Block and Apartment Number, or individual addresses for the sample of houses. Apartment floors could be easily matched to addresses since all remaining Tower Blocks followed the same layout: 13 floors with 8 apartments on each floor. Apartment Numbers 1-8 would hence be found on the first floor, numbers 9-16 on the second floor, and so forth. The data was then anonymized and forwarded to me for the analysis.

Results

For the post-treatment analysis I further stratified the sample into likely Labour supporters and unattached voters based on information about individual partisanship. Houses that belonged to known supporters of rival parties were excluded from the database by the Labour campaign before the anonymized data was sent to me for randomization.[4] Information on partisanship is based on the local party's canvassing of constituents over the past years. As paper one in this thesis has discussed, voter ID data collected through canvassing should be a valid and reliable, albeit a slightly noisy measure of partisanship in the United Kingdom. I decided to post-stratify on partisanship because if high-quality door hangers really work as a reminder to turn out, I expected the treatment to be more effective for Labour supporters than for unattached voters.

Table 4 displays the differences in proportions between treated and untreated units. P-values and confidence intervals are estimated using randomization inference (Gerber and Green 2012; Aronow and Samii 2012). The baseline turnout rate in the control group was around 54 percent, 56 percent in the tower blocks and 49 percent in the houses.

[4] Since treatment assignment in the Tower Blocks was carried out at the floor-level, volunteers happened to treat a small number of rival party supporters in the Tower Blocks, but the Labour data base does not provide any information about them.

Table 4 ITT of door hangers on turnout, by housing type and subject partisanship

	Houses			**Tower Blocks**			**Combined**		
Party ID	Lab	Unattached	Comb	Lab	Unattached	Comb	Lab	Unattached	Comb
				Unadjusted					
Door hanger	.187* [-.05, .43]	-.010 [-.13, .11]	.101† [-.03, .23]	.058 [-.17, .29]	-.013 [-.09, .06]	-0.01 [-.09, .07]	.090 [-.09, .27]	.001 [-.06, .07]	.032 [-.04, .10]
				Covariate-adjusted					
Door hanger	.177† [-.07, .42]	-.001 [-.12, .12]	.113* [-.01, .24]	.059 [-.18, .30]	-.013 [-.09, .06]	-.006 [-.08, .08]	.099 [-.08, .27]	.005 [-.06, .07]	.036 [-.03, .11]
N	88	141	229	95	282	377	183	423	606

† p-value < .10; * p-value < .05; ** p-value < .01; *** p-value < .001 (based on one-tailed test of sharp null hypothesis using randomization inference). Randomization-inference-based 95%-CIs in brackets. Note: Covariates are gender, age, party id, and household size. The latter is also included in the unadjusted estimates.

By replicating the random assignment procedure, the simulated sampling distribution of test-statistics under the sharp null hypothesis fully accounts for the cluster- und blocked randomized experimental design. Since we cannot know if subjects saw the door hanger, all treatment effects will be Intent-to-Treat Effects only. Since there are very few four and five voter households, I display the ITT estimates adjusted by household-size only, as well as a fully covariate-adjusted ITT including additional information on gender, age and partisanship. Column 1 shows the results for the block of houses, column 2 shows the results for the five remaining tower blocks and column 3 displays a pooled average of all experimental blocks. I also display the Conditional Intent-to-Treat Effects for Labour supports and non-Labour supporters.

Table 4 shows that, averaged over all experimental blocks, subjects that lived in households or on floors that received a door hanger were more than 3 percentage-points more likely to turn out and vote (3.2 percentage-points unadjusted and 3.6 percentage-points with covariate adjustment). This effect is different from zero with $p < .16$ using a one-tailed hypothesis test. We can therefore conclude that under repeated random assignment, we would expect to find an ITT as large or larger than 3.6 percentage points due to sampling variability in 16 out of 100 samples. The lack of conventional statistical significance is most likely due to the limited sample size (in part resulting from the failed treatment in tower block 6). The substantive size of the overall ITT effect positions itself somewhere in the middle between Nickerson's (2005) lower bound of 1.4 percentage-points and Lenoir's (2013) higher bound of 5 percentage-points.

Moreover, the overall effect masks differential effects conditional on partisanship and housing type. For Labour supporters, the estimated ITT reaches 9-10 percentage-points, and is different from zero with $p < .14$ using a one-sided hypothesis test. There are no detectable positive or negative effects for subjects who do not support the Labour Party.

Moreover, the treatment effects on Labour supporters are almost double the size in family homes than in apartment blocks. Labour supporters who live in family homes and who were assigned to receive a door hanger are 19 percentage-points more likely to turn out than Labour supporters in family homes that have been assigned to the control group. This effect is statistically significant at conventional levels with $p < .05$.

The substantially smaller and statistically insignificant effect for subjects who reside in apartment blocks could have occurred for a multitude of reasons: Sampling variability, spill-over effects between apartment floors, or composition differences in the electorate between the tower blocks and the houses. Importantly, the share of identified Labour supporters was much larger among houses than in the tower blocks, which explains why the overall effect of the door hangers in the tower blocks approaches zero.

Discussion and Conclusion

The results of the two experiments reported in this paper give us reasons to be both pessimistic and optimistic about the effectiveness of impersonal, but noticeable, GOTV methods in partisan election campaigns. First, we report the first null finding of a randomized text message experiment. Some might say that these results are due to the low salience of the election or the small sample size. However, Malhotra and colleagues (2012) find statistically significant results on a similarly sized sample of habitual voters (the only subsample for which they expect to find any effect) in an election that has an even lower baseline turnout rate. What we show is that in low turnout elections, the number of constituents that can be moved by text messages is so small, that their large-scale use might be relatively ineffective. Still, given that some parties hold detailed information about prior turnout history, texting habitual voters might increase their propensity to vote in very low salience elections. Moreover, the relatively large effect sizes among habitual

Labour and Unattached voters provides some indicative support for noticeable-reminder theory in the version defended by Malhotra and colleagues (2011) that claims that the effects of noticeable reminders depend on an interaction between individual turnout propensity and election salience.

Moreover, I find more encouraging results when evaluating the effectiveness of partisan door hangers in the 2014 UK European and local elections. The EP and local elections, which took place on the same day, are of higher salience than the Police and Crime Commissioner elections, with a baseline turnout rate in the control group of around 54 percent. Although, partly due to an implementation mistake, the estimates do not reach conventional levels of statistical significance, the size of the estimates is substantial and should encourage replication with larger sample sizes, and based on a sample of party supporters only. Although the results for the most part prove inconclusive, I cannot dismiss the possibility that noticeable reminders can provide a significant boost to turnout. We need to further explore the conditions under which they can be used effectively. Based on the results of the experiments reported in this paper and earlier work on noticeable reminders, it looks like non-personal, but noticeable methods of voter contact have the potential to increase turnout among voters who are likely to respond given their partisan identity, the particular electoral context and the expected baseline turnout rate. Experiments using impersonal methods of voter contact should therefore be carefully targeted to receptive audiences. It is therefore crucial that careful experimentation is combined with strong predictive modelling of heterogeneous treatment responses. For future experiments, researchers could also think about randomly allocating interactive elements such as bar codes printed on the door hangers that would potentially allow us to locate a sub-sample of Compliers, and to identify the effectiveness of different elements of campaign communication in more detail.

List of references

Aronow, Peter M., and Cyrus Samii. 2012. ri: The R package for performing randomization-based inference for experiments. R package version 0.9. http://CRAN.R-project.org/package=ri

Bedolla, Lisa Garcia, and Melissa R. Michelson. 2012. *Mobilizing inclusion: Transforming the electorate through get-out-the-vote campaigns*. Cambridge, MA: Yale University Press.

Dale, Allison, and Aaron Strauss. 2009. "Don't Forget to Vote: Text Message Reminders as a Mobilization Tool." *American Journal of Political Science* 53: 787–804.

Fieldhouse, Edward, David Cutts, Peter John, and Paul Widdop. 2013. "When Context Matters: Assessing Geographical Heterogeneity of Get-Out-The-Vote Treatment Effects Using a Population Based Field Experiment". *Political Behavior:* published online.

Fieldhouse, Edward, David Cutts, Peter John, and Paul Widdop. 2012."Do impersonal mobilisation methods work? Evidence from a nationwide Get-Out-the-Vote experiment in England." *Electoral Studies* 32(1): 113-123.

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* 94(3): 653-663.

Green, Donald P., and Alan S. Gerber. 2008. *Get Out the Vote: How to Increase Voter Turnout*, 2nd edition. Washington, DC: Brookings Institution.

Lenoir, Brandon W. 2013. *Studying the effects of mobilization messages in a strategic environment: Applying the receive-accept-sample (RAS) model*. Diss. University of Pittsburgh.

Malhotra, Neil, Melissa R. Michelson, Todd Rogers, and Ali A. Valenzuela. 2011. “Text Messaging as Mobilization Tools: The Conditional Effect of Habitual Voting and Election Salience.” *American Politics Research* 39: 664-681.

Nickerson, David W. 2005. "Partisan mobilization using volunteer phone banks and door hangers." *The Annals of the American Academy of Political and Social Science* 601(1): 10-27.

Nickerson, David W. 2007: “Quality Is Job One: Professional and Volunteer Voter Mobilization Calls,” *American Journal of Political Science* 51(2): 269–282.

Nickerson, David W., and Todd Rogers. 2010. "Do you have a voting plan? Implementation intentions, voter turnout, and organic plan making." *Psychological Science* 21(2): 194-199.

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): 85-97.

Pattie, Charles J., and Ronald J. Johnston. 2003. "Hanging on the telephone? Doorstep and telephone canvassing at the British general election of 1997." *British Journal of Political Science* 33(2): 303-322.

Potter, Philip B.K. and Julia Gray. 2008. “Does Costly Signaling Matter? Prel: iminary Evidence from a Field Experiment.” BCSIA Discussion Paper 2008-09, Kennedy School of Government, Harvard University.

Rogers, Todd, Alan S. Gerber, and Craig R. Fox. 2012. "Rethinking Why People Vote: Voting as Dynamic Social Expression." in Shafir, Eldar (ed.) *The behavioral foundations of public policy*. Princeton University Press: 91-107.

Rolfe, Meredith. 2012. *Voter Turnout: A Social Theory of Political Particip*ation. Cambridge University Press, Cambridge, UK.

Sinclair, Betsy. 2012. *The Social Citizen: Peer Networks and Political Behavior*. Chicago: University Of Chicago Press.

Sinclair, Betsy, Margaret McConnell, and Melissa R. Michelson. 2013. "Local Canvassing: The Efficacy of Grassroots Voter Mobilization." *Political Communication* 30(1): 42-57.

Appendix

Figure A.1 Distribution of Tower Block Apartments by Floor



The National Effects of Subnational Representation:

How parties' access to subnational parliaments shapes national elections

Elias Dinas* & Florian Foos

Abstract

According to scholarly wisdom, party competition at the subnational level plays a negligible role in national elections. We provide theory and evidence that qualifies this view. Subnational elections determine entrance into subnational parliaments, which provides essential organizational resources: members and money. Since in most cases the same political actors compete at all levels of government, they can make use of these resources to improve their status in national party competition. We test our argument exploiting two institutional features of the German multi-level electoral context: the discontinuities generated by the five percent electoral threshold in German state elections, and the occurance of German state elections at different times in the federal election cycle. We find that parties that marginally cross the threshold for state parliamentary representation gain more members, and eventually perform better in national elections, but only if the party is given sufficient time to organize between the state and the federal election. Consistent with our organizational explanation, bottom-up effects are more pronounced where state parliamentary parties receive more financial resources.

*Brasenose College, University of Oxford.

Over the last decades many countries have delegated previously centrally coordinated competencies to their subnational authorities (Benz 2009; Deschouwer 2003; Rodden 2004). This decentralization process has been accompanied by the formation and strengthening of subnational representative institutions, as well as the growing number and importance of subnational elections fostering party competition at the subnational level (Deschouwer 2003; Detterbeck 2012; Schakel 2013). An important question that arises from this development is whether representation in this lower tier of politics has had any noteworthy effect on national electoral results.

The bulk of the existing literature suggests that this should not be the case. Most studies point to top-down electoral dynamics, whereby subnational politics are mainly reflections of the national electoral momentum and the national election cycle (Campbell 1991; Dinkel 1977; Gabriel 1989; Hainmueller and Kern 2005; Jeffrey and Hough 2002; Jérôme and Lewis-Beck 1999; Lohmann, Brady and Rivers 1997). Subnational elections are characterized as second-order elections (Marsh 1998; Reif and Schmitt 1980, 9), used by voters as mid-term elections, either to punish incumbent parties (McLean, Heath and Taylor 1996; Pallarés and Keating 2003; Tronconi and Roux 2009) or to accommodate their policy demands through electoral balancing (Erikson and Mikhail 2001; Lutz Kern and Hainmueller 2006; Rodden and Wibbels 2010).

The same conclusion is also reached by a different strand of research, pointing to the nationalization of party organization and election campaigns. According to the cartel party thesis (Katz and Mair 1995; 1997; 2009), party competition becomes dominated by the national level at the expense of the local party organization. For the mass party, members were deemed vital for parties' survival. In an electoral context, however, that promotes centrally-orchestrated campaigns in which the party's message is quickly and effectively disseminated to the electorate through the use of the mass media, the question whether a party engages the party membership is said to be less vital to party competition (King 2002; Mughan 1995).

Drawing on Panebianco's theory of party institutionalization (1988), we develop a theoretical account that challenges this view. Although our departure point builds on the "cartel party thesis" (Katz and Mair 1995), we depart from Katz and Mair (1995) to accommodate

multi-level systems of government. The mechanism driving the cartel party thesis is that parties that are part of the political system —so-called "insiders"— gain an electoral edge over their competitors by taking advantage of state resources (Katz and Mair 1994; 1995; 1997). Resources are deemed vital for institutionalization, the process by which political parties ensure their long-term survival (Panebianco 1988). We extend this logic of institutionalization into multi-level government. Outcomes of subnational elections influence national electoral dynamics because they determine whether parties enhance their organizational capacity by gaining access to financial and human resources at lower levels of government. Since typically the same political parties compete at all levels of government, they can use the money and members gained from entering subnational parliament to improve their standing in a national election.

A empirical test of this idea involves a comparison between parliamentary "insiders" and "outsiders," i.e. parties that do not usually enter parliament. The problem, however, is that these two groups of parties differ on many accounts apart from whether they have access into the resources provided by subnational representative institutions. Since entrance into parliament is not randomly assigned, it is difficult to disentangle the effect of subnational parliamentary representation on national electoral performance from all other potentially confounding differences between the two groups.

We address this problem by carefully selecting our case of study. In particular, we focus on the German federal system. In German state ("Land") elections entrance into the state parliament is determined by the presence of an electoral threshold. Parties need to receive at least 5% of the vote in order to gain access to the state legislature. Parties that receive less than 5% of the vote stay out. We thus construct our comparison groups by using the discontinuities generated by this five percent threshold that applies to all state elections in Germany.[1]

Even if this design permits the identification of the effect of parliamentary representation on national electoral performance, it still leaves open one important alternative explanation.

[1] The same threshold applies in federal ("Bund") elections, but we do not use this here as electoral performance at the federal level is only used as our outcome of interest.

Entering state parliament might help parties' future electoral performance not through the provision of material resources but simply because it operates as a signal of electoral viability. In so doing, it may yield spillover effects. Electoral success in one arena might lead to more success in another arena because voters want to be on the winning side or because there is a feeling of ''momentum" (Bartels 1985; Goidel and Shields 1994; Simon 1954).

The German electoral setting is particularly helpful in also addressing this competing explanation. German state elections occur at different times during the federal election cycle. While the spillover mechanism should be more pronounced if elections follow in quick succession, the organizational benefits stemming from entering subnational parliaments should not materialize immediately. Indeed, organizational effects stemming from institutionalization should be the larger, the more time parties have to put these resources to use. The two mechanisms can thus be observationally distinguished by considering the variation in the time elapsed between subnational and national elections. We unpack the mechanism further by examining two observational implications of our organizational theory. First, we show that the electoral effects of entering state parliament are the larger, the larger the available state funding for parliamentary parties in a given state. Second, we further trace the organizational roots of these effects by testing whether subnational representation affects the number of party members in a given state, which we argue is a good measure of organizational capability.

In what follows we first develop the organizational hypothesis and derive expectations about the effect of subnational parliamentary representation on national electoral performance. We then describe the research design and present the results. We discuss the implications from our findings in the concluding section.

The organizational returns to subnational parliamentary representation: money and members

In a recent study, Bechtel (2012) argues that there are important bottom-up effects from subnational politics to national politics that have previously been ignored. As he acknowledges, "we know virtually nothing about whether and how subnational elections influence

vote intentions at the national-level" (Bechtel 2012, 3). We posit that one of the ways in which subnational elections affect national electoral outcomes is by determining which parties gain access to subnational representative institutions. The logic is exemplified in Panebianco's following remark about the way in which different electoral contests are interlinked (Panebianco 1988, 9):

> *Party arenas are interdependent and can be conceived as a network of "relevant" environments. Resources obtained in one arena can be spent in another, and success at one gambling table —the exchange of resources under favorable conditions— often affects the extent of one's success at other tables.*

Multi-level systems of government provide multiple points of access to two types of organizational resources [2] that are key to a party's electoral success: money and party members. In decentralized systems, a larger share of these resources is available at the subnational level of government, and varies conditional on entrance into state parliament: state parliaments provide parties with money to pay for the professionalization of their staff and political activities, which in turn increases the attractiveness of the party to potential activists and supporters. Money and members enable parties to better weather political eventualities and national electoral dynamics. Let us explore these two types of benefits in more detail.

The first organizational resource available to state parliamentary parties is money. State parliaments provide generous budgets to parliamentary parties for their organization and day-to-day business. Parties use this money to pay their party leadership, state representatives, parliamentary researchers, special advisors, and constituency case workers (Katz and Mair 2009). Germany represents a typical case of this practice. Data that we collected from the federal budget and the 16 state budgets show that in 2010 parliamentary parties received a

[2] The assumption that holding elected office comes with resources that are electorally meaningful is mirrored in the literature on US Congressional elections, which shows that direct office holder benefits contribute to explaining the incumbency advantage (Levitt and Wolfram 1997; Serra 1994).

total of around 186 million euros for the upkeep of their parliamentary work, of which around 108 million euros were paid to state parliamentary parties (see Table A.6 in the Appendix for a break-down by state). This compares to a total of 132 million euros at all levels of government that was available through the direct public party financing mechanism, which does not discriminate between parliamentary insiders and outsiders (Bundestag 2011). It is no wonder then that Katz and Mair (Katz and Mair 2009) increasingly see the roots of cartelization in the parliamentary party. Scholars such as von Arnim (2011; 2013) have argued that the funding provided to parliamentary parties in the German system constitutes a form of opaque and hidden party financing, which is more important than official public party funding. This is the case because the levels of funding for parliamentary parties is uncapped by law, and therefore higher than the official funding parties receive from the state (von Armin 2011). Moreover, there is much evidence that parties use these funds made available for the parliamentary party to support the work of the party on the ground (ibid). Although this practice is strictly-speaking "unconstitutional", there exists a large grey area, which allows parties to use these funds for political purposes other than parliamentary work. Moreover, parliamentary parties use part of these funds to raise the salaries of some parliamentarians, chiefly benefitting their leadership. State funds are also used to equip parliamentary and constituency offices, from where state representatives provide a range of services to constituents. Importantly, all these financial resources stay in the state and support the state party's political operation. If a party loses parliamentary representation in a given state, the money stops flowing and all employees of the state parliamentary party will be made redundant, meaning the party will not only lose all its members of parliament, but also twice as many professional staff, which vastly outnumber staff in central party offices.

Second, entrance into legislatures improves party organization by increasing incentives for state party membership and grassroots activism (Panebianco 1988). There are at least three types of motivations for joining a party in the parliament. First, winning or losing a seat affects personal efficacy, the prospects of a political career. Whiteley and Seyd (1998), and Fisher and colleagues (2006) find that personal efficacy is a key element explaining why winning a constituency election contest boosts party membership and activism in UK par-

liamentary seats. Political careers normally begin at the subnational level, especially in the context of multilevel government (Katz and Mair 2009). Second, presence in parliament increases group efficacy, the prospects of influencing policy (Fisher, Denver and Hands 2006). Seyd and Whiteley (1992) and Whiteley, Seyd and Richardson (1994) provide survey evidence that many activists join a party because they feel they can affect policy —a perception that is greatly enhanced once a party enters parliament. If subnational levels of government provide avenues for affecting policy, the same mechanism should hold in multi-level systems. Third, it is also possible that parties reward their members through patronage appointments in ministries, or the public sector more widely (John and Poguntke 2012; Poguntke 1994; Smith 1979). Although parties present in government probably have a larger influence over patronage appointments, in a consensus democracy such as Germany, at least some appointments are thought to be the outcome of negotiations between government and opposition (Smith 1979).[3]

Increased grassroots membership benefits parties in a multitude of ways. Although in the short run new parties might be able to attract activists without institutionalized structures of support on grounds of shared ideology and collective interests alone, in the long run institutionalization and the ability to affect policy and provide a political career are crucial to the survival and success of the party (Panebianco 1988, 166). In multilevel systems, state party members who joined as a result of a successful subnational election will of course be available to campaign for the party in higher-order, national election contests. Party members and supporters have a positive impact on voter mobilization and persuasion (Cox 2010; Seyd and Whiteley 1992; Whiteley, Seyd and Richardson 1994). They are ready to distribute leaflets, put out electoral advertising and knock on doors. Moreover, party members will spread the party's message in their households, neighborhoods and work places. As Aldrich (1995) has

[3] To be sure, partisan appointments might not always reflect clientelistic practices, but the attempt to control decision making. Indicatively, Kopecký, Mair and Spirova (2012) present evidence that patronage within the civil service in advanced democracies is increasingly used to promote civil servants with the right party affiliation to exercise control over policy making, and less to reward faithful party members with jobs in the regional administration. Irrespective of the exact motives driving such appointments, the fact that they are decided on partisan grounds is likely to increase supply of membership.

shown at the example of the United States at the beginning of the 19th century, state party organization is key to electoral success at the national level. By entering state parliament, a political party passes an important "organizational threshold" (Panebianco 1988, 193) that helps foster further institutionalization.

A key characteristic of organizational resources is that positive electoral externalities are not felt immediately, but need some time to materialize. Electoral support is expected to rise as the number of voters who benefit from services provided to constituents increases over time. Improvements in the quality of candidates recruited by the state party will need some time to capture media attention. An extended membership base is more likely to affect election results after the party organizes its grassroots support and new members gain campaign experience. Dissemination of party messages also needs time to reach a wider spectrum of constituents. Thus, there are often important long-term effects to improved party organization (Fisher, Denver and Hands 2006).

It is important to note that our line of thinking is driven by general and widely applicable benefits accompanying entrance into subnational legislatures. Our argument does not depend on the direct public subsidies provided to political parties in some countries. In other words, the benefits discussed here are not contingent upon direct public party financing. They exist even in the absence of direct subsidies to political parties because parliamentary parties have found ways to allocate funds to themselves that are not as easily scrutinized under party financing rules. Although the empirical evidence about the electoral impact of public party financing rules is still ambiguous,[4] the simultaneous presence of public party financing might conflate the observed effects attributed to parliamentary representation. As is explained below, the choice of Germany addresses this concern as official public party financing rules in Germany do not discriminate between parliamentary insiders and outsiders.

---

[4] While Katz and Mair (1994) originally located the roots of cartelization in the introduction of direct public financing, in recent years they have raised doubts about its impact on cartelization (2009; see also Pierre, Svasand and Widfeldt (2000); Scarrow (2006)).

An Alternative Explanation: spillover effects

Focusing on the importance of regional elections, Bechtel (2012) alludes to their role as information cues transmitted by the national media about state party performance, candidates and electoral success. The bandwagon effects literature on US Presidential primaries goes one step further suggesting that electoral success or failure in one state's primary election produces spillover effects if there is another primary following in quick succession (Bartels 1985; Simon 1954). It has been argued that electoral success leads to more success because people want to be on the winning side (Goidel and Shields 1994). This mechanism has also been associated with increasing media reports and the way individuals recall the information, on which they base their voting decisions (Iyengar 1990; Mutz 1995). Individuals are more likely to recall information to which they have been exposed more recently (Zaller 1992). Extending this line of argument, one could argue that with the exception of major parties, entering into the state parliament might signal success with possibly important side-effects for parties' future electoral performance.[5]

Despite leading to equifinal expectations, the spillover mechanism differs from the organizational hypothesis in one important respect: time. While electoral spillover effects can explain electoral momentum from one election to the next if the two elections follow in quick succession, they do not account for the financial and organizational advantages provided to parties that perform well in state elections and manage to enter parliament. The media hype often only lasts a couple of weeks and an electoral bounce after a successful performance is likely to fade away.[6] But substantial benefits of representation, be it service provision in

[5] An important critical view to the electoral importance of the spillover mechanism is Mutz (1995). While Mutz agrees that increased media attention leads to an incentive to consider the respective party, she cautions that this need not necessarily lead to a voting decision in favor of the party. People "reassess their own views in light of this new information. From this perspective, momentum appears far less pernicious than typical references to bandwagon phenomena would suggest" (Mutz 1995, 120).

[6] A nice illustration of how spillover effects operate is provided by the German Pirate Party, which after entering the state parliament of Berlin in late 2011, attracted much media attention, skyrocketed in national polls and entered 2 further state parliaments in quick succession scoring 7.4 and 8.4 percent of the vote. However, one year and a half after the

the constituency or organizational and institutional privileges, are there to stay. More importantly, in contrast to spillover effects, they are more likely to yield fruits only after some period has passed since the party entered parliament.

Data and research design

To test the effect of state parliamentary representation on federal party success in Germany, we created our own data set, which includes the electoral results of all parties that participated in German state and federal elections from 1947 up to 2009. The data set is based on the official state and federal election results published by the federal election commission (Bundeswahlleiter 2009). For the period from 1947 up to 1990 it includes the complete results for the 10 states that made up the former Federal Republic before reunification. For the period between 1990 and 2009 the data set includes all electoral results for the 10 old states and the 6 new states (including the city state of Berlin).

Federal elections are held every 4 years and state elections every 4 to 5 years. Each of the old states hence contributes 16-18 state election results and 17 federal election results, while the new states contribute data on 6 state elections and 6 federal elections. The resulting dataset gives us more than 50 small parties that have at any point in time achieved an electoral percentage near the 5% threshold. No party provides more than 20% of observations. The four parties that contribute most of the observations are the FDP, the Greens, the NPD, and the Republikaner.

Since parliamentary insiders differ from outsiders in various unsobservable ways, we make use of the discontinuities generated by the electoral rules used in German state elections. German parties can only enter the state parliament if they have gathered more than five percent of the total state vote. We use the presence of this electoral threshold, which applies to all states for the whole period under investigation, as a way to identify the effect of

party's initial success, it fell back to around 3 percent in national polls and ended up with a poor 2.2% in the 2013 federal election.

state parliamentary representation on federal electoral performance.[7] In particular, we focus on parties with vote shares in the neighborhood around the electoral threshold, comparing parties above this threshold (treated) with parties below the threshold and hence without state representatives (controls).[8]

As in all regression discontinuity (RD) designs, our identification strategy is based on the assumption that agents—in our case German political parties— cannot precisely manipulate the assignment variable—their vote share in the state election—so that it lie just below or above the cutoff value—the 5% threshold. Under this assumption, variation in treatment status in the neighborhood of the threshold is "as good as randomly assigned" (Lee and Lemieux 2009).[9] For instance, the FDP in 1980 received 4.98 percent of the vote in the state election of Niedersachsen and was thus absent from the state parliament. The counterfactual question we try to answer is whether this party would have done better in that state in the following federal election of 1983 if it had achieved only 0.02 percent more so as to make it to state parliament. Reversely, we ask how this party would have done in the 1994 federal election in Hamburg, had it performed marginally worse in the 1991 state election, where it surpassed the threshold with a tight 5.4 percent of the vote. Such instances are not very rare

[7] There are two exceptions: Following a constitutional court ruling in 1955, the threshold does not apply to the party of the Danish ethnic minority SSW, which is always guaranteed representation in Schleswig-Holstein. After 1955 the SSW is therefore excluded. Moreover, an exceptional rule has been applied in the state parliament of Bremen, where a party that fails to pass the threshold state-wide, but passes the 5% threshold in the city of Bremerhaven, gets 1 representative. This was the case for the DVU in the 1999 and the 2003 state elections and for the FDP in the 2003 state election. Results are not sensitive to the inclusion or exclusion of these cases (available upon request).

[8] Although the findings cannot be extrapolated to major parties or to very small parties, this is not particularly problematic here since for these parties parliamentary representation is almost invariably held constant. Major parties are always guaranteed entrance in the state parliament and very small parties are always left out. Therefore, it cannot be parliamentary representation that accounts for spatial and over-time variation in their electoral performance in federal elections.

[9] This assumption does not mean that there no other determinants of parties' state vote share in the federal election. Rather, it means that the association of all these factors with the outcome is smooth, so there is no other covariate that would cause a discontinuous jump in the conditional distribution of the outcome.

in German state elections.

Importantly, focusing on the discontinuities generated by the 5% threshold applied in German state elections allows us to estimate the effect of state parliamentary representation net of the effect of direct public subsidies. According to the party financing regulations in Germany, parties receive public subsidies once they obtain 1% of the state vote. The amount of these subsidies grows in a linear fashion, i.e. in proportion to party's vote share. Thus, as a result of the direct public party financing mechanism there should be no gap between parliamentary insiders and outsiders.

We estimate the effects with local linear regression, which has been shown to have attractive bias properties in estimating regression functions at the boundary (Fan and Gijbels 1992) and enjoys rate optimality (Porter 2003). Following standard convention (Imbens and Lemieux 2008; Lee and Lemieux 2009), instead of implementing two seperate regressions below and above the threshold, we run a pooled regression on both sides of the cut-off point:

$$Y_{i,s,t+1} = \alpha_l + \tau D_{i,s,t} + \beta_l(X_{i,s,t} - c) + (\beta_r - \beta_l)D_{i,s,t}(X_{i,s,t} - c) + \epsilon_{i,s,t+1},$$

where $Y_{i,s,t+1}$ represents party i's vote share in the federal election $t+1$ at state s, X_{is} denotes party i's vote share in the state election s at t, c denotes the 5% threshold and $c-h \leq X \leq c+h$. D is a dummy that switches on for the parties that acquire representation in the state parliament. Using a triangular kernel, the estimation amounts to a weighted linear regression within a bin h, whereby more weights are put on observations closer to the threshold. We all the slope coefficients of each side to differ, thus we denote them here by β_l for the left side of the cuttof point and β_r for the right side respectively. α_l denotes the regression intercept at the left side of the cutoff point, whereas α_r denotes the intercept at the right of the cutoff point, as evaluated at $X = c$. Hence, $\tau = \alpha_r - \alpha_l$, i.e. the difference between the two intercepts as (minimally) extrapolated at the point of the discontinuity.

The key question is the choice of the bandwidth, h. We use Imbens and Kalyanaraman's (2011) optimal bandwidth algorithm (IK), which performs well in recovering experimental (Green et al. 2009), observational and simulation (Imbens and Kalyanaraman 2011) benchmarks. However, since this estimator is data driven, it sometimes exceeds the four percent

window below the threshold. This means that parties below 1% might be included in some of the analyses. Given that public subsidies start with 1% percent of the vote, there is a possibility for a within-group jump. We address this problem in four ways. First, we examine the sensitivity of the estimates across different bandwidths. Second, we follow Cattaneo, Frandsen and Titiunik (2013) in providing local randomization-inference based estimates, focusing only on the very few observations exactly in the neighborhood of the 5% threshold and well above the 1% threshold. Details regarding the window selection process and the balance tests accompanying our treatment effect estimates are provided in the Appendix.[10] Third, we explicitly test for within-group jumps. Fourth, we also present estimates using a fixed 4% bandwidth ($h = 4$), which results in a comparison of insiders and outsiders, all of whom receive state subsidies made available through the official state financing system of political parties.[11] We also conduct the standard diagnostic checks on the validity of the RD design: a test for sorting in the forcing variable and several balance tests, which we report in a separate section.

As a final point, we need to clarify how we try to disentangle our organization-driven effects from the possibly confounding spill-over effects. We make use of a particular feature of the German federal political system: the occurrence of state elections at different times during the federal election cycle.[12] The dates of the state elections originally differed because the states that later joined the Federal Republic of Germany were constituted at different dates prior to the founding of the Federal Republic in 1949, and were part of four different occupation zones. All states occupied by the United States held early elections in 1946, while the majority of British and French occupied states followed only in 1947. Moreover, snap

[10] Focusing only on these observations makes it also easier to satisfy the exchangeability criterion embedded in our understanding of the RD design as a local randomized experiment.

[11] As a further robustness check, we replicate our main analyses excluding elections between 1959 and 1969, since during this period—the first period in which state subsidies were introduced—access to this source of financing was partially contingent upon parliamentary representation (Duebber 1962, 80). The results, reported in the Appendix (Table A.2), are very similar to our main results discussed here.

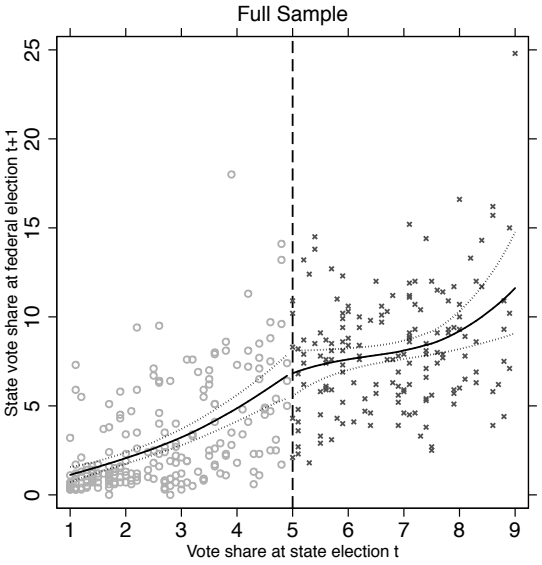
[12] It is relatively uncommon for German state prime ministers to call snap elections. With a few exceptions, elections happen as scheduled at the natural end of each state parliament.

federal elections were called in 1972, 1983 and 2005, which rearranged the order of state elections in the federal election cycle. We assume that spillover effects are more likely to emerge when the period between state and federal elections is small and less likely to appear when the two elections are quite far apart. The institutionalization effects hypothesized here need time to materialize, as parties need to organize on the ground. Thus, using the distance between state election and the next federal election, we can check whether the effect of crossing the five percent threshold are higher when the next election is closer or whether this happens when there is a greater time span between the two elections. In the second case, the result would be more easily compatible with the idea of organizational benefits than with the argument about the presence of spillover effects.

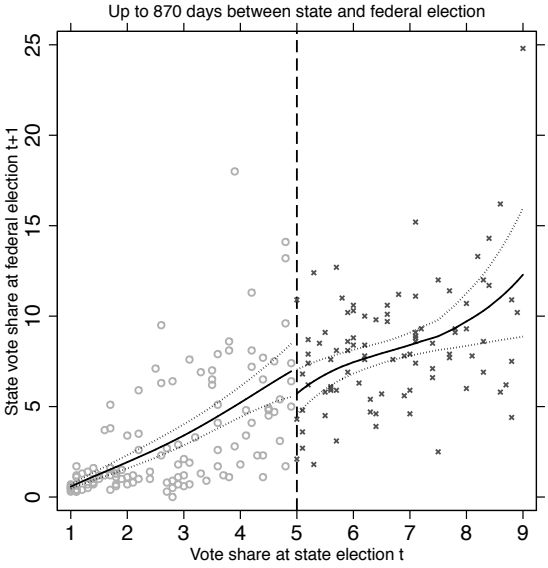
Results

Figure 1 presents three scatterplots. In each of them, the x-axis represents parties' vote share in the state election and the y-axis presents the vote share of the same parties in the subsequent federal election in the same state. Parties not having crossed the electoral threshold in the state election appear with small empty cycles and parties above this threshold are shown with small pluses. The two local linear regression curves denote the conditional expectation of parties' vote share in the federal election in the state in question, given their vote share in the prior state election. Only parties with vote share above one and below nine percent in the state election at time t are included in these graphical illustrations. The vertical line represents the cut-off point of parliamentary representation.

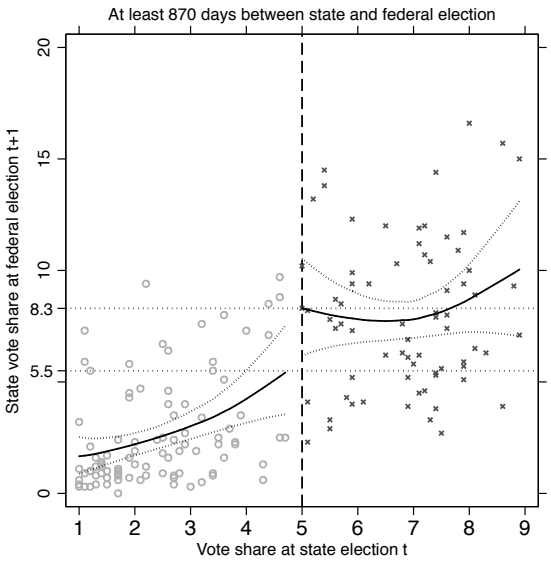
Figure 1a presents all observations within the chosen margin. There is remarkable continuity in the votes of parliamentary insiders and outsiders. No significant gap is observed. Figures 1b and 1c qualify this view, taking into consideration the time frame between the state and the federal election. In a first step we split groups with respect to the median number of days intervening between the state and the federal election, 870 days. Thus, we distinguish between cases in which the national election took place within 870 days (2 years and 4 months) after the state election (Figure 1b) and cases in which the national election



(a) Full Sample



(b) Up to 870 days between state and federal election



(c) More than 870 days between state and federal election

Figure 1: The effect of crossing the 5% threshold in a state election on a party's subsequent federal election vote share in the same state.

Note: Dots denote parties without state parliamentary representation and pluses denote parties with parliamentary representation. The local regression (black solid) curves trace the mean vote share of $Y_{i,s,t+1}$ for parties below and above the threshold and the dotted curves denote the 95% bootstrapped confidence intervals.

took place more than 870 days after the state election (Figure 1c). Finding a gap in the first case but not in the second would provide evidence for the presence of spillover effects but no support to our organizational hypothesis. Reversely, finding a greater gap in the second than the first case would lend support to our organizational hypothesis. This is what we find here:

parties that make it to the state parliament at election t enjoy higher vote shares in that state in the next national election, $t + 1$. However, this effect is confined to those cases, where a relatively long period has passed between the state and the national election.

Table 1 summarizes the findings from Figure 1. It presents the results from the local linear regression estimation. The first column of the Table uses all observations from our dataset. The last two columns classify observations into two groups: those in which the federal election and the state election where less than 870 days apart (column 2) and those where the two elections were less than 870 days apart (column 3). Thus, whereas the second column tests the spill-over hypothesis, the third column tests our organizational hypothesis. The first part of the table uses the IK bandwidth, whereas the second and the third parts of the table employ half and twice that bandwidth respectively. The fourth section employs a common fixed bandwidth ($h = 4$) to facilitate comparisons between the three columns. The fifth part of the table presents treatment effect estimates using randomization inference, based on observations directly surrounding the 5% threshold.[13]

The estimates from all different bandwidth selections point to the same conclusion. Looking at the impact of state parliamentary representation when the upcoming federal election approaches, we find no significant gap. If anything, outsiders seem to outperform insiders if the time period between the two elections is short. On the contrary, looking at those cases where the federal election comes at least two years after the state election, we find a significant effect, which seems to range between 1.8 and 2.4 percentage points.The estimates seem robust to different bandwidths.[14]. There seems to be an upward gap associated

[13] We follow the standard randomization inference procedure to calculate p-values, and confidence intervals outlined at length in Ho and Imai (2006) and in Gerber and Green (2012). The choice of the window in which our local "randomization-type condition" holds is data-driven, based on sequential testing of nested windows of different sizes starting from the smallest and ending at the largest in which the sharp null hypothesis cannot be rejected for any window contained in it (Cattaneo, Frandsen and Titiunik 2013). A detailed explanation of randomization inference, as well as the specifications and results of the window selection procedure are shown in the Appendix (Table A.1).

[14] Figure A.1 in the Appendix shows the variability in the estimates as the bandwidth widens from 2% to 7%. There is remarkable continuity in the estimated effect, which strengthens our confidence that our conclusions are not driven by the choice of the bandwidth

Table 1: The impact of state-level representation on federal vote share in the same state.

	Full Sample	# days $\leq$ 870	# days > 870
		IK Bandwidth	
Treatment Effect	0.71 (0.65)	-0.67 (0.98)	2.23 (1.01)**
CIs	[-0.57—1.99]	[-2.59—1.25]	[-0.24—4.22]
Bandwidth (h)	5.45	3.83	5.90
n [treated]	195	89	93
n [control]	531	103	209
		Half-IK Bandwidth	
Treatment Effect	-0.03 (0.99)	-1.14 (1.44)	2.09 (1.71)
CIs	[-1.97—1.91]	[-3.96—1.68]	[-1.26—5.45]
Bandwidth (h)	2.73	1.92	2.95
n [treated]	126	52	62
n [control]	113	47	52
		Double-IK Bandwidth	
Treatment Effect	0.87 (0.55)	0.13 (0.75)	1.86 (0.86)**
CIs	[0.87—0.55]	[-1.34—1.60]	[0.182—5.45]
Bandwidth (h)	10.90	7.66	11.79
n [treated]	238	112	111
n [control]	531	324	209
		Fixed Bandwidth ($h = 4$)	
Treatment Effect	0.37 (0.81)	-0.61 (0.96)	2.30 (1.39)*
CIs	[-1.23—1.96]	[-2.50—1.27]	[-0.42—5.02]
Bandwidth (h)	4	4	4
n [treated]	161	90	73
n [control]	199	112	88
		Randomization Inference	
Treatment Effect	1.65 [0.28—3.03]**	0.78 [-0.85—2.44]	2.39 [-0.13—4.95]*
Window	1%	0.9%	1.1%
n [treated]	56	30	23
n [control]	35	25	10
		Placebo Outcome, $V_{s,t-1}$: # days > 870	
	IK bandwidth	Half-IK Bandwidth	Double-IK Bandwidth
Treatment Effect	0.41 (0.65)	0.52 (0.84)	0.31 (0.57)
CIs	[-0.86—1.68]	[-1.13—2.17]	[-0.81—1.43]
Bandwidth (h)	6.53	3.26	13.05
n [treated]	89	59	100
n [control]	154	52	154

Note: The entries denote the treatment effect (τ) of entering the state parliament on the party's vote share in the same state in the next federal election. *p<.10; **p<.05, two-tailed tests.

with state parliamentary representation but only in those cases where the federal election takes place at least two years after the state election. This result is clearly consistent with the idea that parliamentary representation in the state legislature may help parties' vote share in

the next federal elections.[15]

The last part of Table 1 presents a simple placebo test. If the effect identified here is due to parliamentary representation, we should find no similar gap between insiders and outsiders when looking at parties' vote share at the previous federal election, $t-1$. Finding such an effect might signal the presence of unobserved heterogeneity generating path dependence in patterns of electoral support in state and federal elections. As indicated in the last rows of the table, however, no such effect is observed, irrespective of the chosen bandwidth. The results are very similar in all three columns.[16]

As a next step, we engage into a more detailed exploration of the relationship between state parliamentary representation and the timing of the federal election. Figure 2 presents the magnitude of the effect conditional on the distance between the state and the next federal election.[17] The black solid curve denotes the estimated difference in the vote share of a

[15] In the Appendix (Table A.3) we extend this evidence, using as a dependent variable the difference between the party's vote share in the state minus its overall vote share in federal election, $t+1$. This dependent variable allows us to account for possible time-trends hitting parties in a homogenous way across all states. The results from this exercise are substantively identical to those presented here. Moreover, Table A.2 of the Appendix shows that the results remain substantively intact when excluding the years 1959-1969, in which the threshold for participating in the official public financing system for political parties coincided with the representation threshold, from the analysis.

[16] Since our main interest lies in the subset of observations with more than 870 days between state and federal election (column 3), we focus only on those observations when implementing the placebo test. The results are substantively identical when all observations are included in the analysis.

[17] The estimates stem from a linear regression model that includes only observations within the chosen window. It expands the local linear specification by interacting all terms with a distance counter measuring the number of days betweeen the two elections. We also include a second polynomial to allow for non-linearities in this pattern:

$$\begin{aligned} Y_{i,s,t+1} = \beta_0 + \beta_1 D_{i,s,t} + \beta_2 (X_{i,s,t} - c) + \beta_3 D_{i,s,t} (X_{i,s,t} - c) + \\ \beta_4 Distance_{i,s} + \beta_5 D_{i,s,t} Distance_{i,s} + \beta_6 (X_{i,s,t} - c) Distance_{i,s} + \\ \beta_7 (X_{i,s,t} - c) D_{i,s,t} Distance_{i,s} + \beta_8 Distance_{i,s}^2 + \beta_9 D_{i,s,t} Distance_{i,s}^2 + \\ \beta_{10} (X_{i,s,t} - c) Distance_{i,s}^2 + \beta_{11} (X_{i,s,t} - c) D_{i,s,t} Distance_{i,s}^2 + \epsilon_{i,s,t+1}, \end{aligned}$$

Based on these estimates, we obtain predictions for the expected vote share of a parliamentary insider, as extrapolated at the point of the threshold, conditional on the the distance between state and federal election. We do so by collecting estimates for various values of

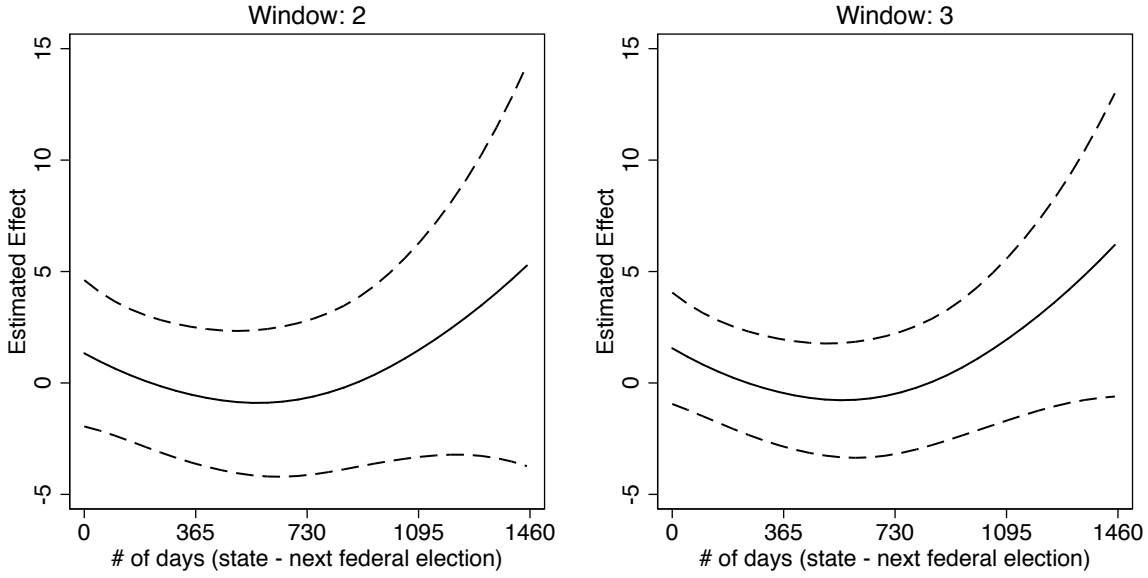


Figure 2: A closer look at the distance between state and federal election and the impact of state representation on federal vote share.

Note: The black solid curve presents the average difference between state parliamentary insiders and outsiders in their state vote share in the next federal election. The dotted curves present the 95% CIs.

marginal state parliamentary insider compared to an outsider. The non-linear but monotone ascending trend along the horizontal axis indicates that the gap between the two groups of parties is widened as the distance between state and federal election increases. In the first panel, we use only a window of two percentage points, thus including observations with higher a vote share than 3% and lower than 7% in the state election. In the second panel, we include only parties with a vote share higher than 2% and lower than 8% percentage points. The overall pattern is almost identical.[18] Although the level of uncertainty accompanying these estimates is high, the results highlight the conditioning role of the timing of a federal election with respect to state parliamentary representation. We find hardly any effect within the first two years, but remarkable (though imprecisely estimated) effects when the two elections are more than two years apart. Moreover, the longer the distance, the greater this effect appears to be. The overall pattern suggests that parliamentary representation needs time to produce electoral fruits, but eventually pays off.

$Distance_{i,s}$ from the following expression: $\beta_1 + \beta_5 Distance_{i,s} + \beta_9 Distance_{i,s}^2$.

[18] In the Appendix (Figure A.1) we show that the same pattern is observed with wider windows.

Robustness checks and diagnostics

Four potential criticisms need to be addressed at this point. First, we have thus far assumed that our observations around the threshold are exchangeable. This assumption might be violated if some parties can manipulate their vote share so as to be just below or (more likely) above the threshold. Although in free and fair elections it is hard to see how this can be done, previous studies on incumbency advantage have suggested possible sorting mechanisms (Caughy and Sekhon 2011).[19] We test for sorting in the forcing variable using the McCrary (2008) density test. The test refutes this possibility. The p-values for the null of no-sorting are 0.58 and 0.97 when using a bandwidth $h = 4$ and the IK optimal bandwidth ($h = 3.84$) respectively.[20] Figure A.2 in the Appendix illustrates the absence of sorting in parties' vote shares in state elections either above or below the threshold.

Second, it is important to check for covariate balance at the cutoff point. It might be that some parties, years or states are over- or under-represented in each part of the cut-off point, confounding our estimates. To see whether this is the case, we examine whether a significant gap at the cutoff point is observed in various seemingly irrelevant outcomes. As outcome variables we use a set of dummy variables —for each of the 16 states, Eastern region, FDP party, Green party, and extreme right party— as well as year, both as a continuous variable and as a pre/post-1989 binary outcome. The results are shown in Table 2. In total we implement 21 balance tests and none of them proves marginally significant.[21] Importantly, neither the FDP, nor the Green party dummies are statistically significant. This result helps us discount the possibility that the two larger parties, CDU and SPD, strategically "lend" votes to

[19] That said, it is clear that our design is not based on the marginality of the election. Any type of election could provide us with parties near the 5 percent threshold. Thus, the criticisms related to studies using the margin of victory as the forcing variable do not apply in our design.

[20] The test has been implemented only for state elections that take place more than 870 days before the next federal election, since it is the results from this analysis we need to scrutinize.

[21] In the Appendix (Table A.4), we repeat this procedure focusing only on the observations with more than 870 days between state and federal election. The results are very similar to those reported here.

potential coalition partners in order to aid them in crossing the 5%-threshold. While this type of strategic voting might occur occasionally, there is no evidence that the FDP in particular sorts on the right side of the representation threshold.

Table 2: Balance Checks

	RD Estimates	95% Lower Bound	95% Upper Bound	IK Bandwidth
Extreme right	-0.041 (0.010)	-0.237	0.154	2.74
East Germany	0.091 (0.102)	-0.109	0.290	2.51
Year	0.670 (3.436)	-6.035	7.434	6.23
Post-1989	0.101 (0.137)	-0.167	0.369	3.04
Greens	0.230 (0.148)	-0.061	0.520	1.90
FDP	-0.039 (0.148)	-0.329	0.251	2.51
Baden-Wuttemberg	-0.011 (0.052)	-0.091	0.113	2.77
Bayern	0.041 (0.095)	-0.146	0.228	2.79
Berlin	0.051 (0.045)	-0.038	0.139	2.38
Brandenburg	0.025 (0.018)	-0.010	0.061	2.26
Bremen	0.025 (0.064)	-0.101	0.150	2.69
Hamburg	-0.191 (0.125)	-0.437	0.055	2.22
Hessen	0.005 (0.073)	-0.138	0.147	2.35
Mecklenburg-Vorpommern	-0.020 (0.056	-0.129	0.089	2.30
Niedersachsen	-0.100 (0.084)	-0.264	0.065	2.72
Nordrhein-Westfalen	0.042 (0.089)	-0.132	0.216	2.45
Rheinland-Pfalz	-0.014 (0.081)	-0.173	0.145	2.38
Saarland	0.083 (0.057)	-0.029	0.195	2.14
Sachsen	0.076 (0.053)	-0.029	0.181	2.15
Sachsen-Anhalt	0.075 (0.054)	-0.0315	0.182	1.58
Schleswig-Holstein	0.025 (0.096)	-0.164	0.213	2.59
Thueringen	-0.046 (0.053)	-0.150	0.058	2.01

Note: The entries in the first column denote the treatment effect (τ) of entering the state parliament on various placebo outcomes. Standard errors in parentheses. *p<.10; **p<.05, two-tailed tests.

Coalition dynamics in Germany open one further possibility that needs addressing: Although participation in coalition government is necessarily a consequence of parliamentary representation, we concede that our theorized mechanism does not necessarily rest on participation in state government. To be sure, parties in state government should be given even more privileged access than opposition parties to the organizaitonal benefits alluded to above.

However, we believe that being part of the state executive is not a necessary condition for having access into organizational resources. Consequently, we expect that our results are not driven only by parties that by entering into the legislature become also coalition partners. To test whether this is the case, as a third robustness check we replicate the analysis in Table 1 excluding the liberal party FDP from the analysis. The FDP has been the main party of government entering frequent coalitions with both CDU and SPD at state and federal level. If results were merely driven by participation in governments, excluding all FDP cases should lower the estimated effect of parliamentary representation. However, as Table A.5 in the Appendix shows, estimates are actually larger when excluding the major party of government.

Fourth, we check whether there are within-group jumps. If the reason for the observed gap between insiders and outsiders is parliamentary representation, we should not find significant discontinuities in the dependent variable among observations that do not differ in terms of their treatment value. To detect such discontinuities, we focus separately on parties below and on parties above the 5% electoral threshold. We then split each group in two subgroups, treating parties with higher vote shares as treated and parties with lower vote shares as controls.

To maximize statistical power, the benchmark for this within-group division is the median. Thus, among parties with $0 < X_{i,s,t} < 5$, none of whom achieves parliamentary representation, those below the median are compared to those above the median. The same comparison takes place separately for parties with $5 < X_{i,s,t} < 10$, all of whom enter the state parliament. We also divide parties by using the midpoint of the scale, i.e. 2.5% for the group below the threshold and 7.5% for the group above the threshold. For each analysis we use two different bandwidths. The results appear in Table 3. We find no instance of within-group jumps.

Table 3: Robustness Check: Within-Group Jumps

	Within-Group Jumps		Within-Group Jumps	
	$c = 0.8$ (Median)	$c = 2.5$	$c = 7.2$ (Median)	$c = 7.5$
	IK Bandwidth			
Treatment Effect	0.153 (0.234)	-0.662 (0.844)	0.206 (1.489)	-1.074 (2.013)
95% CIs	[-0.306—0.612]	[-2.317—0.992]	[-2.712—3.124]	[-5.020—2.879]
Bandwidth (h)	0.933	2.505	1.833	0.901
n [treated]	46	83	33	17
n [control]	102	102	32	23
	Bandwidth: $h = 2.5$			
Treatment Effect	0.277 (0.295)	-0.660 (0.846)	0.494 (1.304)	0.424 (1.314)
95% CIs	[-0.301—0.855]	[-2.32—0.996]	[-2.061—3.050]	[-2.151—3.000]
n [treated]	83	45	43	32
n [control]	102	147	38	47

Note: The first two columns compare parties with $c \leq X_{i,s,t} < 5\%$ and parties with $0\% < X_{i,s,t} < c$. None of these parties has achieved parliamentary representation. The last two columns compare parties with $5\% \leq X_{i,s,t} < c$ and parties with $c \leq X_{i,s,t} < 10\%$. All parties enter the parliament in the last two colums. We use the median vote share of parties with $0\% < X_{i,s,t} < 5\%$ and with $5\% < X_{i,s,t} < 10\%$ to create the comparison groups in columns one and three respectively. We limit our observations to the 0-10% interval. We also use the mid points, i.e. $c = 2.5$ for parties with $X_{i,s,t} < 5\%$ and $c = 7.5$ for parties with$X_{i,s,t} > 5\%$.

Evidence regarding the mechanism

The results presented thus far indicate that representation in the state parliament offers a comparative advantage in the subsequent federal election. The fact that this effect needs time to materialize points to its organizational nature, but does direct evidence support that inference? In addressing that question we first consider the financial resources available to parliamentary parties in each state. It has been widely reported that some German states have been more generous when it comes to the appropriation of public money to parliamentary parties than other states (von Armin 2011). Does subnational parliamentary representation lead to higher national vote shares in states with higher per capita parliamentary party funding?

The absolute amount of state funding made available to parties present in state parliament can be found in the yearly state budgets. We divide the absolute amount of state funding by the number of parties present in the legislature and by the number of inhabitants in each state to get an idea about how many additional Euro a parliamentary party has at its disposal to

spend on organization, service delivery, and on promoting its political goals.[22] The amount of state funding a parliamentary party receives per inhabitant ranges from 10 cents in Baden-Wuerttemberg to 1.7 Euros in Bremen.

We divide our sample into two groups based on whether states fall above or below the median in parliamentary party funding per inhabitant and repeat the analysis on these two subsamples. This step will show whether the consequences of state-level parliamentary election vary depending upon the general level of a state's financial support for political parties. The results displayed in Table 4 strongly support the organizational mechanism, showing that insiders benefit more in electoral terms in those states where parliamentary funding as a rule is more generous.

Table 4: The Role of Parliamentary Benefits

	States with low levels of parliamentary party funding	States with high levels of parliamentary party funding
	IK Bandwidth	
Treatment Effect	2.094 (1.201)*	3.920 (1.685)**
95% CIs	[-0.260—4.448]	[0.617—7.222]
Bandwidth	7.17	7.35
n [treated]	80	20
n [control]	135	72
	Fixed Bandwidth ($h = 4$)	
Treatment Effect	2.203 (1.667)	5.340 (2.180)**
95% CIs	[-1.064—5.470]	[1.067—9.613]
n [treated]	59	12
n [control]	48	39

Note: The analysis distinguishes between the eight states with lower (left column) and the eight states with higher (right column) parliamentary funding. We focus only on those cases where the distance between the state and the federal election is at least 870 days, because it is for these cases that we find an effect of parliamentary representation. Using half or double of the IK bandwidth produces substantively identical results. *p<.10; **p<.05, two tailed tests. We cannot employ randomization inference for this subgroup analysis since the number of observations falling within the appropriate window would be miniscule.

Second, we explore the effects of state parliamentary representation on future state party membership. If the organizational mechanism holds, parties that enter state parliament

[22] We use data from the 2010 fiscal year because this is the earliest year all German states have made their budgets available online. Where data is available from earlier years, we check if the observed rankings of states hold, and it looks like there is a more general pattern of some states providing more generous parliamentary party funding than others. The 2010 data used to rank states according to generosity of parliamentary party funding can be found in Table A.6 in the Appendix.

should attract more state party activists than parties that fail to pass the 5%-threshold. A growth in the membership base after entering state parliament would thus be evidence for the existence of increasing organizational capabilities in the state.

In order to test if a party attracts more members in those states, where it enters state parliament, we use yearly data on FDP and Green Party membership in the period between 1990 and 2010.[23] Those 20 years were a turbulent period for both parties, in which they lost and gained representation in several state parliaments. Often only a few 1000 votes decided the parties' fates in a state election.

We analyze these data employing again the discontinuities stemming from the 5% threshold. We thus compare the membership rates of the FDP and the Greens in cases where they (marginally) enter the state parliament to cases in which they are (marginally) left out. Since our membership data vary by year, we can examine the impact of entering the state parliament at year t on parties' membership rates at every year until the next state election. Our outcome of interest thus becomes state party membership (the number of members per 1 million inhabitants) in years $t+1$, $t+2$, $t+3$ and $t+4$. We use the IK bandwidth, half and twice this bandwidth, a common fixed bandwidth of $h = 4$, and local randomization inference to overcome potential small-sample problems arising from the fact that there are only few observations exactly below and above the cutoff point for the shorter period between 1990 and 2010.[24]

The results, reported in Table 5, illustrate the way parliamentary representation yields significant organizational benefits. In cases where the two parties marginally cross the threshold, they have on average around 600 more members per million inhabitants than in cases in which they marginally fail to cross the threshold. Despite the uncertainty surrounding this estimate, the effect of parliamentary representation on future state party membership is robust to different bandwidths. Using the $\pm 1.3\%$ window around the cutoff point, in which

[23] Our use of the specific parties and time period is only due to data availability reasons. These were the only reliable data we could find on party membership in Germany.

[24] We again engage in sequential testing of windows from the smallest to the largest to find that the largest window within which the local randomization assumption holds is within 1.3 percentage points on either side of the representation threshold.

the local randomization assumption seems to hold, randomization-based inference provides a very similar picture. In the last section of the Table, we further explore the difference in membership rates before and after a state election, conditional on parliamentary status. Parties that make it into the state parliament gain approximately 100 more members in the first two years after the election, compared to when they fail to cross the 5% threshold. In line with our hypothesized mechanism, this number increases over time and until the next state election. In absolute terms, this means that if the FDP enters the state parliament of Northrhine-Westphalia (the largest state in Germany of around 18 million inhabitants), it gains around 2,900 members in the course of the parliament compared to the year before the election. This amounts to around 17% of its average overall membership in this state.

The sixth part of Table 5 tests whether the effect attributed to state parliamentary presence is due to state-specific unobserved heterogeneity. If we simply capitalize on differential membership rates of both parties across states, we should observe the same gap in the years preceding the state election. We test this possibility and find no significant difference in the membership rates of parties at time $t-1$, $t-2$, $t-3$, and $t-4$ according to whether they entered parliament in election t. Apart from not attaining statistical significance at any conventional level, the estimates are of much lower magnitude, and are also changing sign at $t-4$.

When combined, these sets of findings suggest that the membership effects identified here are due to parties' entrance into the state parliament. Gaining more members, parties have more resources to organize on the ground. More members means more people to distribute leaflets, put up signs and talk to their friends come Election Day. It seems highly plausible that these figures have an impact on the electoral fortunes of the party in national elections.

Table 5: The impact of state-level representation on party membership, FDP and Greens, 1990-2010

	$Members_{t+1}$	$Members_{t+2}$	$Members_{t+3}$	$Members_{t+4}$
IK Bandwidth				
Treatment Effect	720.0**	651.8**	569.6*	575.5*
	(361.5)	(329.6)	(305.6)	(306.1)
Bandwidth (h)	3.449	3.567	4.254	4.266
n [treated]	62	60	61	60
n [control]	48	48	49	49
Half IK Bandwidth				
Treatment Effect	641.3	577.4	553.3	537.9
	(419.4)	(382.3)	(362.0)	(357.9)
Bandwidth (h)	1.725	1.784	2.127	2.133
n [treated]	36	35	40	39
n [control]	23	23	26	26
Double IK Bandwidth				
Treatment Effect	651.9**	636.7**	609.4**	604.2**
	(303.4)	(269.8)	(246.0)	(245.6)
Bandwidth (h)	6.898	7.134	8.508	8.531
n [treated]	84	83	84	83
n [control]	50	50	49	49
Fixed Bandwidth ($h = 4$)				
Treatment Effect	709.5**	642.9**	583.4*	590.7*
	(350.6)	(321.2)	(312.2)	(312.8)
n [treated]	66	64	60	59
n [control]	50	50	49	49
Randomization Inference ($Window = 1.3\%$)				
Treatment Effect	545.7*	515.6*	466.2*	481.32*
	[-173.6—1145.9]	[-134.1—1041.0]	[-171.0 —940.8]	[-157.8— 954.4]
n [treated]	31	30	30	29
n [control]	15	15	14	14
Placebo Tests:	$Members_{t-1}$	$Members_{t-2}$	$Members_{t-3}$	$Members_{t-4}$
Fixed Bandwidth ($h = 4$)				
Treatment Effect	275.0	161.1	51.9	-157.0
	(344.9)	(383.6)	(457.7)	(566.6)
n [treated]	55	49	46	46
n [control]	47	46	46	45
	Change in Members from $t-1$ *to*:			
($h = 4$)	$t+1$	$t+2$	$t+3$	$t+4$
Treatment Effect	81.72	130.11*	150.64*	161.89
	(54.05)	(70.19)	(90.91)	(109.4)
n [treated]	54	52	48	47
n [control]	47	47	46	46
	Randomization Inference : Change in Members from $t-1$ *to*:			
($Window = 1.3\%$)	$t+1$	$t+2$	$t+3$	$t+4$
Treatment Effect	99.3**	140.8**	152.3**	170.3*
	[15.4—188.1]	[25.4—258.0]	[5.1—300.7]	[-19.0—352.2]
n [treated]	22	21	21	20
n [control]	14	14	13	13

The entries denote differences in members per 100,000 inhabitants according to whether the party achieved state parliamentary representation. *p<.10; ** p<.05, two tailed tests. State party membership data source: Niedermayer, Oskar (2011). Parteimitglieder in Deutschland: Version 2011. Arbeitshefte aus dem Otto-Stammer-Zentrum, Nr. 18.

Discussion

Subnational parliamentary representation can provide a political advantage for parties that contest nationwide elections if the national election does not follow immediately after the successful subnational election. Our results indicate that after 3 or 4 years of crossing the representation threshold in German state parliaments, the electoral advantage in federal elections amounts to around 2.3 percentage points.

What explains this effect? Evidently, it is quite difficult to reconcile this lagged effect of state parliamentary representation with the idea of bandwagon effects or positive media coverage in the immediate aftermath of a successful election. In contrast, the only mechanism that seems to fit these results is institutional, and organizational in nature. Parties that are present in state parliaments can offer the material and status benefits of a political career, hire full time staff and use generous funds to build their party organization. The party is therefore better able to attract activists and appeal to the electorate. Activists who set their eyes on a political career find opportunities to get ahead and can in turn be employed by the party for mobilization and persuasion efforts. Voters can address their concerns to state representatives, whose offices provide a range of services to them. However, addressing grievances within the electorate takes time, as does the use of parliamentary funds to build an effective organization. This finding has four important implications that merit some elaboration.

The first implication alludes to the role of subnational institutions as moderators of top-down effects. The more institutionalized a party, the less disruptive the impact of environmental uncertainty (Panebianco 1988, 109). If national trends work against a party, electoral losses will be moderated in those states in which the party has been present in state legislatures. The same holds for gains parties make thanks to of national factors. In the long run, this pattern can lead to the geographical sorting of small parties.

Second, the results presented here directly challenge the consensus, according to which politics in decentralized political systems can be described as a one-way street from the higher to the lower level. Such a characterization is overly simplistic. As subnational entities gain more powers, and subnational parliamentary parties gain more resources, political

parties gain an electoral advantage from having access to these resources. Since it is easier for smaller political parties to access indirect funding provided to parliamentary parties at the subnational than at the national level, future research can assess whether the multi-level structure of the state might contribute to explaining the varying success of smaller and new parties across established democracies. Although most of the literature on the emergence of new parties has so far concentrated on the proportionality of the electoral system, the degree of decentralization might be another factor that might warrant consideration from researchers.

Third, this paper offers strong support for a crucial assumption of cartelization theory, namely that political parties benefit from being inside the political system, and inside subnational assemblies. However, the mechanism driving this advantage is not necessarily dependent on the systems of direct public party financing. Parties benefit because the state pays for their parliamentary work, because they use state resources to organize, and because access to the state provides parties with high-status and materially lucrative positions to distribute amongst its leadership. In light of these results and previous work, it seems that the mechanism of how established parties benefit in Western democracies is more complex than the debate about official public party financing suggests.

Fourth, the paper calls for a qualification of the dominant view that subnational elections are simply imperfect reflections of party competition at the national level. Previous studies have alluded to such a top-down process either when examining elections at lower (Schakel and Jeffery 2012) or higher (van der Eijk and Franklin 1996) level than the national level. The resulting theory of second-order elections stated that voters often transfer their national voting decision rules into subnational elections. This study reveals that even if this is the case, the final outcome of subnational elections may have indirect effects. By getting access to state funds, parliamentary parties have the opportunity to become more visible and useful to their electorate at different levels of government. Previous evidence also from Germany suggests that voters' gratitude to local services (Bechtel and Hainmueller 2011) lasts for at least one election term. Thus, when there is time for this organizational advantage to materialize, parties that have gained access to the decentralized state eventually transfer this

benefit from the subnational to the national domain.

Although our empirical results are based on data from German state and federal elections, our theoretical predictions should also apply to other decentralized political systems. The conditions are the existence of an integrated party system across multiple levels of government, and subnational representative bodies that provide access to substantial state resources. Therefore, as subnational entities continue to gain more competencies over the center, one can deduce that apart from the well-documented process of the nationalization of subnational politics, a parallel process of decentralization of national politics is also taking place. As decentralization advances in most established democracies, the political machinery created through representation in subnational institutions gradually transforms national politics by bringing subnational dynamics back into the national political arena.

References

Aldrich, John H. 1995. *Why Parties? The Origin and Transformation of Political Parties in America*. Chicago: University of Chicago Press.

Bartels, Larry M. 1985. "Expectations and Preferences in Presidential Nominating Campaigns." *American Political Science Review* 79(3):804–815.

Bechtel, Michael. 2012. "Not Always Second Order: Subnational Elections, National-level Vote Intentions, and Volatility Spillovers in a Multilevel Electoral System." *Electoral Studies* 31(1):804–815.

Bechtel, Michael and Jens Hainmueller. 2011. "How Lasting is Voter Gratitude? An Analysis of the Short- and Long-Term Electoral Returns to Beneficial Policy." *American Journal of Political Science* 55(4):852–868.

Benz, Arthur. 2009. "Ein gordischer Knoten der Politikwissenschaft? Zur Vereinbarkeit von Foederalismus und Demokratie." *Politische Vierteljahresschrift* 50(1):3–22.

Bundestag, Deutscher. 2011. "Festsetzung der staatlichen Mittel für das Jahr 2010.".

Bundeswahlleiter. 2009. Wahl zum 17. Deutschen Bundestag am 27. September 2009: Ergebnisse und Vergleichszahlen frueherer Bundestags-, Europa- und Landtagswahlen sowie Strukturdaten fuerr die Bundestagswahlkreise. Technical report Wiesbaden.

Campbell, James. 1991. "The Presidential Surge and its Midterm Decline in Congressional Elections, 1868–1988." *Journal of Politics* 53(2):477–487.

Cattaneo, Matias D., Brigham Frandsen and Rocio Titiunik. 2013. "Randomization Inference in the Regression Discontinuity Design: An Application to Party Advantages in the U.S. Senate." University of Michigan.

Caughy, Devin and Jasjeet Sekhon. 2011. "Elections and the Regression Discontinuity: Lessons from Close U.S. House Races, 1942-2008." *Political Analysis* 29(4):385–408.

Cox, Garry W. 2010. "Swing voters, core voters, and distributive politics." In *Political Representation*, ed. Ian Shapiro, Susan Stokes, Elizabeth Jean Wood and Alexander S. Kirshner. Cambridge: Cambridge University Press pp. 342–357.

Deschouwer, Kris. 2003. "Political Parties in Multi-Layered Systems." *European Urban and Regional Studies* 10(3):213–226.

Detterbeck, Klaus. 2012. *Multi-Level Party Politics in Western Europe*. Basingstoke: Palgrave Macmillan.

Dinkel, Rainer. 1977. "Der Zusammenhang zwischen Bundes- und Landtagswahlergebnissen." *Politische Vierteljahresschrift: Zeitschrift der Deutschen Vereinigung für Politische Wissenschaft* 18(2/3):349–359.

Duebber, Urlich. 1962. *Parteifinanzierung in Deutschland*. Koeln und Opladen: Westdeutscher Verlag.

Erikson, Robert S. and Filippov Mikhail. 2001. "Electoral Balancing in Federal and SubNational Elections: The Case of Canada." *Constitutional Political Economy* 12(4):312–331.

Fan, Jianqing and Irene Gijbels. 1992. "Variable Bandwidth and Local Linear Regression Smoothers." *The Annals of Statistics* 20(4):2008–2036.

Fisher, Justin, David Denver and Gordon Hands. 2006. "Party Membership and Campaign Activity in Britain: The Impact of Electoral Performance." *Party Politics* 12(4):505–519.

Gabriel, Oscar W. 1989. "Federalism and Party Democracy in West Germany." *Publius* 19(4):65–80.

Gerber, Alan and Donald P. Green. 2012. *Field Experiments: Design, Analysis, and Interpretation*. New York: WW Norton.

Goidel, Robert K. and Todd G. Shields. 1994. "The Vanishing Marginals, the Bandwagon, and the Mass Media." *Journal of Politics* 56(3):802–810.

Green, Donald P., Terence Y. Leong, Holger L. Kern, Alan S. Gerber and Christopher W. Larimer. 2009. "Testing the Accuracy of Regression Discontinuity Analysis Using Experimental Benchmarks." *Political Analysis* 17(4):400–417.

Hainmueller, Jens and Holger Lutz Kern. 2005. Incumbency Effects in German and British Elections: A Quasi- Experimental Approach. Public economics EconWPA.

Ho, Daniel E. and Kosuke Imai. 2006. "Randomization Inference with Natural Experiments: An Analysis of Ballot Effects in the 2003 California Recall Election." *Journal of the American Statistical Association* 101(475):888–900.

Imbens, Guido and Karthik Kalyanaraman. 2011. "Optimal Bandwidth Choice for the Regression Discontinuity Estimator." *Review of Economic Studies* 79(3):933–959.

Imbens, Guido and Thomas Lemieux. 2008. "Regression Discontinuity Designs: A Guide to Practice." *Review of Economic Studies* 142(2):615–635.

Iyengar, Shando. 1990. "The Accessibility Bias in Politics: Television News and Public Opinion." *International Journal of Public Opinion Research* 2(1):1–15.

Jeffrey, Charlie and Dan Hough. 2002. "The Electoral Cycle and Multi-Level Voting in Germany." *German Politics* 10(2):73–98.

Jérôme, Bruno and Michael Lewis-Beck. 1999. "Is Local Politics Local? French Evidence." *European Journal of Political Research* 35(2):181–197.

John, Stefanie and Thomas Poguntke. 2012. "Party Patronage in Germany." In *Party Patronage and Party Government in European Democracies*, ed. Petr Kopecký, Peter Mair and Maria Spirova. Oxford: Oxford University Press pp. 342–357.

Katz, Richard S. and Peter Mair. 1994. *How Parties Organize: Change and Adaptation in Party Organizations in Western Democracies*. London: Sage Publising.

Katz, Richard S. and Peter Mair. 1995. "Changing Models of Party Organization and Party Democracy: The Emergence of the Cartel Party." *Party Politics* 1(1):5–28.

Katz, Richard S. and Peter Mair. 1997. *Party Organisation, Party Democracy and the Emergence of the Cartel Party*. Oxford: Oxford University Press.

Katz, Richard S. and Peter Mair. 2009. "The Cartel Party Thesis: A Restatement." *Perspectives on Politics* 7(4):753–766.

King, Anthony. 2002. *Leaders' Personalities and the Outcomes of Democratic Elections*. Oxford: Oxford University Press.

Kopecký, Petr, Peter Mair and Maria Spirova. 2012. *Party Patronage and Party Government in European Democracies*. Oxford: Oxford University Press.

Lee, David S. and Thomas Lemieux. 2009. Regression Discontinuity Designs in Economics. Working Paper 14723 National Bureau of Economic Research.

Levitt, Steven D. and Catherine D. Wolfram. 1997. "Decomposing the Sources of Incumbency Advantage in the U.S. House." *Legislative Studies Quarterly* 22(1):45–60.

Lohmann, Susanne, David W. Brady and Douglas Rivers. 1997. "Party Identification, Retrospective Voting, and Moderating Elections in a Federal System." *Comparative Political Studies* 30(4):420–449.

Lutz Kern, Holger and Jens Hainmueller. 2006. "Electoral balancing, divided government and 'midterm'loss in german elections." *The Journal of Legislative Studies* 12(2):127–149.

Marsh, Michael. 1998. "Testing the Second-Order Election Model after Four European Elections." *British Journal of Political Science* 28(4):591–607.

McCrary, Justin. 2008. "Manipulation of the Running Variable in the Regression Discontinuity Design: A Density Test." *Journal of Econometrics* 142(2):698–714.

McLean, Iain, Anthony Heath and Bridget Taylor. 1996. "Were the 1994 Euro-and local elections in Britain really second-order? Evidence from the British election panel study." *British Elections & Parties Yearbook* 6(1):1–20.

Mughan, Anthony. 1995. "Television and Presidentialism: Australian and US Legislative Elections Compared." *Political Communication* 12(3):327–342.

Mutz, Diane. 1995. "Mechanisms of Momentum. Does Thinking Make It So?" *Journal of Politics* 57(1):104–125.

Pallarés, Francesc and Michael Keating. 2003. "Multi-Level Electoral Competition Regional Elections and Party Systems in Spain." *European Urban and Regional Studies* 10(3):239–255.

Panebianco, Angelo. 1988. *Political Parties: Organization and Power*. Cambridge: Cambridge University Press.

Pierre, Jon, Lars Svasand and Anders Widfeldt. 2000. "State Subsidies to Political Parties: Confronting Rhetoric with Reality." *West European Politics* 23(2):1–24.

Poguntke, Thomas. 1994. "Parties in a Legalistic Culture: The case of Germany." In *How Parties Organize: Change and Adaptation in Party Organizations in Western Democracies*, ed. Robert S. Katz and Peter Mair. London: Sage Publications pp. 185–215.

Porter, Jack. 2003. "Estimation in the Regression Discontinuity Model." Department of Economics, University of Wisconsin.

Reif, Karl Heinz and Hermann Schmitt. 1980. "Nine Second-Order Elections–A Conceptual Framework for the Analysis of European Election Results." *European Journal of Political Research* 8(1):3–44.

Rodden, Jonathan. 2004. "Comparative Federalism and Decentralization: On Meaning and Measurement." *Comparative Politics* 36(4):481–500.

Rodden, Jonathan and Eric Wibbels. 2010. "Dual accountability and the nationalization of party competition: Evidence from four federations." *Party Politics* 17(4):629–654.

Scarrow, Susan. 2006. "Party subsidies and the freezing of party competition: Do cartels work?" *West European Politics* 29(4):619–639.

Schakel, Arjan H. 2013. "Congruence Between Regional and National Elections." *Comparative Political Studies* 46(5):631–662.

Schakel, Arjan H. and Charlie Jeffery. 2012. "Are regional elections really 'second-order' elections?" *Regional Studies* 47(3):323–341.

Serra, George. 1994. "What's in It for Me?: The Impact of Congressional Casework on Incumbent Evaluation." *American Politics Research* 22(4):403–420.

Seyd, Patrick and Paul Whiteley. 1992. *Labour's Grass Roots the Politics of Party Membership*. Oxford: Clarendon Press.

Simon, Herbert. 1954. "Bandwagon and Underdog Effects and the Possibility of Election Predictions." *Public Opinion Quarterly* 18(3):245–253.

Smith, Gordon. 1979. *Democracy in Western Germany: Parties and Politics in the Federal Republic*. London: Heinemann.

Tronconi, Filippo and Christophe Roux. 2009. "The political systems of Italian regions between state-wide logics and increasing differentiation 1." *Modern Italy* 14(2):151–166.

van der Eijk, Cees and Mark N. Franklin. 1996. *Choosing Europe? The European Electorate and National Politics in the Face of Union*. Ann Arbor: The University of Michigan Press.

von Armin, Hans Herbert. 2011. *Der Verfassungsbruch: Verbotene Extra Diaten-Gefrassige Fraktionen*. Berlin: Duncker und Humbolt.

von Armin, Hans Herbert. 2013. *Die Selbstbediener: Wie bayerische Politiker sich den Staat zur Beute machen*. Munich: Wilhelm Heyne Verlag.

Whiteley, Paul and Patrick Seyd. 1998. "The Dynamics of Party Activism in Britain – A Spiral of Demobilisation?" *British Journal of Political Science* 28(113-137).

Whiteley, Paul, Patrick Seyd and Jeremy Richardson. 1994. *True Blues: The Politics of Conservative Party Membership*. Oxford: Oxford University Press.

Zaller, John. 1992. *The nature and origins of mass opinion*. Cambridge: Cambridge University Press.

Appendix

Randomization Inference

To determine the window in which the localized random assignment assumption holds, we test for sequentially nested windows whether any covariate imbalances between parties that marginally cross the 5%-threshold (the treated observations) and parties that marginally do not enter parliament (the untreated observations) are greater than those we would expect to occur from sampling variability alone (Cattaneo, Frandsen and Titiunik 2013). For the general analysis, we first extract the f-statistic regressing parliamentary presence on the following pre-treatment covariates: state vote share at time t-1, year, extreme right, and East/West. For party membership the following covariates are available: party membership at time t-1, year and East/West. We then compare the resulting f-statistic to the mean of the f-statistics we obtain after simulating local random assignment around the 5%-threshold within the chosen window 10.000 times. The sampling distribution of any test statistic (in our case the f-statistic) can be approximated by a large random sample of all possible local random assignments. The collection of f-statistics hence "represents the exact sampling distribution under the null hypothesis that no covariates have any effect on the assigned treatment" (Gerber and Green 2012, 107). The resulting p-value indicates whether we can reject the sharp null hypothesis that the covariates taken together have no effect on localized treatment assignment within the chosen window. A low p-value would therefore indicate that we can reject that imbalances could have occurred due to sampling variability. We choose a conservative significance level of α = 0.15 to test whether the local randomization assumption is rejected in each window. We start at a 0.5 percentage points window on each side of the threshold, and raise it sequentially by .1 percentage points on either side of the threshold testing symmetric windows of 1.1, 1.2, 1.3 and so forth until the sharp null hypothesis is rejected with a p-value equal or lower than 0.15. The detailed results of this exercise are shown in (Table A.1).

For the analysis we follow the procedure outlined in Ho and Imai (2006) and in Ger-

ber and Green (2012). We first calculate the differences-in-means between treated and untreated observations (within the chosen window for which the local randomization assumption holds) using the observations in our existing data set. In order to see if estimated differences in the outcome variable could simply be due to sampling variability, we simulate 10.000 hypothetical local randomization outcomes around the 5%-threshold under the assumption of the "sharp null hypothesis" that the LATE is zero for all units. Under the sharp-null missing outcomes are equal to the observed outcomes (Gerber and Green 2012, 62). We then go on to compare the differences-in-means estimate that we received from our data to the mean of all differences-in-means estimates over 10.000 hypothetical local randomization outcomes under the assumption of no treatment effect for no observation. To construct 95% confidence intervals we approximate the full schedule of potential outcomes assuming a localized version of SUTVA and that the treatment effect is the same for all observations directly surrounding the 5% threshold (Cattaneo, Frandsen and Titiunik 2013). Adapting the method of forming confidence intervals outlined in Gerber and Green to local quasi random assignment, for observations within the window below the threshold, missing "treated" values are imputed by adding the estimated LATE to the observed untreated values (2012, 66f). For treated observations, missing "untreated" values are imputed by substracting the estimated LATE from the observed treated values. We then list the estimated LATE from each local randomization in ascending order, where the 2.5th percentile marks the bottom of the 95% confidence interval, and the estimate at the 97.5th percentile marks the top (ibid). The interpretation of this confidence interval is that if we imagine a series of hypothetical reruns of our state elections under identical conditions, 95 out of 100 local randomized outcomes in the window directly surrounding the threshold will generate intervals that bracket the true LATE.

Table A.1: Randomization Inference Window Test
(Cut-Off p $\leq$ .15; chosen window in bold)

% on each side	0.5%	0.6%	0.7%	0.8%	0.9%	1.0%	1.1%	1.2%	1.3%	1.4%
Balance test: Full sample										
p-value	0.784	0.817	0.803	0.822	0.406	**0.169**	0.1073	n/a	n/a	n/a
n [treated]	25	30	34	36	45	49	50	n/a	n/a	n/a
n [control]	13	19	20	24	25	28	31	n/a	n/a	n/a
Balance test: # days $\leq$ 870										
p-value	0.800	0.521	0.572	0.395	0.361	**0.316**	0.105	n/a	n/a	n/a
n [treated]	15	19	22	23	26	30	n/a	n/a	n/a	n/a
n [control]	9	13	14	18	18	19	n/a	n/a	n/a	n/a
Balance test: # days > 870										
p-value	0.520	0.649	0.643	0.601	0.253	0.158	**0.319**	0.080	n/a	n/a
n [treated]	10	11	12	13	19	19	15	20	n/a	n/a
n [control]	4	6	6	6	7	9	10	14	n/a	n/a
Balance test: Party membership										
p-value	0.363	0.481	0.680	0.782	0.718	0.743	0.734	0.283	**0.155**	0.025
n [treated]	10	12	15	15	17	18	19	23	23	23
n [control]	6	7	7	10	11	12	12	13	14	18

The entries in the top of every section denote the p-value associated with the sharp-null stemming from balance tests across sequential windows around the 5% threshold.

Table A.2: Replicating the analysis from Table 1, excluding observations between 1959 and 1969.

	Full Sample	# days $\leq$ 870	# days > 870
		IK Bandwidth	
Treatment Effect	0.54 (0.69)	-0.58 (0.87)	2.31 (1.07)**
95% CIs	[-0.81—1.88]	[-1.53—0.371]	[0.22—4.40]
Bandwidth (h)	5.18	4.54	6.00
n [treated]	166	90	75
n [control]	505	171	189
		Half-IK Bandwidth	
Treatment Effect	-0.13 (1.03)	-1.14 (1.28)	2.28 (1.73)
CIs	[-2.14—1.88]	[-2.66—1.37]	[-1.10—5.66]
Bandwidth (h)	2.59	2.27	3.00
n [treated]	105	60	48
n [control]	102	57	45
		Double-IK Bandwidth	
Treatment Effect	0.81 (0.57)	0.16 (0.72)	1.78 (0.92)*
CIs	[-0.30—1.93]	[-1.26—1.58]	[-0.03—3.59]
Bandwidth (h)	10.36	9.08	9.00
n [treated]	204	113	87
n [control]	505	316	189
		Fixed Bandwidth ($h = 4$)	
Treatment Effect	0.24 (0.83)	-0.69 (0.97)	2.37 (1.44)*
CIs	[-1.37—1.86]	[-2.59—1.22]	[-0.45—5.18]
Bandwidth (h)	4	4	4
n [treated]	140	83	57
n [control]	187	109	78
		Randomization Inference	
Treatment Effect	1.65 [0.28—3.03]**	0.78 [-0.85—2.44]	2.39 [-0.13—4.95]*
Window	1%	0.9%	1.1%
n [treated]	56	30	23
n [control]	35	25	10
	Placebo Outcome, $V_{s,t-1}$: # days > 870		
	IK bandwidth	Half-IK Bandwidth	Double-IK Bandwidth
Treatment Effect	0.34 (0.73)	0.24 (0.94)	0.19 (0.60)
CIs	[-1.09—1.77]	[-1.60—2.07]	[-0.45—0.83]
Bandwidth (h)	6.66	3.33	13.322
n [treated]	69	46	79
n [control]	140	53	140

Note: The table replicates Table 1 of the main text, excluding all observations between 1959 and 1969, because during this period direct public financing was partially contingent upon parliamentary representation. Estimation and interpretation follow the logic explicated in the main text. *p<.10; **p<.05, two-tailed tests.

Table A.3: Accounting for Parties' National Ups and Downs: $Y \equiv X_{i,s,t} - \bar{X}_{i,t}$

	Full Sample	# days $\leq$ 870	# days > 870
		IK Bandwidth	
Treatment Effect	0.01 (0.47)	-0.74 (0.56)	1.36 (0.62)**
95% CIs	[-0.91—0.93]	[-1.83—0.35]	[0.14—2.57]
Bandwidth (h)	4.898	5.489	5.782
n [treated]	184	107	91
n [control]	405	324	207
		Half-IK Bandwidth	
Treatment Effect	0.26 (0.64)	-0.60 (0.75)	1.28 (0.74)*
95% CIs	[-0.99—1.51]	[-2.06—0.87]	[-0.18—2.73]
Bandwidth (h)	2.449	2.744	2.891
n [treated]	115	71	56
n [control]	105	67	49
		Double-IK Bandwidth	
Treatment Effect	-0.06 (0.39)	-0.75 (0.49)	1.01 (0.57)*
95% CIs	[0.82—0.71]	[-1.72—0.22]	[-0.10—2.12]
Bandwidth (h)	9.796	10.977	11.563
n [treated]	235	131	107
n [control]	531	324	207
		Fixed Bandwidth ($h = 4$)	
Treatment Effect	-0.04 (0.53)	-0.75 (0.65)	1.26 (0.71)*
95% CIs	[-1.07—0.99]	[-2.03—0.53]	[-0.19—2.65]
n [treated]	161	90	71
n [control]	199	112	87
	Placebo Outcome, $V_{s,t-1} - \bar{V}_{t-1}$: # days > 870		
	IK bandwidth	Half-IK Bandwidth	Double-IK Bandwidth
Treatment Effect	0.64 (0.81)	0.83 (1.25)	0.55 (0.68)
95% CIs	[-0.94—2.22]	[-1.62—3.28]	[-0.77—1.88]
Bandwidth (h)	5.193	2.597	10.386
n [treated]	77	44	96
n [control]	153	38	153

Note: The table replicates Table 1 of the main text, using a different dependent variable. Instead of using the vote share of the party in the next federal election in the state in question, we use the difference between its state vote share and its national average. By taking the national vote share into consideration we attempt to account for nation-wide ups and downs in parties' electoral appeal over time. *p<.10; **p<.05, two-tailed tests.

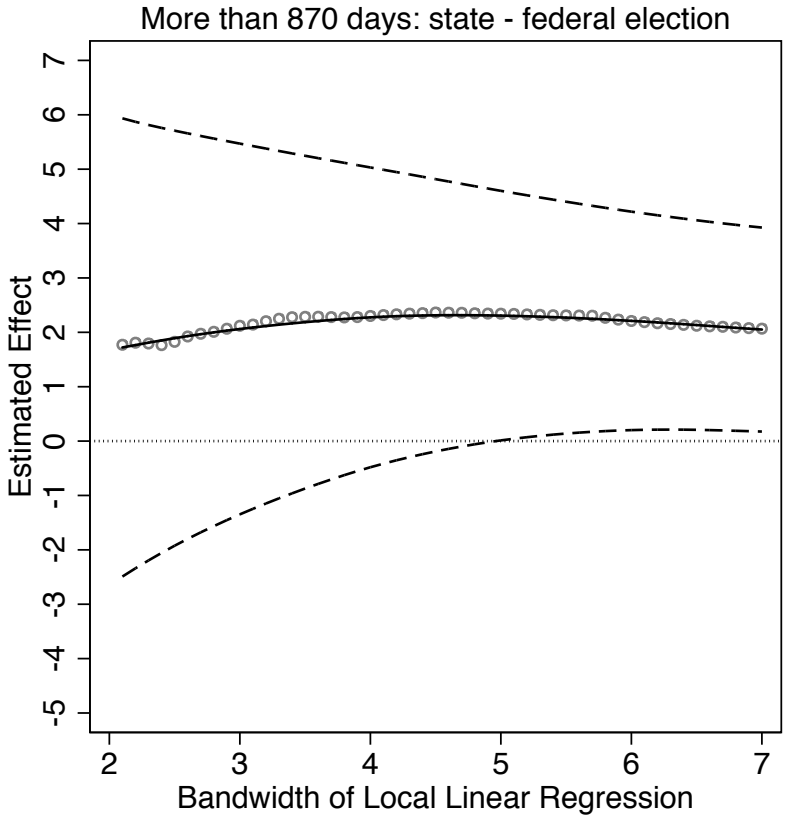


Figure A.1: Inspecting the Sensitivity of the Estimates to Different Bandwidths.
Note: The gray dots denote the actual RD estimates for various different bandwidths, as indicated by the horizontal axis. The loess solid curve and the dashed curves summarize the point estimates and their associated 95% CIs. Whereas the decrease in the number of observations widens the upper and lower bounds of the RD estimates in smaller bandwidths, the variation in their magnitude is very small.

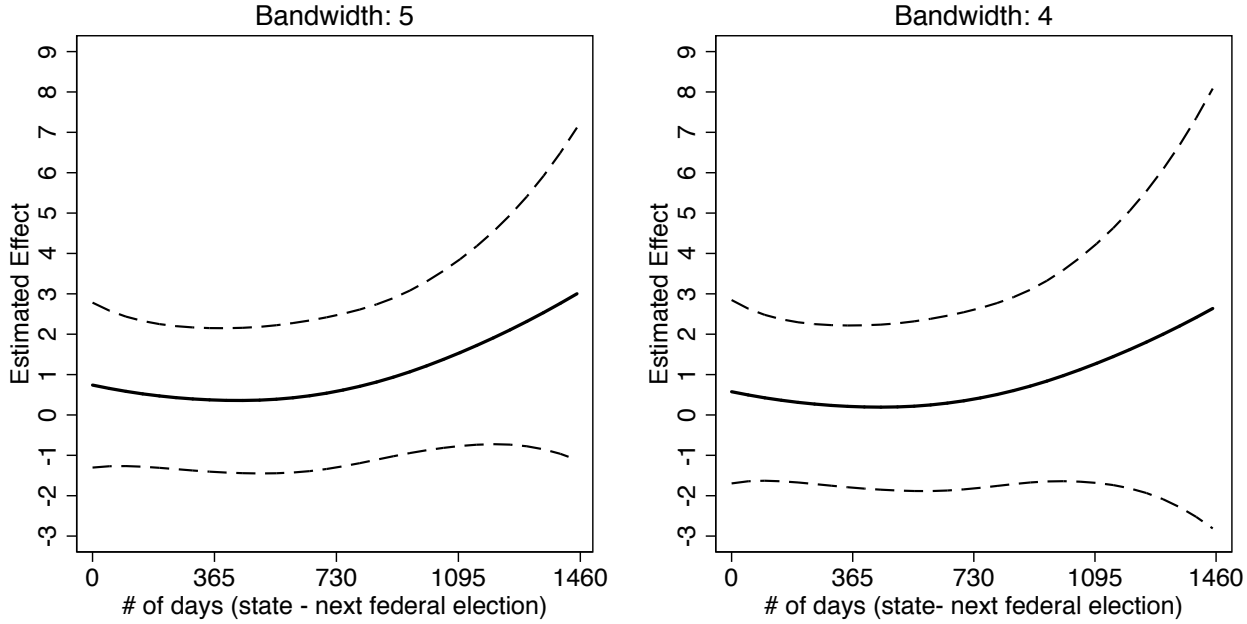


Figure A.2: Treatment Effects and Distance between State and Federal Election: Using a window of 4 and 5 percentage points.

Note: The solid curve denotes the difference in the vote share of parties that enter state parliament compared to parties without regional parliamentary representation. The estimation draws on the procedure outlined in the main text. The chosen windows define the upper and lower bounds of state vote share a given party needs to have in order to be included in the analysis.

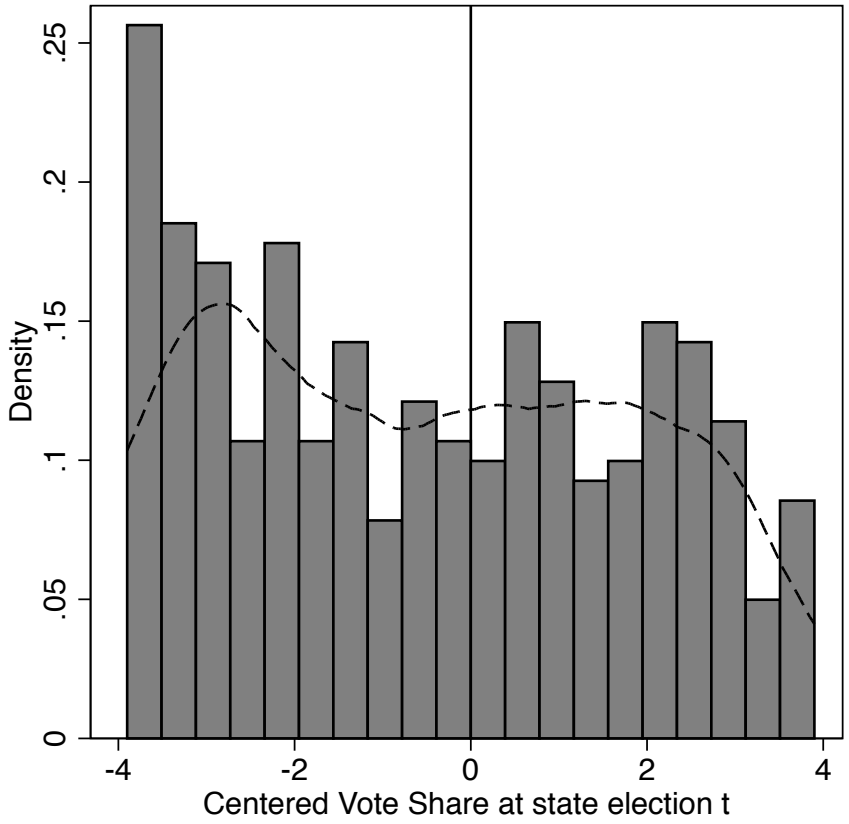


Figure A.3: A Density Inspection for Manipulation in the Forcing Variable

Note: Bin estimation determined by the McCrary sorting test. A kernel density curve illustrates the continuity in the density around the cut-off point (electoral threshold).

Table A.4: Balance Checks: Focusing on $\#Days > 870$.

	RD Estimates	95% Lower Bound	95% Upper Bound	IK Bandwidth
Extreme right	0.025 (0.161)	-0.290	0.340	2.990
East Germany	0.034 (0.107)	-0.176	0.244	3.508
Year	-5.505 (5.705)	-16.687	5.677	6.42
Post-1989	-0.097 (0.156)	-0.403	0.208	5.351
Greens	0.165 (0.217)	-0.343	0.145	9.523
FDP	0.224 (0.153)	-0.075	0.523	7.406
Baden-Wuttemberg	0.066 (0.051)	-0.035	0.167	10.930
Bayern	0.178 (0.105)*	-0.028	0.385	2.769
Berlin	0.151 (0.110)	-0.063	0.366	2.430
Brandenburg	0.017 (0.013)	-0.007	0.042	4.277
Bremen	-0.011 (0.149)	-0.304	0.282	3.221
Hamburg	-0.048 (0.055)	-0.156	0.059	2.645
Hessen	-0.157 (0.188)	-0.525	0.211	3.139
Mecklenburg-Vorpommern	-0.116 (0.065)*	-0.244	0.010	9.891
Niedersachsen	0.033 (0.178)	-0.315	0.380	1.979
Nordrhein-Westfalen	-0.150 (0.184)	-0.510	0.210	2.566
Rheinland-Pfalz	-0.060 (0.163)	-0.388	0.249	3.050
Saarland	-0.078 (0.122)	-0.317	0.160	2.745
Sachsen	0.001 (.015)	-0.029	0.030	5.433
Sachsen-Anhalt	-0.037 0.046	-0.128	0.054	10.784
Schleswig-Holstein	0.261 (0.164)	-0.060	0.583	3.098
Thueringen	0.004 (0.026)	-0.048	0.055	4.513

Note: The table replicates Table 2 of the main text, using only the subset of observations where the distance between state and next federal election is more than 870 days. Standard errors in parentheses. *p<.10; **p<.05, two-tailed tests. When excluding the two states with significant gaps (Bayern and Mecklenburg-Vorpommern), the estimated effect of parliamentary representation remains very similar to the effects reported in Table 1 of the main text as well as Tables A.2 and A.3 of the appendix: With the IK bandwidth ($h = 6.022$), $\tau = 2.301$ with 95% CIs [0.71—4.53]. With half the IK bandwidth ($h = 3.011$), $\tau = 2.300$ with 95% CIs [-1.27—5.87]. Finally, with double the IK bandwidth ($h = 12.04$), $\tau = 1.896$ with 95% CIs [0.031—3.76].

Table A.5: Replicating the analysis from Table 1, excluding the FDP.

	Full Sample	# days $\leq$ 870	# days > 870
		IK Bandwidth	
Treatment Effect	0.733 (0.810)	-0.302 (0.861)	2.436 (1.591)
95% CIs	[-0.855—2.32]	[-1.988—1.385]	[-0.682—5.554]
Bandwidth (h)	5.216	4.927	5.024
n [treated]	87	50	35
n [control]	490	253	192
		Half-IK Bandwidth	
Treatment Effect	0.538 (1.103)	-0.785 (0.970)	2.947 (2.540)
95% CIs	[-1.623—2.699]	[-2.686—1.115]	[-2.031—7.925]
Bandwidth (h)	2.608	2.463	2.512
n [treated]	54	29	22
n [control]	80	41	35
		Double-IK Bandwidth	
Treatment Effect	0.876 (0.656)	0.501 (0.754)	1.348 (1.202)
95% CIs	[-0.409—2.162]	[-0.976—1.979]	[-1.010—3.704]
Bandwidth (h)	10.433	9.853	10.047
n [treated]	115	63	52
n [control]	490	298	192
		Fixed Bandwidth ($h = 4$)	
Treatment Effect	0.670 (0.952)	-0.478 (0.936)	2.595 (1.904)
95% CIs	[-1.197—2.536]	[-2.312—1.356]	[-1.137—6.327]
n [treated]	70	40	30
n [control]	158	86	72
	Placebo Outcome, $V_{s,t-1} - \bar{V}_{t-1}$: # days > 870		
	IK bandwidth	Half-IK Bandwidth	Double-IK Bandwidth
Treatment Effect	-0.140 (1.013)	0.458 (1.384)	-0.411 (0.800)
95% CIs	[-2.125—1.845]	[-2.256—3.172]	[-1.979—1.156]
Bandwidth (h)	6.587	3.294	13.174
n [treated]	35	22	44
n [control]	138	39	138

Note: The table replicates Table 1 of the main text, excluding all observations from the FDP. *p<.10; **p<.05, two-tailed tests.

Table A.6: Parliamentary Funding per state, adjusted by number of parties in the parliament and by state population size, 2010 data

State	Parliamentary Funding per capita & per party	Total Parliamentary Expenditure (in million euros)
Baden-Wuerttemberg	0.102	5.488
Nordrhein-Westfalen	0.110	9.802
Niedersachsen	0.117	6.837
Bayern	0.234	14.673
Schleswig-Holstein	0.248	4.211
Hessen	0.276	8.383
Rheinland-Pfalz	0.331	3.973
Sachsen	0.410	10.221
Berlin	0.415	7.172
Hamburg	0.447	3.972
Brandenburg	0.460	5.762
Sachsen-Anhalt	0.559	5.192
Saarland	0.629	3.200
Thueringen	0.671	7.494
Mecklenburg-Vorpommern	0.685	5.629
Bremen	1.670	5.516

Note: The dashed horizontal line denotes the distinction between states with low and states . with high levels of parliamentary funding. Data Sources: "Landeshaushalte" as published online: http://www.haushaltssteuerung.de/landeshaushalte.html

CONCLUSION: THE LIMITATIONS AND POSSIBILITIES OF (STUDYING) PARTISAN CAMPAIGN INFLUENCE

This thesis reports the first randomized trials of voter contact with party-affiliated field campaigns in the United Kingdom. Although the five papers included in this thesis cannot present a comprehensive picture of constituency campaigning to the reader, they advance our theoretical understanding of partisan campaigns, and are important from both a comparativist and a political perspective. In what follows I will outline the major conclusions from the experiments that have been conducted, and will reflect on how this research might lead to greater engagement between political practitioners and researchers.

Testing causal theories of voter decision-making

Research based on randomized field experiments is sometimes criticized because of its apparent lack of engagement with theory (Heckman and Smith 1995; Deaton 2010). Some would say that GOTV experiments have focused on "what works", i.e. testing the effects of different campaign interventions, at the expense of contributing to our understanding of the political world. This portrayal of experimental field research is, arguably, outdated. To give a few examples, research using randomized field experiments has made important contributions to our understanding of politics as a social instead of an individually rational process (Green et al. 2008; Nickerson 2008; Sinclair et al. 2012). It has contributed to our understanding of attitude in-stability (LaCour and Green 2014), and to the study of campaign communication and media effects (for reviews see Arceneaux 2010 and Green et al. 2014). By integrating insights about heuristic and social decision-making into the experimental campaign literature, this thesis formulates testable hypotheses about the direct and indirect effects of political cues on mobilization and persuasion.

The papers included in this thesis start from the assumption that it is difficult to understand how party identification or individual candidate preferences shape electoral-decision making if we do not account for the role of the political environment in structuring those identities and preferences (Kulinski et al. 2001). The political environment includes political institutions, but also political actors such as parties and candidates that communicate directly or indirectly with constituents (Mondak 1993). The first premise is therefore that campaigns voluntarily or involuntarily provide signals to voters that impose structure, and therefore constraints, on an individual's electoral decision-making process. The most important signal provided to voters during election campaigns is the party affiliation of the candidate. The party cue is so powerful because it relates to longstanding social stereotypes that voters hold about the typical representative of a party (Green et al. 2002). In order to understand partisan constraints on choice and behavior, my starting point is therefore not the partisan identification of the individual, but the signals provided to the individual by parties and candidates. Party is the most dominant, but not the only cue provided by election campaigns and used by voters to make sense of elections. Campaigns send many competing verbal and non-verbal signals, ranging from the appearance of the candidate, over demographic attributes to ideological references.

Theoretical contributions

Campaigns provide cues to voters, and competing theories make different predictions about how individuals will respond to the availability of such cues. Rational choice theory, which hypothesizes a direct causal pathway from candidate choice to turnout (Downs 1957), would for instance predict that party cues should help individuals choose between candidates, and hence increase turnout. This causal theory has clear observational

implications that are testable using field experimentation. I find that the provisions of party cues in a favorable, low-information environment does not cause an increase in turnout, not even among Labour supporters and unattached voters who should, and are in fact, receptive to campaign messages including such cues. Others have hypothesized that partisan messages should increase turnout by activating partisan identities (Rogers et al. 2012). This mechanism again, has clear observational implications, in that party cues should increase turnout among partisans, but not among unattached voters. I do not find any support for the identity-activation mechanism either. In contrast, the results are consistent with Zaller's (1992) acceptance-rejection axiom. In the first paper Eline de Rooij and I argue that a candidate's party affiliation works like an acceptance-rejection cue in field campaigns. Individuals make their decision of whether to follow or to reject a suggested behavior dependent on whether they disagree with the source. If the source is agreeable, they are happy to follow through with a suggested action. Unaffiliated GOTV tactics that have been deployed effectively in non-partisan contexts should hence have similar effects on turnout when employed by party-affiliated groups and targeted to party supporters. Another theory that is consistent with our findings assumes that partisan content results in partisan cross-pressure (Fiorina 1976) that negates the positive impact of the GOTV elements of the message. According to this interpretation of the results, partisans do not reject a campaign message, but the positive (GOTV) effects, and the negative (partisan) effects result in an observed null effect. We could further disentangle these two remaining possibilities by fielding a treatment that does not balance the GOTV and the partisan elements, but clearly privileges one over the other.

Second, manipulating the partisan intensity of campaign messages allows me to perform an indirect test of the controversial conflict-avoidance hypothesis. The conflict-avoidance mechanism is at the heart of the disagreement between scholars, some who view

the exposure to diverse opinions as beneficial for political participation (Huckfeldt et al. 2004), and others who think that partisan disagreement in heterogeneous networks inhibits participation (Mutz 2006). Mutz (2006) argues that individuals are conflict averse, and likely to withdraw from politics in order to avoid arguments with people they are close to. Again, the conflict avoidance hypothesis has clear observational implications, and I use an innovative intervention to subject this hypothesis to a fair test. The test is implicit because I assume that an openly partisan campaign call affiliated with the Labour Party results in a higher potential of partisan disagreement between a household member who supports the Labour Party and one who support the Conservative Party, than a message that does not mention parties or partisan politics. Following this strategy, I find that explicit partisan content that is favorable to one household member, but not to the other, does not negatively impact on mobilization within heterogeneous households. Instead, a campaign phone call from the Labour Party seems to increase indirect mobilization in households consisting of a Labour and a Conservative supporter.

Third, I investigate whether the provision of positive, non-party signals affects electoral decision-making among supporters of rival parties. Party is the principal, but not the only cue, that individuals use to arrive at electoral decisions (Lau and Redlawsk 2001). One often-discussed non-verbal cue is the appearance of the candidate. Some have said that individuals judge candidates within milliseconds based on their looks and facial features (Olivola and Todorow 2010). While some scholars argue that good looks help candidates get elected, others have hypothesized that "beauty is beastly", or that appearance might prime gender-related consideration, including gender stereotypes.

I find that providing supporters of rival parties with a photo of a rival parliamentary candidate does not seem to have any systematic positive or negative effect on how they evaluate her candidacy, or whether they are willing to vote for her. What seems to matter,

based on the preliminary results of this pilot-experiment, is that non-verbal cues match the verbal cues provided in a letter accompanying the photo. If the letter emphasizes that the candidate is a listener, the treatment is ineffective if it shows her talking on the photo.

The Role of Cues in Structuring Electoral Behavior

Following the Cue – The Mobilization of Core Supporters

The papers included in this thesis provide the first reliable causal evidence that a major party's telephone canvassing campaign mobilized supporters in the United Kingdom. Although we have good reasons to think that results from non-partisan GOTV studies would translate into partisan environments, the empirical evidence on the effectiveness of partisan GOTV campaigns from the United States (McNulty 2005; Cardy 2005; Michelson 2006; Bailey et al. 2013) from France (Pons 2014, Pons and Leigey 2013), and from Italy (Kendal et al. 2014) is at best mixed compared to what we would expect given the well-established positive findings from the non-partisan GOTV literature (Gerber and Green 2008; Green et al. 2013).

Nevertheless, the experiments conducted in two UK constituency campaigns suggest that direct and indirect campaign methods can be effective when targeted at supporters of the party that initiates contact. The Birmingham telephone experiment showed that Labour telephone calls significantly increased turnout amongst registered (non-postal) Labour and unattached voters in the 2012 Police and Crime Commissioner Election. Although based on a smaller sample, the Southampton door-hanger experiment provides some substantive evidence that Labour door-hangers increased turnout among Labour supporters, but not among unattached voters, in the 2014 European and Council elections. Moreover, the effects of campaign contact spill over to other household

members, and these indirect effects are as large as the direct effects that result from the personal interaction with a campaign volunteer.

"Not even a reply": The Constraints to Campaign Influence

The four campaign experiments included in this thesis are the first to demonstrate that supporters of rival parties (and to some extent unattached voters) are less responsive to the GOTV element of campaign messages than individuals who support the party that initiates contact. This result, established in the Birmingham experiment, has recently been replicated in the Southampton experiments using door-hangers and hand-written letters. While this finding makes intuitive sense, some theories suggest that campaign messages are likely to backfire and mobilize supporters of rival parties (Lupia and McCubbins 1998; Arceneaux and Kolodny 2009), or that they both persuade and mobilize those supporters. I do not find evidence for either mechanism. Moreover, by subtracting the entire partisan content from the campaign message, the Birmingham experiment provides evidence that these differences in responsiveness between partisan groups are not due to unobserved confounders, but due to an acceptance-rejection mechanism triggered by the partisan elements of a campaign message. The finding that out-partisans are reluctant to meet, and even to reply to a hand-written invitation sent by a parliamentary candidate might be surprising to many readers. It was surprising both to me and to the Labour candidate who wrote the letters. While I expected that most voters would not change their vote intentions after meeting the candidate, I expected that a substantial minority would take up the invitation to a meeting if the candidate were willing to invest substantial amounts of her time. Voters often complain that politicians are not willing to sit down and listen to their concerns. But even if they were, thinking in partisan categories seems to stand in the way of direct communication across the partisan divide. This result suggests that in the current

political environment candidates really do have problems communicating with constituents who do not initially support them.

Overcoming Partisan Bias – Direct versus Indirect Influence

This thesis reports a long list of failure to influence the behavior of out-party supporters: Telephone calls from Labour Party volunteers, and hand-written letters from a Labour Party candidate failed to increase turnout among rival party supporters, and rival partisans even failed to reply to letters inviting them to meet a Labour Party candidate. Nevertheless, the conclusion that campaigns therefore have to abandon the idea of reaching-out to voters who are not their natural constituents, and instead "go where the ducks are" as Bartels (1998) wrote, would be premature.

First, unattached voters and supporters of smaller rival parties seem more receptive to campaign communication than supporters of the main competitor. Unattached voters turn out in greater numbers as a result of Labour-affiliated GOTV phone calls, and supporters of smaller parties such as the LibDems, UKIP and the Green Party state that they are more likely to vote for the Labour candidate after they were contacted by mail. These indirect persuasion effects endured at least for a couple of weeks. Moreover, we would only reach the conclusion that campaign contact with rival party supporters was ineffective if we ignored the indirect effects, the backchannels, through which campaign messages reach and influence individuals who are initially opposed to the party.

The analysis of campaign spillover-effects makes it clear that rival party supporters are not immune to campaign interventions. Spillover-effects are especially pronounced in heterogeneous partisan households. Not surprisingly, individuals are more accepting of conflicting campaign messages if they are mediated by a person they trust than when they are presented by someone they do not know, and who belongs to a party they do not like.

The results of this thesis therefore suggest that mediation by household members can overcome partisan biases in the acceptance of campaign messages. While it undeniably structures election campaigns, partisanship is therefore not the be all and end all of what campaigns can achieve. To the average voter, politics is not everything, and personal relationships might in the end outweigh partisan differences.

Are these findings externally valid? In the classical sense of replication in different settings, this is a question that needs to be answered by future research (Morton and Williams 2010). However, if one subscribes to a view of external validity that is not purely predictive, but has also analytical elements (Garcia and Wantchekon 2010), then testing hypotheses derived from existing theories, confirming some, and rejecting others, improves our general understanding of campaign interactions in a partisan environment.

Towards a comparative perspective

Over the past few years, the experimental study of election campaigns has become increasingly comparative in its focus. As Green, McGrath and Aronow (2013) emphasize in their extensive review of GOTV experiments, randomized trials have been conducted in ten countries. Moreover, experiments in collaboration with political parties have been reported in five continents, Europe (France, Italy), North America (United States, Canada, Mexico), South America (Brazil), Africa (Benin, Nigeria) and Asia (Japan). However, the results of these experiments raise at least as many questions, as they are able to answer. One important question that is raised by the comparative literature is why large and well-powered partisan campaign experiments have so far failed to show any effect of partisan canvassing on electoral mobilization in European countries (Pons and Leigey 2013; Pons 2014; Kendal et al. 2014). There is also US-evidence that partisan direct mail is less effective at increasing turnout than non-partisan direct mail (Green et al. 2013). By

manipulating important elements of the political environment in which voters respond to campaign stimuli, this thesis might help to make sense of some of these unexpected findings.

Partisan campaigns do not intend to mobilize the entire electorate, but only those individuals who support them. But experiments that randomize contact at both the aggregate (Pons 2014; Pons and Liegey 2013; Kendall et al. 2014) and the individual level (Nickerson 2005) have problems identifying heterogeneous mobilization effects based on partisanship. At the aggregate level, null effects might mask that the intervention was effective with party supporters. This possibility is particularly pronounced in multi-party systems such as France and Italy. Maybe a majority of voters contacted by Pons (2014) and Kendall et al (2014) were supporters of smaller, rival parties that are open to persuasion, but not to mobilization? Due to data limitations, experiments conducted at the individual level might also run into the dangers of contacting the "wrong voters" (Nickerson 2005). The high-quality canvassing data on individual partisanship that was made available in all four experiments is hence a prerequisite for identifying individuals who should be responsive to partisan interventions. While this measure still has to be verified by third parties, it is strongly correlated with party ID results based on telephone surveys.

At the same time, experiments conducted in Europe have shown significant effects on the vote shares of the parties that initiated contact (Pons 2014; Kendal et al. 2014). The evidence from the Southampton experiment that hand-written letters persuaded supporters of smaller parties to vote for the Labour candidate is consistent with Pons' (2014) finding that canvassing increased the vote share of the French Socialist Party, and decreased the vote share of the Front National without increasing overall turnout.

It also appears to be the case that non-partisan door-to-door and telephone canvassing experiments conducted in the United Kingdom (John and Brennan 2008;

Fieldhouse et al. 2012) have produced larger treatment effects on turnout than GOTV experiments that have been conducted in the United States and in other European countries. The partisan experiments reported in this thesis are no exception to this rule, and might lead us to wonder if we are still drawing samples from the upper end of the distribution, or if these results might really be a function of the British electoral context. Since studies have yet to field the same treatments in different countries, a direct comparison of the effects between the UK, the US, France and Italy remains difficult (Green et al. 2013).

Moreover, the results from the Birmingham and Southampton experiments also contribute to our comparative understanding of impersonal campaign techniques. The Birmingham text message experiment reports the first null finding from a randomized GOTV campaign based on text messages. The null finding is the more interesting, given that the telephone GOTV experiment that I conducted on a similar sample, with the same campaign, in the same election, proved highly effective. How can we make sense of these findings, especially given that another recent text message experiment conducted by Bhatti et al. (2014) in Sweden reports a statistically significant 2 percentage-points increase in turnout? These results do not necessarily contradict each other. According to Malhotra et al. (2011), impersonal, but noticeable GOTV methods rely on a previous intention to participate. Bhatti et al (2014) operated in a high turnout election with a baseline turnout rate of 60% in the control group. This almost matches the turnout rate in the control group reported by Dale and Strauss (2009) who find a treatment effect of 3 percentage-points.

These results speak to the importance of the electoral context, the political environment, as an explanation for why similar methods succeed in some contexts, but fail in others. However, since we are effectively looking at a treatment-by-covariate interaction, we still do not know if the baseline turnout rate really affects the effectiveness

of non-partisan mobilization methods. One way of approaching this question within the experimental framework would be to rely on downstream experimentation, and test if subjects that were randomly assigned to an effective GOTV treatment in a previous experiment are more likely to respond to an impersonal GOTV treatment in the next election than subjects who were assigned to the control group. We know from previous research that voting in one election increases the likelihood of voting in the future (Gerber et al. 2003; Cutts et al. 2009). We could hence exploit this insight to identify the causal interaction effect between turnout propensity, and impersonal mobilization methods.

The comparative experimental study of partisan campaigns is still in an early phase. We need to take variations in the political environment, both within and between countries, seriously when replicating the results of earlier campaign studies. A research program that does not only rely on examining treatment-by-covariate interactions, but that tries to manipulate both the campaign intervention and the conditioning variable of interest, would be a promising way forward. A future research program that tries to answer under which conditions campaigns can mobilize and persuade voters, might also be an opportunity to combine naturally occurring experiments, such as the lowering of the voting age in the Scottish independence referendum to 16, with fully randomized campaign interventions.

The Limitations and Possibilities of Embedded Experimentation

I am not the first researcher in Britain to partner with the Labour Party in order to test the effects of constituency campaigns in the field. Unfortunately, Bochel and Denver's (1972) canvassing study has remained the only imbedded field experiment in Britain for more than 40 years. The divorce between theory and practice, between knowledge of the scientific method and real-world experience, has been an unfortunate development for political practitioners and for researchers studying election campaigns alike. Political

practitioners have for too long relied on received wisdom or on unreliable heuristics in judging "what works" in election campaigns. Problems such as selecting on the dependent variable (we won, hence the campaign must have been effective), or path dependency (we do something because we have always done it) are commonplace. Even if campaigns use data to judge the effectiveness of their canvassing, they tend to make biased comparisons. It is not uncommon for local parties to compare turnout rates among constituents who were contacted to turnout rates among constituents who were not contacted. This biased comparison often leads practitioners to have an exaggerated idea about the effectiveness of their canvassing or literature drops. However, they are not alone in making biased comparisons. Political Scientists have long studied campaign interventions using observational survey data. Although they might have used more sophisticated methods to "control" for observed confounders, their models are very similar to the biased difference-in-proportions estimator that campaign managers might scribble on the back of a leaflet. Moreover, Political Scientists are interested in testing causal theories about election campaigns. But without manipulating the very campaign intervention that we intend to study, we cannot be confident about the causal and ecological validity[1] of the causal inference made.

While this unsatisfying situation has been the case for too long, change has happened over the last decade. Campaign analytics has played an increasing role in the United States in the past two Presidential campaign cycles (Issenberg 2012). This data-driven revolution has not gone unnoticed in the UK. Many candidates and campaign organizers have heard about the GOTV experiments conducted by Gerber and Green (2008) in the United States and the influence they have had on how parties' field operations are conducted. Campaigns have a limited budget, both in money and time, and

[1] Morton and Williams (2010) break internal validity down into construct, causal and ecological validity.

have an inherent interest in using these resources more effectively. They are also interested in understanding how voters 'tick' and how to best engage with them. There is hence a clear incentive for organizers and candidates to collaborate with researchers.

Scientists, if they are willing to leave their offices, and engage with the challenging world of campaign politics, have a lot to offer to campaigns. They know the scientific methods that are increasingly in demand. While in the United States, campaigns for state-wide or national office now often employ in-house analysts and behavioral scientists (Issenberg 2012), this development has been more limited in the United Kingdom. Researchers have therefore a lot to offer to political practitioners. In return for evaluating a campaign intervention that is happening anyway, researchers might be allowed to manipulate elements of this intervention in order to test causal theories of voter decision-making. Moreover, sometimes candidates would like to try new methods of campaigning, but there is too much uncertainty about the costs and benefits of reallocating resources. Campaigns in the UK are not afloat with money, and are facing real trade-offs. Does a campaign rather invest in direct mail, or in door-hangers? Is the time of a candidate most effectively spent on the door-step, and talking to voters directly, or is he or she no more effective at mobilizing and persuading voters than other party volunteers? In such situations, where trade-offs are real and there is a lot of uncertainty, campaigns are inherently risk-averse. They would often emphasize the uncertainty surrounding the effectiveness of a new method, but they would not recognize that the "tried and tested" method, the status quo, has been tried, but not tested. In reality, the status quo is often surrounded by as much uncertainty than the new method. This is where the researcher comes in. Sometimes campaigns can be convinced to test new methods in randomized pilot projects before agreeing to deploy them on a large scale. Experimentalists are therefore usually on the side of innovators, and of those who question the status quo. During pilot

projects, the researcher has often more leverage in designing a campaign intervention in collaboration with the candidate or campaign manager. Interventions that result are often informative both from a theoretical and a practical point of view.

Collaborations between researchers and practitioners moreover have the advantage that they are highly cost-effective for both sides. Campaigns that are run on a budget save the money of hiring expensive consultants and in-house analysts, while researchers, and those who provide funds to researchers, save the money of implementing their own interventions. Critics have often objected to field experimentation because of its high costs (Deaton 2010). Moreover, the large grants that are necessary to run a field experiment without the collaboration of political practitioners often unfairly restrict entrance into field experimentation and make the methods unfeasible for junior researchers. The only costs I incurred conducting multiple field experiments over the past three years were travel costs and expenses.

Of course embedded experimentation has its drawbacks for both researchers and practitioners. Researchers often need to compromise with campaigns about interventions, and they do not have full control over the implementation of the experiment. While the former problem can be overcome by creativity and good experimental design, the latter is an important drawback that results in too many failed experiments. A failed experiment can cause frustration for both partners, and might be especially harmful to a doctoral student whose thesis is dependent on the successful execution of an experiment. Researchers who conduct embedded experiments have therefore invested much thought into how this situation can be improved for both parties. Although evidence is based on observational data alone, concise written agreements, field protocols and records, and the possibility that the researcher is allowed to train volunteers and supervise the

implementation of the experiment, all seem to be crucial ingredients of successful implementation.

Practitioners, on the other side, often need to assign control groups, invest additional resources to implement an experiment, and live with the fact that results will often not be instantly available, and that the experimental method might cause public controversy. The stakes of experimentation are therefore usually higher for practitioners than for academics, and researchers need to be aware of these constraints. As Loewen and colleagues (2010) suggest, researchers should therefore think about interventions that can deliver timely results. It might also be preferable to experiment "in cycle", that means in the period between two first order elections. Moreover, following the public outcry over the 'Facebook experiment', we need to have a public debate about whether it is preferable if organizations and firms continue treating individuals without knowing the effects of that treatment, or if it might not be preferable that treatment effects on human subjects are known, and published in the public sphere (Watts 2014). All these problems, and trade-offs are real, but they need to be weighed against the low financial costs of embedded experimentation, and against the gain in ecological validity that is only possible when manipulating the behavior of real political actors.

Is it ethical from the perspective of the researcher to conduct experiments in partnerships with organizations that have clear partisan goals? In my understanding, the answer to this question will be context-dependent. It depends on the specific intervention that is being studied, the nature of the organization that a researcher collaborates with, and the distinct advantages that embedded experimentation provides over other alternative methods of inquiry. In general, the transparency of experimentation, which is increased through the participation of researchers, should not be underestimated. It is in the public interest that experimental results are in the public domain to be studied by all interested

parties instead of being conducted by parties in-house away from public scrutiny, and with no public benefits. The advantages of experimentation for causal inference are clear. The incentives are great for political parties. The most important question then is if we will continue the path that we have so far followed in Britain, the path of making results available to the public, or if we go down the American path where analysts who are employed by campaigns run most randomized trials behind closed doors. I think it is clear which option is more beneficial to society, and to Political Science.

For embedded experiments to function, expectations from both sides should be clearly defined and articulated, and the conduct has to be professional and respectful (Peyton 2012). Life and research is a repeated game, and particularly so when working with large organizations.

In conducting experiments, researchers should think about other researchers who might come after them. It is important that experiments open doors to future collaboration, and do not close doors for other researchers who might be interested in following suit. It looks like the experiments that I conducted, some of them in collaboration with colleagues, have only been the beginning of campaign experimentation in Britain. Some experiments that I report in this thesis are pilot projects based on small samples. Some larger experiments will follow, and others have already followed. Labour candidates are no longer the only political practitioners who are willing to conduct experiments. A campaign affiliated with the Conservative Party welcomed researchers and just conducted the first campaign experiment in Somerset during the 2014 European elections. The experimental frontier is wide open. It is now on researchers and practitioners to make use of the great possibilities and mutual benefits that research collaboration provides.

References

Arceneaux, Kevin. 2010. “The Benefits of Experimental Methods for the Study of Campaign Effects.” *Political Communication* 27(2): 199-215.

Arceneaux, Kevin, and Robin Kolodny. 2009. “Educating the Least Informed: Group Endorsements in a Grassroots Campaign.” *American Journal of Political Science* 53(4): 755–770.

Bailey, Michael A., Daniel J. Hopkins and Todd Rogers. 2013. “Unresponsive, Unpersuaded: The Unintended Consequences of Voter Persuasion Efforts”, http://scholar.harvard.edu/files/todd_rogers/files/unresponsive.pdf

Bartels, Larry M. (1998). Where The Ducks Are: Voting Power in a Party System. In Geer, John G. (ed). *Politicians and Party Politics.* Baltimore: John Hopkins University Press: pp. 43-79.

Bhatti, Yosef, Jens Olav Dahlgaard, Jonas H. Hansen, and Kasper M. Hansen. 2014. “How voter mobilization spread in households and families – The use of short text messages on cell phones to boost turnout.” Paper presented at the Annual Meeting of the American Political Science Association in Washington, DC.

Bochel, John M., and David T. Denver. 1971. "Canvassing, turnout and party support: an experiment." *British Journal of Political Science* 1(3) 257-269.

Brown, Steven D., Andrea ML Perrella, and Barry J. Kay. 2010. "Revisiting Local Campaign Effects: An Experiment Involving Literature Mail Drops in the 2007 Ontario Election." *Canadian Journal of Political Science* 43(1): 49-67.

Cardy, Emily Arthur. 2005. "An Experimental Field Study of the GOTV and Persuasion Effects of Partisan Direct Mail and Phone Calls." In Donald P. Green and Alan S. Gerber (eds). The Science of Voter Mobilization. *The Annals of the American Academy of Political and Social Science* 601: 28-40.

Cutts, David, Edward Fieldhouse, and Peter John. 2009. "Is voting habit forming? The longitudinal impact of a GOTV campaign in the UK." *Journal of Elections, Public Opinion and Parties* 19(3): 251-263.

Dale, Allison, and Aaron Strauss. 2009. "Don't Forget to Vote: Text Message Reminders as a Mobilization Tool." *American Journal of Political Science* 53: 787–804.

Deaton, Angus S. 2010. "Instruments, randomization, and learning about development." *Journal of economic literature*: 424-455.

Downs, Anthony. 1957. "An Economic Theory of *Democracy*. New York: Harper and Row.

Druckman, James N., Donald P. Green, James H. Kuklinski, and Arthur Lupia, 2006. "The growth and development of experimental research in political science." *American Political Science Review* 100(4): 627-635.

Fiorina, Morris P. 1976. "The voting decision: instrumental and expressive aspects." *The Journal of Politics* 38(2): 390-413.

Ford, Robert, and Matthew J. Goodwin. 2014: *Revolt on the Right: Explaining Support for the Radical Right in Britain*. London: Routledge.

Garcia, Fernando Martel, and Leonard Wantchekon. 2010. "Theory, external validity, and experimental inference: Some conjectures." *The ANNALS of the American Academy of Political and Social Science* 628(1): 132-147.

Gerber, Alan S., Donald P. Green, and Ron Shachar. 2003. "Voting may be habit-forming: evidence from a randomized field experiment." *American Journal of Political Science* 47(3): 540-550.

Green, Donald P., and Alan S. Gerber. 2008. *Get out the vote: How to increase voter turnout*. Brookings Institution Press.

Green, Donald P., Peter M. Aronow, and Mary C. McGrath. 2013. “Field Experiments and the Study of Voter Turnout.” *Journal of Elections, Public Opinion & Parties*: published online.

Green, Donald P., Brian R. Calfano, and Peter M. Aronow. 2014. Field Experimental Designs for the Study of Media Effects. *Political Communication* 31(1): 168

Heckman, James J., and Jeffrey A. Smith. 1995. “Assessing the Case for Social Experiments”. *Journal of Economic Perspectives* 9(2): 85-110.

Huckfeldt, Robert R., Paul E. Johnson, and John D. Sprague. 2004. *Political disagreement: The survival of diverse opinions within communication networks*. Cambridge: Cambridge University Press.

Issenberg, Sasha. 2012. *The Victory Lab: The Secret Science of Winning Campaigns.* New York: Crown Publishers.

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(3): 565-574.

Kendall, Chad, Tomasso Nannicini and Francesco Trebi. 2014. How Do Voters Respond to Information? Evidence from a Randomized Campaign. Working paper.

Kuklinski, James H., Paul J. Quirk, Jennifer Jerit and Robert F. Rich. 2001: "The political environment and citizen competence." *American Journal of Political Science* 45(2): 410-424.

LaCour, Michael J., and Donald P. Green. 2014. "Messages, Messengers, and Diffusion of Support for Gay Equality: Results from Two Longitudinal Field Experiments." *Working paper.*

Lau, Richard R., and David P. Redlawsk. 2011. “Advantages and Disadvantages of Cognitive Heuristics in Political Decision Making." *American Journal of Political Science* 45(4): 951-971.

Loewen, Peter John, Daniel Rubenson and Leonard Wantchekon. 2010. “Help Me Help You: Conducting Field Experiments with Political Elites.” *The Annals of the American Academy of Political and Social Science* 628(1): 165-175.

Lupia, Arthur and McCubbins Matthew. 1998. *The Democratic Dilemma*. Cambridge: Cambridge University Press.

Malhotra, Neil, Melissa R. Michelson, Todd Rogers, and Ali A. Valenzuela. 2011. “Text Messaging as Mobilization Tools: The Conditional Effect of Habitual Voting and Election Salience.” *American Politics Research* 39: 664-681.

McNulty, John E. 2005. "Phone-Based GOTV - What's on the Line? Field Experiments with Varied Partisan Components, 2002-2003." In Donald P. Green and Alan S. Gerber (eds). The Science of Voter Mobilization. Special Editors: *The Annals of the American Academy of Political and Social Science* 601: 41-65.

Michelson, Melissa R. 2005. “Meeting the Challenge of Latino Voter Mobilization”.The ANNALS of the American Academy of Political and Social Science 601(1): 85-101.

Mondak, Jeffery J. 1993. "Source cues and policy approval: The cognitive dynamics of public support for the Reagan agenda." *American Journal of Political Science*: 186-212.

Morton, Rebecca B., and Kenneth C. Williams. 2010. *Experimental political science and the study of causality: From nature to the lab*. Cambridge: Cambridge University Press.

Mutz, Diana C. 2006. *Hearing the Other Side: Deliberative versus Participatory Democracy*. Cambridge: Cambridge University Press.

Nickerson, David W. 2005. "Partisan mobilization using volunteer phone banks and door hangers." *The Annals of the American Academy of Political and Social Science* 601(1):

10-27.

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): 85-97.

Nickerson, David W. 2008. “Is Voting Contagious? Evidence from Two Field Experiments.” *American Political Science Review* 102(1): 49–57.

Olivola, Christopher Y., and Alexander Todorov. 2010. "Elected in 100 milliseconds: Appearance-based trait inferences and voting." *Journal of Nonverbal Behavior* 34(2): 83-110.

Peyton, Kyle. 2012. Ethics and Politics in Field Experiments. The Experimental Political Scientist. Newsletter of the APSA Experimental Section 3(1): 20-37.

Pons, Vincent. 2014. Does Door to door canvassing affect vote shares: Evidence from a Countrywide Field Experiment in France. Working paper: http://economics.mit.edu/files/9000.

Pons, Vincent and Guillaume Liegey. 2013. Increasing the Electoral Participation of Immigrants. Experimental Evidence from France. Working paper: http://economics.mit.edu/files/8885.

Rogers, Todd, Alan S. Gerber, and Craig R. Fox. 2012. Rethinking Why People Vote: Voting as Dynamic Social Expression. In Eldar Shafir (Ed.), *Behavioral Foundations of Policy* (pp. 91-107). Princeton: Princeton University Press.

Sinclair, Betsy, Margaret McConnell, and Donald P. Green. 2012. “Detecting Spillover Effects: Design and Analysis of Multi-level Experiments.” *American Journal of Political Science* 56: 1055-1069.

Watts, Duncan J. 2014. Stop complaining about the Facebook study. It's a golden age for research. The Guardian, 7 July 2014:

http://www.theguardian.com/commentisfree/2014/jul/07/facebook-study-science-experiment-research

Zaller, John. 1992. *The nature and origins of mass opinion*. Cambridge: Cambridge University Press.